JMIR Research Protocols

Impact Factor (2022): 1.7
Volume 11 (2022), Issue 3 ISSN 1929-0748 Editor in Chief: Xiaomeng (Simone) Ma, PhDc, MS, BS

Contents

Protocols

Original Paper

Corrigenda and Addendas

Protocol

Home Telemonitoring of Arterial Hypertension With Antihypertensive Treatment Titration: Protocol for a Randomized Controlled Prospective Trial (HOROSCOPE Study)

Sonia Ben Hafaiedh[1*], MD; Yosra Ben Daya[1*], MD; Amina Hadjer Radoui[1*], BSc, MSc; Mohamed Bouchoucha[1*], MD; Rabie Razgallah[2*], MD; Semir Nouira[3*], MD[‡]

[1]Medical Department, MédiS Laboratories, Tunis, Tunisia

[2]Dacima, Tunis, Tunisia

[3]Emergency Department and Laboratory Research (LR12SP18), Fattouma Bourguiba University Hospital, Monastir, Tunisia

[‡]Horoscope study group

[*]all authors contributed equally

Corresponding Author:
Semir Nouira, MD
Emergency Department and Laboratory Research (LR12SP18)
Fattouma Bourguiba University Hospital
Rue du 1ier juin 1955, Fattouma Bourguiba University Hospital
Monastir, 5000
Tunisia
Phone: 216 73106046
Email: semir.nouira@rns.tn

Abstract

Background: Despite the availability of effective treatment, the control of hypertension remains insufficient. Telemonitoring in the management of hypertension would be an effective way to improve blood pressure control.

Objective: The aim of our study will be to evaluate the effects of telemonitoring with antihypertensive treatment titration on blood pressure control in Tunisian patients with hypertension.

Methods: Our trial will be a prospective, rater-blinded randomized controlled trial carried out with primary care physicians in the Sahel region of Tunisia. Patients will be eligible for enrollment if they are aged over 35 years, are newly diagnosed with hypertension, or are known to be poorly controlled on antihypertensive therapy. Participants will be randomly assigned in a 1:1 ratio to the telemonitoring arm or usual care arm. The telemonitoring arm will involve a weekly telephone call for the collection of the home blood pressure measurements, therapeutic education, and treatment compliance assessment as well as a monthly call for treatment titration and a side effect check. Randomization will be done via the use of an interactive web responsive system, and patients will be stratified by investigation center. Neither participants nor investigators will be masked to the group assignments. The primary outcome will be the change in mean 24-hour systolic blood pressure from baseline to the 6-month follow-up in the 2 groups. All randomized patients who attend the follow-up visit at 6 months and have no missing data for the primary outcome will be included in the analysis.

Results: Recruitment to the trial started in July 2020. The study was initiated with 17 primary care physicians. We expect the inclusion period to last for approximately 6 months. We expect to complete data collection by the end of 2021 and plan to disseminate the results subsequently.

Conclusions: The HOROSCOPE (Home Telemonitoring of Arterial Hypertension With Antihypertensive Treatment Titration: Randomized Controlled Prospective Trial) study will provide important new evidence that could shed some light on the feasibility and impact of telemonitoring and self-monitoring in a Tunisian population of patients with hypertension who consult primary care physicians.

Trial Registration: ClinicalTrials.gov NCT04607239; https://clinicaltrials.gov/ct2/show/NCT04607239

International Registered Report Identifier (IRRID): DERR1-10.2196/26184

(JMIR Res Protoc 2022;11(3):e26184) doi:10.2196/26184

KEYWORDS

telemonitoring; arterial hypertension; primary care; ambulatory blood pressure monitoring; randomized controlled trial

Introduction

Background

Hypertension is one of the main risk factors for cardiovascular disease and the leading cause for global morbidity and mortality [1,2]. Despite the availability of effective treatment, the control of hypertension remains insufficient [3]. It has been suggested that self-monitoring in the management of hypertension is an effective method of improving blood pressure control [4,5]. Nevertheless, isolated self-monitoring has not been associated with better blood pressure control; it would be more beneficial if it was used in combination with other interventions, such as integrating primary care physicians in a telemonitoring approach. Despite conflicting evidence, a Scottish study involving general practitioners reported a significant reduction in blood pressure over a period of 6 months [6]. Furthermore, telemonitoring could help physicians with antihypertensive treatment titration [7], as shown in another trial (TASMINH2 [Telemonitoring and/or Self-monitoring of Blood Pressure in Hypertension 2]), which demonstrated that the telemonitoring with home titration group had lower blood pressure than that of the usual care (control) group after 12 months [8]. Moreover, a recent study (TASMINH4 [Telemonitoring and/or Self-monitoring of Blood Pressure in Hypertension 4]) aimed to determine whether telemonitoring resulted in a lower blood pressure by comparing it to usual care (control) and self-monitoring only [9]. The trial was a parallel randomized controlled trial that was conducted with 142 general practitioners in the United Kingdom and included patients who were hypertensive, aged over 35 years, had a blood pressure of >140/90 mm Hg, and were willing to self-monitor their blood pressure. A total of 1182 patients were randomly assigned to blood pressure self-monitoring (self-monitoring group: n=395), blood pressure self-monitoring with telemonitoring (telemonitoring group: n=393), or usual care (control group: n=394). After 12 months, patients' systolic blood pressure (SBP) was lower in both the self-monitoring and telemonitoring groups compared to that of the usual care group (mean 137, SD 16.7 mm Hg and mean 136, SD 16.1 mm Hg vs mean 140.4, SD 16.5 mm Hg, respectively). No statistically significant differences between the self-monitoring and telemonitoring groups were recorded. This study showed that self-monitoring, with or without telemonitoring, can be used to guide the titration of antihypertensive medicines in primary care. The observed drops in blood pressure in this study indicated an approximate 20% and 10% reduction in the risk of stroke and coronary heart disease, respectively [9]. The results of this study deserve to be confirmed, especially since this study's approach can be economically interesting [10] in the context of a resource-limited country with failing health economics. It would be of great use to evaluate this approach with general practitioners in Tunisia—a country in which the prevalence of hypertension in urban areas is significantly greater than that in rural areas (P<.001) and only 38.1% of patients with hypertension are aware of their hypertension [11].

Research Questions

The aim of the HOROSCOPE (Home Telemonitoring of Arterial Hypertension With Antihypertensive Treatment Titration: Randomized Controlled Prospective Trial) study is to demonstrate the impact of telemonitoring on blood pressure control among patients with hypertension.

The trial will address the following three main research questions: (1) does telemonitoring improve treatment compliance and blood pressure control in newly diagnosed or poorly controlled patients with hypertension, (2) does telemonitoring improve patients' quality of life, and (3) does telemonitoring reduce the incidence of cardiovascular complications?

Methods

Study Design

Our trial will be a multicentric, rater-blinded, prospective randomized controlled trial. The allocation ratio between the intervention and control group will be 1:1. Randomization will be done via the use of an interactive web responsive system, and patients will be stratified by investigation center. Neither participants nor investigators will be masked to the group assignments. However, the cardiologist, who will interpret the results of the primary outcome, will be blinded.

The study is expected to last for an overall duration of 12 months (from the first patient who is enrolled to the last patient who completes the trial). The recruitment period will last for 6 months, and the follow-up interval will be 6 months.

Participants

Patients with hypertension will be recruited from consultations with primary care general practitioners in the central region of Tunisia. The inclusion and exclusion criteria are outlined in Textbox 1.

Textbox 1. Inclusion and exclusion criteria for the HOROSCOPE (Home Telemonitoring of Arterial Hypertension With Antihypertensive Treatment Titration: Randomized Controlled Prospective Trial) study.

Inclusion criteria

- Aged ≥35 years old
- Newly diagnosed hypertension
- Poorly controlled hypertension
  - One single blood pressure value outside the expected range will not be taken as a definitive indicator of clinical deterioration
  - Blood pressure should be >140/90 mm Hg at screening for at least 2 visits to allow the inclusion of a patient

Exclusion criteria

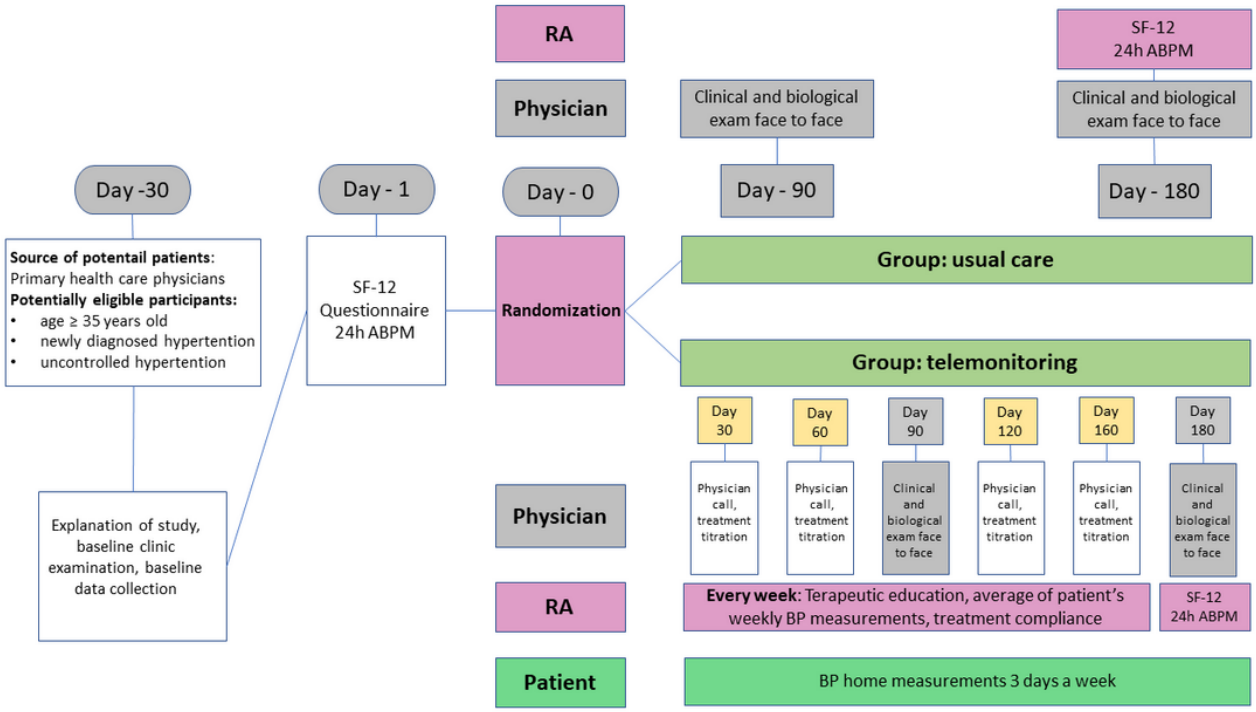
- Orthostatic hypotension
- Chronic renal failure (serum creatinine level: >200 μmol/L)
- Acute coronary syndrome, coronary revascularization, or stroke within the past 3 months
- Known secondary causes of hypertension
- Pregnancy
- New York Heart Association Class III or IV heart failure or left ventricular ejection fraction of <30%
- Dementia or any other cause that could prevent the implementation of remote monitoring (telemonitoring)

Recruitment, Screening, and Informed Consent Procedures

Participants will be recruited directly from primary health care centers by general practitioners. Participants will be included during their visit to the outpatient clinic after the verification of their eligibility criteria. When patients meet the eligibility criteria, the investigator will provide them with a study information sheet. Patients showing interest in taking part in the study will be included.

Material compensation will not be offered for participating in the study. Potential participants will be informed about the scientific benefits of their participation and will not be offered any other incentives. In order to obtain informed consent, the investigator will inform the participants about the possible benefits and potential side effects of participating in this study as well as the study duration. Patients will also be informed that their participation will be strictly voluntary and that they can withdraw from the study at any time without stating any reason. They will also be informed that withdrawing from the study will not affect subsequent medical assistance and treatment. Participants will be advised that their personal data will be used by other authorized health care professionals other than their treating physicians as part of the clinical trial. They will be given sufficient time to reflect extensively on their participation decision. The informed consent form will be signed and dated by the investigator following the signature of the participant. The consent form will be retained as part of the study records. A signed copy will be provided to each study participant. Overall patient flow throughout the study is depicted in Figure 1.

Figure 1. Participant flow throughout the study. ABPM: ambulatory blood pressure monitoring; BP: blood pressure; CRA: clinical research associate; SF-12: 12-Item Short Form Survey.



Sample Size

A sample size of 199 subjects per group is required for 80% power, assuming a mean difference in the reduction of SBP of at least 5 mm Hg between the two study groups and an SD of 15 mm Hg. Considering a patient dropout ratio of 15%, the study will need to recruit 230 subjects in each group.

Baseline and Follow-up Clinic Examinations

Patients will undergo an initial baseline clinic examination, during which clinical measurements, demographic data, and past medical histories will be collected and electrocardiogram exams and biological examinations will be conducted.

A validated blood pressure monitor (ROSSMAX X3; Rossmax Swiss GmbH) will be used for blood pressure measurements. After 1 month, a satisfaction questionnaire—the Medical Outcomes Study Short Form-12 questionnaire [12]—will be completed by participants, so that we can assess their quality of life.

Twenty-four–hour ambulatory blood pressure monitoring (ABPM) will be performed for all included patients to analyze their blood pressure profiles through the use of a Holter monitor (Contec ABPM 50; Contec Medical Systems). Ambulatory blood pressure measurements will be scheduled to acquire a blood pressure reading every 30 minutes from 6 AM to 10 PM and every 60 minutes from 10 PM to 6 AM. The cuff will be placed on participants' nondominant arms. Of the 48 possible measurements, at least 32 (67%) must be available. Otherwise, ABPM will be repeated. The Holter monitor must not get wet (eg, via a swimming pool, bath, or shower), and usual activities must be maintained. After installing the Holter monitor, study participants can return to their usual activities. They will be invited to return to the clinic 24 hours later, so that we can collect the data from their monitors. The results will be interpreted by a cardiologist.

Randomization will take place 1 month after participants' inclusion, that is, after the collection of at least 32 interpretable ABPM measurements. All patients will be asked to attend 2 follow-up clinics within 6 months. The clinical follow-up data to be collected are detailed in Table 1.

Table 1. Timetable for patient data collection throughout the trial.

Collected data	Baseline (inclusion visit)	After 29 days (1 day prior to the randomization visit)	After 3 months (follow-up visit 1)	After 6 months (follow-up visit 2)
Sociodemographics	✓			
Past medical history	✓			
Clinical examination[a]	✓	✓	✓	✓
Biological tests[b]	✓		✓	✓
ECG[c]	✓	✓		✓
SF-12[d]		✓		✓
ABPM[e]		✓		✓
Treatment compliance		✓	✓	✓
Side effects		✓	✓	✓

[a]The clinical examination will mainly include the assessment of blood pressure (systolic and diastolic blood pressure), BMI, and heart failure New York Heart Association stage.

[b]Biological tests will include those for serum creatinine, glycemia, lipid profiles, and brain natriuretic peptides.

[c]ECG: electrocardiogram.

[d]SF-12: 12-Item Short Form Survey.

[e]ABPM: ambulatory blood pressure monitoring.

Randomization

Patients who meet the eligibility criteria will be randomly assigned to either of the study groups in a 1:1 ratio via the use of an interactive web responsive system.

The study groups are as follows:

- Usual care group (control group): patients in this group will receive their antihypertensive treatment according to the usual approach during each consultation appointment made with the attending physicians.
- Telemonitoring group (intervention group): patients in this group will receive their usual care along with at-home self-monitoring, telemonitoring, and the titration of antihypertensive treatment.

Intervention

Control Group

Participants randomized to the control group will receive usual care (ie, physician visits every 3-6 months and therapy according to standard treatment guidelines). In contrast to the intervention group, control group participants will not receive phone calls for therapeutic education, a treatment compliance assessment, treatment titration, or a side effect check. In cases of any worsening symptoms, patients will have to contact their general physician individually and independently.

Intervention Group

Participants randomized to the intervention group will receive therapeutic education, including information on diet, sleep, physical activity, and other lifestyle advice, from the research associate. Patients will also be provided with a validated blood pressure monitor (ROSSMAX X3), and they will be taught how to use the blood pressure monitor at home. Patients will be asked to measure their blood pressure 3 times in the morning and 3 times in the evening for 3 days per week.

Once per week, the research associate will call the patients to check their treatment compliance and verbally collect the blood pressure measurements. Compliance evaluation will be conducted by asking the patients whether they have been taking their medication correctly every day on time for the past week. Patients will be asked to keep their empty treatment boxes. The remaining pills will be counted, and patient compliance will be assessed. The research associate will also remind patients about lifestyle and dietary measures. Subsequently, the research assistant will fill in the data entry fields in the electronic platform, which can be checked and validated by the investigators. Patients will be contacted every month via telephone by their attending physician for 6 months to titrate the antihypertensive treatment according to the data collected by the research associate. At day 90 and day 180, the medical visit will be conducted face-to-face. For the titration of antihypertensive medicines, the decision criterion will be the percentage of home blood pressure readings that meet the target. If at least 75% (375/500) of the readings since the last visit meet the blood pressure target, no changes in treatment will be considered. If less than 75% (375/500) of the readings meet the target, we will recommend readjusting the treatment. During the telephone follow-up calls, physicians will encourage participants to adhere to the treatment. Regardless of blood pressure control, if patients experience side effects, the involved antihypertensive drug will be stopped or changed, or its dose will be reduced.

Target Blood Pressure

The target blood pressure will be based on the 2018 European Society of Cardiology and European Society of Hypertension Guidelines [13].

The home blood pressure target will be ≤135/85 mm Hg, and the general practitioners' office blood pressure target will be ≤140/90 mm Hg. The mean ambulatory blood pressure targets will be ≤135/85 mm Hg at daytime, ≤120/70 mm Hg at nighttime, and ≤130/80 mm Hg for 24 hours.

Statistical Analysis

Categorical variables will be expressed as frequencies and percentages, while continuous variables will be expressed as means with SDs. An analysis of variance will be performed to compare the two study groups in terms of the primary end point. The normality of continuous data will be assessed with the Shapiro-Wilk test. Furthermore, we will use a general linear model for repeated measures to compare the marginal means of Gaussian quantitative dependent variables before and after the intervention (a Bonferroni test will be considered). A nonparametric test (the Friedman test) will be used for the comparison of non-Gaussian quantitative variables before and after intervention. The Cochran Q test will be performed to compare binary dependent variables before and after intervention. Subgroup analyses will be performed by age (<65 years and ≥65 years), sex, and the severity of hypertension. Differences between continuous variables will be evaluated with a 2-tailed Student *t* test or by its correction, as appropriate. The differences among the categorical variables will be evaluated with the Pearson chi-square test. All tests will be considered significant at the cutoff value of $P<.05$.

Data will be collected via an electronic data capture system (Dacima Clinical Suite; Dacima Software Inc) that is compliant with international data security requirements (ie, those of the US Food and Drug Administration [FDA] Title 21 of the Code of Federal Regulation [CFR], part 11; Health Insurance Portability and Accountability Act [HIPAA]; International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use [ICH]; and General Data Protection Regulation [GDPR]). Dacima Clinical Suite is authorized by the local body of personal data protection. The electronic case report form will be used to perform quality checks on the data (checks on simple controls [eg, range check], eligibility criteria, and complex controls) to increase the accuracy of the data. Double data entry will not be scheduled, as data will be directly captured in the electronic data system. Investigators and research assistants will use monitoring data tools as an audit trail for tracking the data modification log, data queries, and electronic signatures. Research assistants will use source document verification tools to ensure the accuracy of the collected data. Automatic alerts will be implemented to notify the responsible investigator about any data or protocol violations.

Participants who do not have data on the primary outcome at the 6-month follow-up assessment will be labeled as those who are missing data for the primary outcome at this time point. A further sensitivity analysis will include participants with 1 missing measurement. The frequency (percentage) of losses to follow-up (defaulters and withdrawals) after 6 months will be reported by randomized group and compared between the two groups. The availability of the outcome data for the primary and secondary outcomes will be summarized for the two randomized groups. The mixed-effects model implicitly accounts for data missing at random; however, the missing data mechanism will be explored. A logistic regression model will be used to explore any associations between baseline characteristics and the availability of data on the primary outcome. Any changes to the assumptions made in the primary analysis (ie, data missing at random) will be considered in a sensitivity analysis.

Outcomes

Primary Outcome

Our primary outcome will be the percentage of patients who exhibit a change of 10 mm Hg in average 24-hour SBP from baseline to the 6-month follow-up.

Secondary Outcomes

The following will be our secondary outcomes:

- The percentage of patients who exhibit a change of 5 mm Hg in average 24-hour diastolic blood pressure from baseline to the 6-month follow-up
- The percentage of patients who exhibit a change of 10 mm Hg and 5 mm Hg in average 24-hour blood pressure from baseline to the 6-month follow-up
- The change in blood pressure load (Textbox 2) percentages from baseline to the 6-month follow-up
- The change in dipping (Textbox 2) percentages from baseline to the 6-month follow-up
- The percentage of patients taking more than 80% (400/500) of the antihypertensive medications at the 6-month follow-up
- The percentage of patients with changes in mean 12-Item Short Form Survey scores from baseline to the 6-month follow-up
- Clinical outcomes, including hospitalization rates and emergency room admissions related to high blood pressure, adverse effects, or other reasons

Textbox 2. Blood pressure load and dipping definitions.

Definitions

- Blood pressure load: the percentage of abnormally elevated blood pressure readings, which is usually provided in the ambulatory blood pressure monitoring report. Normal values should be <40% [14,15].
- Dipping: the difference between the mean systolic pressure during the day and mean systolic pressure during the night. This is expressed as a percentage of the daytime mean, and the accepted normal ranges between 10% and 20% [16].

Ethics and Dissemination

Full ethical approval for the trial was obtained from the local ethics committee of Monastir University of Medicine (reference number: IORG 0009738 N° 50/098060279). The study is registered on the ClinicalTrials.gov registry under the identifier NCT04607239.

The electronic data capture system that was provided to the study is compliant with confidentiality requirements (ie, those of the US FDA Title 21 of the CFR, part 11; HIPAA, ICH; and GDPR), and local regulatory authorization was obtained before starting the project. Personal data that are not relevant to the study will not be collected. Other personal data will be secured by obtaining specific rights for their use in follow-up monitoring only. Investigators and research associates will sign a confidentiality agreement before starting the project.

The final data set will be extracted from the electronic data capture system along with anonymized personal data. The primary investigator will be responsible for access to the final data set for medical writing purposes. The Contract Research Organization team of Dacima is authorized to access the final data set to program statistical analyses and provide the final results of the project according to the statistical analysis plan. The primary investigator will be responsible for reviewing the data analysis and validating the final clinical study report.

Results

Recruitment to the trial started in July 2020. The data collection period is expected to last for approximately 12 months (from the first patient who is enrolled to the last patient who completes the trial). The length of the study per patient will be 6 months. We expect to complete data collection by the end of 2021 and plan to disseminate the results subsequently.

Discussion

Study Contributions

This paper describes the protocol for the HOROSCOPE study—a randomized controlled trial that will assess the impact of telemonitoring on the control of blood pressure. To our knowledge, this will be the first Tunisian study to compare telemonitoring to usual care in terms of the control of blood pressure. If telemonitoring is found to be successful, we will extend the study to other investigation centers throughout the Tunisian territory. We can, in the future, develop connected electronic devices that can send the results of blood pressure measurements directly to doctors.

Strengths and Limitations

Randomized controlled trials are considered the “gold standard for evaluating the efficacy of different interventions in clinical research and constitute evidence for medical treatment” [17]. By conducting such a trial and ensuring internal validity, we will maximize the robustness of our study.

A possible limitation is that enrolled patients may greatly vary in terms of their clinical profiles. In addition, subjects may feel overloaded by the daily measurements and questionnaires that need to be answered, which could result in dropouts. Another possible limitation is that the study population may not represent the overall Tunisian population.

Acknowledgments

This research is supported by MédiS Laboratories, who provided the blood pressure monitoring devices, and by Dacima, who provided statistical advice and the electronic data capture system.

Authors' Contributions

HOROSCOPE Study Group:
Zohra Dridi
Lotfi Ghdira
Mohamed Anis Gtari
Moez Cherif
Sahtout Mounira
Marouen Hafsa
Sami Zayen
Moncef Chebil
Lotfi Mhenni
Mounira Farhat
Nawel Belkacem
Nouri Sayadi
Rim Sayadi
Saber Chefi
Samir Ben Souid
Fahima Hassine
Sonia Khedher Mzali
Olfa Harzallah

Riadh Boukef
Hajer Yaakoubi
Ines Abid
Soumaya Charfi
Wafa Kandil
Nejha Ismail
Imen Trabelsi

Conflicts of Interest

None declared.

References

1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
2. Rodgers A, Ezzati M, Hoorn SV, Lopez AD, Lin RB, Murray CJL, Comparative Risk Assessment Collaborating Group. Distribution of major health risks: findings from the Global Burden of Disease study. PLoS Med 2004 Oct;1(1):e27. [doi: 10.1371/journal.pmed.0010027] [Medline: 15526049]
3. Falaschetti E, Mindell J, Knott C, Poulter N. Hypertension management in England: a serial cross-sectional study from 1994 to 2011. Lancet 2014 May 31;383(9932):1912-1919. [doi: 10.1016/S0140-6736(14)60688-7] [Medline: 24881995]
4. McManus RJ, Mant J, Haque MS, Bray EP, Bryan S, Greenfield SM, et al. Effect of self-monitoring and medication self-titration on systolic blood pressure in hypertensive patients at high risk of cardiovascular disease: the TASMIN-SR randomized clinical trial. JAMA 2014 Aug 27;312(8):799-808. [doi: 10.1001/jama.2014.10057] [Medline: 25157723]
5. Duan Y, Xie Z, Dong F, Wu Z, Lin Z, Sun N, et al. Effectiveness of home blood pressure telemonitoring: a systematic review and meta-analysis of randomised controlled studies. J Hum Hypertens 2017 Jul;31(7):427-437. [doi: 10.1038/jhh.2016.99] [Medline: 28332506]
6. McKinstry B, Hanley J, Wild S, Pagliari C, Paterson M, Lewis S, et al. Telemonitoring based service redesign for the management of uncontrolled hypertension: multicentre randomised controlled trial. BMJ 2013 May 24;346:f3030 [FREE Full text] [doi: 10.1136/bmj.f3030] [Medline: 23709583]
7. Tucker KL, Sheppard JP, Stevens R, Bosworth HB, Bove A, Bray EP, et al. Self-monitoring of blood pressure in hypertension: A systematic review and individual patient data meta-analysis. PLoS Med 2017 Sep 19;14(9):e1002389. [doi: 10.1371/journal.pmed.1002389] [Medline: 28926573]
8. McManus RJ, Mant J, Bray EP, Holder R, Jones MI, Greenfield S, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): a randomised controlled trial. Lancet 2010 Jul 17;376(9736):163-172. [doi: 10.1016/S0140-6736(10)60964-6] [Medline: 20619448]
9. McManus RJ, Mant J, Franssen M, Nickless A, Schwartz C, Hodgkinson J, TASMINH4 investigators. Efficacy of self-monitored blood pressure, with or without telemonitoring, for titration of antihypertensive medication (TASMINH4): an unmasked randomised controlled trial. Lancet 2018 Mar 10;391(10124):949-959 [FREE Full text] [doi: 10.1016/S0140-6736(18)30309-X] [Medline: 29499873]
10. Monahan M, Jowett S, Nickless A, Franssen M, Grant S, Greenfield S, et al. Cost-effectiveness of telemonitoring and self-monitoring of blood pressure for antihypertensive titration in primary care (TASMINH4). Hypertension 2019 Jun;73(6):1231-1239 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.118.12415] [Medline: 31067190]
11. Marzouk-Aissi W, Alaya-Bouafif NB, Bougatef S, Traissac P, Delpeuch F, Romdhane HB. Prévalence et déterminants de l'hypertension artérielle chez les adultes tunisiens, Tunisie. Rev Epidemiol Sante Publique 2012 Sep;60(Supplement 2):S101. [doi: 10.1016/j.respe.2012.06.210]
12. Ware Jr 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]
13. Williams B, Mancia G, Spiering W, Rosei EA, Azizi M, Burnier M, ESC Scientific Document Group. 2018 ESC/ESH Guidelines for the management of arterial hypertension. Eur Heart J 2018 Sep 01;39(33):3021-3104. [doi: 10.1093/eurheartj/ehy339] [Medline: 30165516]
14. Liu M, Li Y, Wei FF, Zhang L, Han JL, Wang JG. Is blood pressure load associated, independently of blood pressure level, with target organ damage? J Hypertens 2013 Sep;31(9):1812-1818. [doi: 10.1097/HJH.0b013e3283624f9b] [Medline: 23743810]
15. Michel Y, Brandstatter H, Gaspoz JM, Muggli F, Pechère-Bertschi A. [Interpretation of 24h ambulatory blood pressure monitoring by primary care physicians]. Rev Med Suisse 2009 Sep 23;5(218):1876-1880. [Medline: 19852417]
16. Bloomfield D, Park A. Night time blood pressure dip. World J Cardiol 2015 Jul 26;7(7):373-376 [FREE Full text] [doi: 10.4330/wjc.v7.i7.373] [Medline: 26225196]

17. Spieth PM, Kubasch AS, Penzlin AI, Illigens BMW, Barlinn K, Siepmann T. Randomized controlled trials - a matter of design. Neuropsychiatr Dis Treat 2016 Jun 10;12:1341-1349 [FREE Full text] [doi: 10.2147/NDT.S101938] [Medline: 27354804]

Abbreviations

ABPM: ambulatory blood pressure monitoring
CFR: Code of Federal Regulation
FDA: Food and Drug Administration
GDPR: General Data Protection Regulation
HIPAA: Health Insurance Portability and Accountability Act
HOROSCOPE: Home Telemonitoring of Arterial Hypertension With Antihypertensive Treatment Titration: Randomized Controlled Prospective Trial
ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
SBP: systolic blood pressure
TASMINH2: Telemonitoring and/or Self-monitoring of Blood Pressure in Hypertension 2
TASMINH4: Telemonitoring and/or Self-monitoring of Blood Pressure in Hypertension 4

Edited by G Eysenbach; submitted 01.12.20; peer-reviewed by C Dziri, M Burnier, J Hanley; comments to author 05.02.21; revised version received 31.03.21; accepted 19.04.21; published 01.03.22.

Please cite as:
Ben Hafaiedh S, Ben Daya Y, Radoui AH, Bouchoucha M, Razgallah R, Nouira S
Home Telemonitoring of Arterial Hypertension With Antihypertensive Treatment Titration: Protocol for a Randomized Controlled Prospective Trial (HOROSCOPE Study)
JMIR Res Protoc 2022;11(3):e26184
URL: https://www.researchprotocols.org/2022/3/e26184
doi:10.2196/26184
PMID:35230254

Protocol

The Efficacy of a Smartphone Game to Prevent HIV Among Young Africans: Protocol for a Randomized Controlled Trial in the Context of COVID-19

Victor Mudhune[1], MBA, MPH; Gaëlle Sabben[2], MPH; Ken Ondenge[1], MPH; Calvin Mbeda[1], BSc; Marissa Morales[2], BSN, MPH, RN; Robert H Lyles[3], PhD; Judith Arego[1], BA; Richard Ndivo[1], BSc; Robert A Bednarczyk[2], PhD; Kelli Komro[4], PhD; Kate Winskell[2], PhD

[1]HIV Research Division, Center for Global Health Research, Kenya Medical Research Institute, Kisumu, Kenya

[2]Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[3]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[4]Department of Behavioral, Social, and Health Education Sciences, Rollins School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:
Kate Winskell, PhD
Hubert Department of Global Health
Rollins School of Public Health
Emory University
1518 Clifton Road
Atlanta, GA, 30322
United States
Phone: 1 404 7275286
Email: swinske@emory.edu

Abstract

Background: Adolescents contribute slightly less than one-third of all new HIV infections in sub-Saharan Africa. There is a need for more effective intervention approaches to help young adolescents safely navigate through adolescence and into adulthood. We are assessing the efficacy of *Tumaini*, a smartphone game designed to prevent HIV among young Africans. Against the background of COVID-19, meaningful alteration of the research protocol was necessary to ensure successful implementation and retention of the study participants in ongoing research.

Objective: The objective of our protocol is to determine (1) if *Tumaini* delays sexual debut and increases condom use at first sex and (2) whether it influences behavioral mediators of early and unprotected sex.

Methods: Participants were recruited from Kisumu County in Western Kenya. This study is a 2-arm, individual-randomized controlled trial that enrolled 1004 adolescents aged between 12 years and 15 years. The intervention arm participants are playing *Tumaini*, while the control arm is provided with *Brainilis*, a commercially available control game. The study period will last 45 months. At baseline, participants in both arms completed a baseline survey and biological testing for HIV and herpes simplex virus, type 2 (HSV-2); participants will have annual game play periods in years 1-3. They will also complete a total of 12 follow-up surveys. At endline, repeat biological testing will be conducted. Protocol adaptations were necessitated by the COVID-19 pandemic and implemented in accordance with local public health guidelines.

Results: Participants were enrolled between October 2020 and November 2020. We plan to complete study procedures in September 2024. The enrolled participant sample was 50.1% (499/996) female and had a mean age of 14.0 (SD 0.6) years.

Conclusions: This ongoing research demonstrates that, with appropriate revisions to planned protocol activities guided by the need to maintain study integrity, protect both study participants and staff, and adhere to institutional review board and local health authority guidelines, human subject research is possible in the context of a global pandemic. If the trial demonstrates efficacy, *Tumaini* would provide an alternative, remote means of delivering age-appropriate education to adolescents on safer sex, HIV prevention, and effective life skills on a highly scalable, low-cost, and culturally adaptable platform.

Trial Registration: ClinicalTrials.gov NCT04437667; https://clinicaltrials.gov/ct2/show/NCT04437667

International Registered Report Identifier (IRRID): DERR1-10.2196/35117

(JMIR Res Protoc 2022;11(3):e35117) doi:10.2196/35117

KEYWORDS

HIV; young Africans; adolescent; Kenya; serious game; game for health; randomized controlled trial; mHealth; prevention; smartphone; teenager; young adult; Africa; gaming; COVID-19; efficacy

Introduction

HIV prevention interventions need to be targeted and packaged to meet the needs of high-risk groups. Significant progress in the discovery and adoption of HIV prevention and treatment measures has resulted in a global decline in the incidence rate by 23% in the last 10 years [1]. To achieve the Joint United Nations Programme on HIV/AIDS (UNAID) Fast-Track strategy to end the AIDS epidemic by 2030 [2], countries need to use all the powerful tools available to them and a leave-no-one-behind approach. In sub-Saharan Africa, a region accounting for almost two-thirds of the global burden of new HIV infections, adolescent girls and young women account for 26% of new infections [3]. In Kenya, one-third of all new infections occur in young people aged 15 years to 24 years [4], with young women at particular risk.

Young people need comprehensive and correct knowledge and skills prior to sexual debut [5-7], to ensure they are able to protect themselves from HIV at their first sexual experience and onwards. Such knowledge and skills need to be well-packaged and age-appropriate. In sub-Saharan Africa, comprehensive HIV prevention knowledge among adolescents remains below 50% in most countries with available data [7-12]. Reaching adolescents with prerisk prevention interventions may help establish lifelong patterns of safer sexual behavior and avert high-risk behaviors in the future [13-16]. Interventions must promote skills for understanding and managing the risks of HIV infection that address important contextual drivers of adolescent risk and are relevant to local practices and culture [17-21]. Increasingly accessible smartphone technologies in Africa make it possible to engage youth—at scale and at low cost—in culturally adapted prevention interventions that require few resources to implement with consistent quality and motivational appeal and incorporate automated data collection [21-23]. If appropriately grounded in behavioral theory and evidence-based practice, electronic games delivered via mobile phones have the potential to become valuable tools in HIV prevention efforts.

Tumaini is a theoretically grounded, narrative-based game for inexpensive Android smartphones designed in collaboration with US-based and Kenyan specialists in adolescent sexual health, with input from Kenyan preadolescents and their parents. It aims to increase age and condom use at first sex by boosting knowledge about sexual health and HIV; building risk avoidance, risk-reduction skills, and related self-efficacy; challenging HIV stigma and harmful gender norms and attitudes; fostering future orientation, goal setting, and planning; and promoting dialogue with adult mentors [24,25]. *Tumaini* is grounded in (1) theory on narrative and narrative-based applied communication, (2) social behavioral theory and existing evidence-based HIV prevention interventions, and (3) principles of instructional design. Tumaini is made up of 3 integrated parts: (1) a central interactive narrative featuring 6 playable characters (3 male, 3 female) whom players guide into and through adolescence, making decisions that have short- and long-term consequences for the characters' lives, relationships, and health; (2) a set of mini-games that tie into the narrative and support and reinforce its core themes; and (3) "My Story," in which the player sets goals, plans how to achieve them, and reflects on how the game relates to his or her life and how he or she might apply the lessons learned to protect his or her future. In a feasibility study with 60 preadolescents conducted in Kenya [26], the intervention showed significant gains in sexual health–related knowledge and self-efficacy, behavioral intention for risk avoidance strategies and sexual risk communication, and overall increased scores on the behavioral survey measures when compared with the control arm at 6 weeks postintervention [27]. Intervention arm participants spent, on average, over 50% longer playing the game than instructed, while quantitative and qualitative data on user engagement and game appeal, relevance, and acceptability among adolescents and parents were extremely positive [28,29]. A randomized controlled trial (RCT), with a larger sample size (anticipated n=1000) and longer duration of follow-up (45 months), was proposed to evaluate behavioral efficacy of the intervention. A detailed protocol was developed that is the subject of this manuscript.

On January 30, 2020, the World Health Organization (WHO) Emergency Committee declared a global health emergency based on growing case notification rates of SARS-CoV-2, the cause of a respiratory illness designated "coronavirus disease 2019" or COVID-19 [30]. WHO published a comprehensive package of guidance documents for countries covering topics related to the management of the outbreak including infection prevention and control measures against COVID-19 [31,32]. On March 11, 2020, COVID-19 was declared a pandemic, and Kenya confirmed its first case on March 13, 2020 [33]. Various public health infection prevention measures were put in place across the country, including in Kisumu, the site of the study. These were expected to be implemented for an extended period and included wearing of masks in public spaces, a regional travel ban, limitations on public gatherings, night curfews, screenings for fever, massive handwashing campaigns, and closure of schools. Although some remote schooling was possible in certain cases (eg, private schools), for the most part, school-aged children remained out of school until October 2020. Research activities were modified to comply with at-the-time current guidance provided by the Kenyan Government and the institutional review boards (IRBs) overseeing study implementation.

This study, which began implementation in October 2020, is being carried out in Kisumu in the context of the COVID-19 pandemic and related public health disease containment measures. The purpose of this paper is to present the protocol for the efficacy RCT, with particular attention to adaptations

from the previous pilot feasibility study, and to inform others about feasible and acceptable protocol modifications in the context of a respiratory virus pandemic.

Methods

RCT Design

This study is a 2-arm, non-blinded, individual RCT targeted to enroll 1000 adolescent participants aged 13 years to 14 years. The intervention-arm participants are playing *Tumaini*, an interactive, narrative-based electronic game, which is loaded on study-provided, low-cost Android smartphones. Adolescents in the control arm are receiving standard of care (ie, no intervention beyond any existing sex education from family, school, and peers) and allocated Android smartphones loaded with *Brainilis*, a free game that challenges the brain in areas like memory, logic, math, and focus, downloaded from Google Play Store. Selection of *Brainilis* as an attention-control game [34] was based on free download availability in Kenya, age-appropriate educational content, a lack of overlap with *Tumaini* in terms of content and game style, and perceived interest for the age group. The intervention period will last 45 months, with participants in both arms completing a baseline survey and 12 follow-up surveys, baseline and endline biological testing for HIV and herpes simplex virus, type 2 (HSV-2), game play periods during the first 3 years, and periodic qualitative data collection to monitor study acceptability and community concerns and to inform potential future dissemination. The 45-month study period is determined by the need for a high enough proportion of participants to become sexually active over the course of the study to allow us to determine the efficacy of the intervention.

The trial was preceded by cognitive interviews [35] for, and piloting and reliability testing of, the quantitative data collection instrument to be used during the trial. These activities were targeted to engage 200 adolescents aged 13 years to 14 years and 32 of their parents; they are henceforth referred to as "Phase 1" of this study.

Ethics Approval

The IRB of Emory University (STUDY00002974) and Kenya Medical Research Institute (KEMRI) Scientific and Ethics Review Unit (SERU: KEMRI/SERU/CGHR/11/3812) approved this study.

Implementation Setting

The study is being conducted in urban and peri-urban locations in East, West, and Central administrative locations of Kisumu County in Western Kenya. Kisumu is the country's third largest city. Data collection is taking place at KEMRI facilities, health facilities, and community halls.

Against the background of the COVID-19 pandemic, study activities originally intended to be conducted in person were modified to comply, as needed in the context of an evolving epidemic, with current guidance provided by the Kenyan Government. This included guidance from the local Ministry of Health, Ministry of Education, and the Departments of Health and Education of the County Government of Kisumu. Emory IRB and the KEMRI SERU provided further guidance aiming to protect study staff and participants. In cases where these guidelines were not in accord, the most stringent guidance was followed.

Target Population

The study recruited different sets of participants for the 2 phases of the study: survey revision and update activities (Phase 1). All participants had to be residents of Kisumu County and have had no previous exposure to or engagement in any research activities related to *Tumaini*, including the formative research.

Phase 1 Population

For this phase of the study, 200 adolescents and 32 parents of adolescents were recruited. Adolescents were required to have basic English literacy and Grade 3-4 on the Flesch-Kincaid Reading Scale (assessed via a short listening and reading comprehension test at enrollment) and were aged between 12 years and 14 years at recruitment. Parents and guardians for this phase had to be a parent or guardian (henceforth referred to as "parent") to a child aged between 12 years and 17 years old at the time of recruitment. In addition to the aforementioned inclusion criteria, the possibility of remote data collection was necessitated by the COVID-19 situation for the activities related to updating the behavioral measures. Adolescent and parent participants were required to have access to an electronic device capable of running Zoom software [36] (ie, a tablet, smartphone, or computer).

Phase 2 Population

For the RCT, 1000 adolescents were to be recruited, as well as a subset of willing parents to the adolescents. In addition to the English language, lack of previous exposure to *Tumaini*-related activities (including survey revision under Phase 1), and age eligibility criteria, we limited enrollment to 1 child per family. In cases where multiple children in a family were eligible and interested, only the child whose name came first in the alphabet was enrolled.

A sample of the adolescent trial participants (n=16) allocated to the intervention arm of the study and a sample of their parents (n=12; including 6 parent-child dyads) were selected to participate in related and repeated annual qualitative research activities. Details of their selection are described in the following sections in the context of the qualitative activities themselves. At endline, 32 other intervention stakeholders, including health care workers, teachers, and community leaders, will supplement the qualitative sample. These additional stakeholders will be identified through stakeholder engagement and be required to have expertise on the sexual health needs of adolescents by virtue of their profession or status in the community to provide relevant insights.

Recruitment and Enrollment

Prior to initiating recruitment activities, approval was obtained from local Ministry of Health and Ministry of Education officials within Kisumu County. Based on experience from the feasibility pilot study, this study had proposed to recruit adolescents through schools for both Phases 1 and 2, working with head teachers to distribute invitation letters to potentially eligible

adolescents to deliver to parents or caregivers. Follow-up in-person informational meetings were then to be used to identify interested and eligible adolescents and parents for recruitment. Because schools closed due to COVID-19 restrictions, active school-based recruitment was not possible. Therefore, participant recruitment was initiated through a combination of strategies in order to ensure a diverse pool of potential participants and allow for more rapid recruitment.

For Phase 1 (survey revisions), the following strategies were employed: (1) schools distributing invitation letters to potentially eligible parents to inform them of the study using existing school communication channels, as well as sharing parents' contact information after consulting those whose children were eligible based on school grade enrollment (grade 7 was identified as the target grade for ease of identification of potential participants as corresponding most closely to the desired age range); (2) recruitment directly in the community with support from the community advisory board (CAB) members' ongoing activities; and (3) recruitment through parents or participants from previous studies who had consented to future contact and their social networks, employing a snowball approach.

For Phase 2 (the RCT), only strategies (1) and (2) implemented in Phase 1 were employed. The need for participants in this phase to be intervention-naïve, as well as concerns about the potential for contamination between study arms, in particular control-arm adolescents accessing the intervention through pre-existing social networks, precluded the use of snowball sampling from previous participants.

Multiple community meetings were held with the support of CAB members at different times in each area of the city to maximize accessibility to interested parents. KEMRI study staff made study presentations, with parents being invited to volunteer for participation. Community gatherings for recruitment and participant screening were conducted outdoors following pandemic guidance regarding number of participants, social distancing, and use of face coverings.

Following a phone discussion or in-person meetings maintaining social distancing, parents' willingness to volunteer for participation was assessed, and interested participants were screened for eligibility based on adolescent age, grade level, and absence of previous participation in related study activities. Child's age eligibility was confirmed via presentation of any official document giving the child's name and date of birth (eg, birth certificate, baptismal certificate, or passport). Interested and eligible parents provided locator information, including a phone number and physical address, where they could be reached for future consent and assent procedures.

Recruitment slots were generated for the target 1000 adolescents for Phase 2, to ensure more targeted recruitment efforts and appropriate distribution of the sample by gender, age, and location (by geographic subzones within East, West, and Central Kisumu). Recruitment slots were filled as soon as a participant matching the slot characteristics was identified and confirmed willing to participate in the study. Priority was given to adolescents aged 13 years or 14 years old, and the sample was supplemented by 12-year-olds when necessary to achieve enrollment targets. Recruitment materials and consent and assent forms were available in Kenya's national languages (English and Kiswahili) and most common local language within the catchment area (Dholuo) to ensure comprehension and were updated to provide for the possibility of remote consenting via phone or Zoom software, depending on participant's preference and local government guidance. No participant opted for remote consenting, so all consenting was done in person, observing infection prevention measures.

Survey Instrument

The behavioral survey instrument used in this study was adapted from the version used in the feasibility study, which is described elsewhere [27]. The survey was revised to remove or replace items that saw a ceiling effect, add items to assess social desirability bias (drawing from the Marlowe-Crowne scale adapted by Vu et al for use in Kenya [37]), and reinforce measurement of certain behavioral constructs to strengthen mediation analysis. In addition, measures were added that had not been appropriate for inclusion in the feasibility study but were warranted for this study because of its length, longitudinal nature, and older participants. Examples include measures related to sexual risk (number of partners, age differential with partners, marriage and pregnancy history, substance use) borrowed from the Project AIM evaluation conducted with adolescents of a similar age in Botswana [38]. Participants are also completing an additional set of game experience survey items. This is included as a component of the behavioral survey or administered independently after each game play period. This questionnaire was updated from that used during the feasibility study [27,29] to include psychological process measures, such as intrinsic motivation, immersion, and identification, as well as items focusing on communication about the game with peers and members of their households, especially parents, and questions to assess potential contamination across study arms.

As a proxy for socioeconomic status (SES), 2 questions were added regarding food insecurity [39] and 2 about house (walls and roof) building materials, drawn from the survey questionnaire used in the routine KEMRI/Centers for Disease Control and Prevention Health and Demographic Surveillance System [40] in Western Kenya. During survey piloting and reliability testing, there was little variation in responses to the roof materials question; hence, it was removed. The final survey instrument includes 109 questions. No single participant will be exposed to the full 109 items at any one time, due to a combination of skip patterns based on participant gender (eg, "Have you started your menses?"), gate questions (eg, "Have you ever had sexual intercourse, or sex?"), and certain questions only being asked at specific time points (eg, social desirability questions will be annual; experience of pregnancy will only be asked at endline).

The survey is delivered via tablet-based Open Data Kit (ODK) software [41] and includes an audio component to ensure consistent understanding of the questions among those with more limited English literacy. This delivery platform was a modification from the feasibility study that used an audio computer-assisted self-interview (ACASI) system. The change of platform allowed the team flexibility in conducting the surveys outside the KEMRI offices on multiple tablets at no

additional cost, an essential criterion given the number of participants and tight period for data collection during school holidays. This approach accommodated the additional complications related to social distancing and in-person meeting limitations during pandemic-era implementation.

Survey Revision

Parents of children similar to those to be included in the RCT were invited to review the full survey in focus group discussions (FGDs) to ensure its acceptability to parents (women, n=8; men, n=7). Subsequently, the survey underwent cognitive interviewing with adolescents to ensure cultural, linguistic, and age appropriateness, with particular attention paid to new questions as detailed in the previous section. We planned 5 rounds of 4 cognitive interviews, but, due to time constraints, we conducted 13 interviews with adolescents (girls, n=6; boys, n=7) to ensure acceptability, face validity, and comprehension, with revisions after each round. Both the survey and the proposed ODK platform were tested for acceptability, by piloting the final instrument complete with audio of all the questions through 2 rounds with 15 adolescents aged 12 years to 14 years using the tablets and headsets intended for study use. Each pilot participant was debriefed by study staff to identify any remaining barriers to comprehension in the survey or issues navigating the survey interface on the tablet. The survey and instructions were updated between piloting rounds, drawing on this feedback.

Assessment of reliability using test-retest was conducted with 150 adolescents aged 12 years to 14 years who took the survey twice approximately 1 month apart using the ODK tablet platform with audio. This sample size represented a balance between budgetary feasibility and minimum acceptable sample size for assessing questionnaire reliability. Temporal reliability of the survey items using Cohen kappa and Goodman and Kruskal gamma coefficients (in SAS/STAT© software, version 9.4) was calculated for participants' answers over the 2 time points. For items where both coefficient values fell below the recommended threshold for acceptable reliability (typically 0.5) [42], the importance of the item to the overall goals of the study was reviewed to ensure that all thematic and theoretical areas of interest were appropriately assessed. In rare cases where an item was deemed too important to remove from the questionnaire, advice was sought from study staff and survey design experts on the team on strategies to boost reliability, and the item was revised as appropriate.

Randomization

Participants were randomized 1:1 to either the control group (*Brainilis*) or the intervention group (*Tumaini*) at a time point between enrollment (consenting) and the baseline survey. Individual randomization was revealed to adolescents after they had completed all baseline data collection activities. The assignments were generated using PROC PLAN in SAS version 9.4 [43,44] to create gender-stratified randomization codes using block randomization with a block size of 10. Other demographics (age, geographic location, school) were expected to be balanced due to the slot allocation process used during recruitment.

Study Procedures

They were instructed to engage in a minimum of 10 hours of game play over each game play period during the long school holidays in December of the first year of the study and in the 2 subsequent years. However, due to COVID-19 disruptions to the school calendar, participants had a December 2020 game play period that will be followed up by a March 2022 game play period, as only these holidays provided a long enough duration for meaningful intervention exposure. The game automatically collects data on participants' in-game behavior, logging time-stamped records of all participant interaction with the app. These will subsequently be analyzed to assess, for example, time spent playing, scores on knowledge-based mini-games, choices made in the narrative game, and components to which the player was exposed. Adolescent participants enrolled in the control arm receive *Brainilis* loaded on a phone identical to that used for the intervention-arm participants at the same time the intervention game is made available to that study arm. Provision of a control game to the nonintervention participants is intended to ensure that any intervention effect was due to the content of the intervention rather than to a phone and game more generally. All enrolled adolescents will complete 13 behavioral surveys over the 45-month period via ODK. Following each game play period, they will also complete a game experience survey (n=3) and indicate which game player profiles are theirs and which were created by family members and friends, for analytical purposes. A schedule of activities is presented in Figure 1. Similarly, following disruption to the school calendar and subsequent reduction of most school holidays to about 1 week until 2023, there was a need to shift resources to accommodate administration of surveys for all participants within the 1-week window.

Figure 1. Schedule of study activities. FGD: focus group discussion; HSV-2: herpes simplex virus, type 2.

	2020		2021				2022			2023			2024		
	Pre-T1	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	Post-T13
Phase 1-survey revision															
FGD with parents (acceptability)	X														
Cognitive interviews (validity)	X														
Pilot test	X														
Test-retest (temporal reliability)	X														
Phase 2- Randomized controlled trial															
Behavioral survey		X	X	X	X	X	X	X	X	X	X	X	X	X	
Biological testing (HIV and HSV-2)		X												X	
Game experience survey			X				X			X					
Discussions with intervention-arm			Semi-structured interviews once a year												FGDs
Discussions with additional intervention sample/stakeholders															X

All study participants completed baseline, and will complete endline, HIV and HSV-2 testing. These procedures are carried out by trained nursing staff at the participants' nearest health facility of their choice. Pre- and posttest HIV counselling is provided as per national guidelines. HIV testing follows a modified national algorithm, involving parallel approved antibody-based rapid test kits at the health facilities, and is confirmed (for discordant results) by HIV DNA polymerase chain reaction at a centralized laboratory. The rationale for the modified HIV-testing algorithm is to reduce chances of false-positive test results, given the relatively low HIV prevalence in the target population, and the attendant risk of causing needless distress to adolescent participants. HSV-2 testing is conducted using HSV-2 enzyme-linked immunosorbent assay (ELISA) kits at a centralized laboratory. Study staff followed up with the participant and testing facility to ensure linkage to care for those who were newly diagnosed, who were contacted 2 weeks after testing for workup to initiate antiretroviral therapy if this could not be done immediately. Participants were retained for the study regardless of their baseline biological test results.

Staff Training and Information Materials

All research personnel completed ethics training and underwent an intensive protocol training before each round of scheduled activities. They reviewed data collection tools, had access to a phone loaded with *Tumaini* and *Brainilis* ahead of the recruitment to familiarize themselves with the games, and reviewed study-specific data collection procedures and a manual of study standard operating procedures. All research team members involved in qualitative research data collection had previous experience, including relevant training during the feasibility study of the intervention. The site management team had previous training in the provision of adolescent-friendly services for sexual and reproductive health research. Targeted health facility nursing staff were already trained on biospecimen collection and adolescent-friendly HIV counselling and testing procedures. The study team includes individuals trained in counselling who assist with identifying emotional distress and refer participants for gender-based violence support as needed.

The study team developed informational handouts based on feedback and experience from the pilot research and from staff training. These were provided to the recruitment and enrollment staff, enabling them to respond to parents' and adolescents' questions and concerns about the study and the game in a uniform way and accurately represent the goals of the study.

Due to the ongoing pandemic, all study personnel also received training on COVID-19–related protocols for personal and participant health and safety, including the proper use of personal protective equipment (PPE). Key study personnel were already familiar with Zoom teleconferencing software. Staff training was conducted on recognition and documentation of signs and symptoms related to COVID-19, including the need for self-isolation and access to appropriate services as recommended by the local Ministry of Health in case of an exposure or confirmed case of COVID-19.

Phone Setup

Tumaini and *Brainilis* were downloaded and programmed into the intervention and control phones, respectively. Similar to the setup used during the feasibility test, all other phone functions were disabled using a parental control app. The same phone will be issued to the same participant during each game play period in order to allow the participant to resume prior game play, if desired. Where phones are lost or cannot be restored, a replacement will be logged and issued. The phone identification number and participant identification number are linked in a separate phone log, which will be used as reference for every phone allocation visit and to match in-game data from *Tumaini* log files to participants' other study data in the LabKey data management platform. To complement data collected through the ODK behavioral surveys, intervention-arm players' in-game log files (paradata) will be downloaded upon return of the phones each year. These files contain time-stamped details of player interaction with the game, which will be used to examine exposure to the intervention.

Qualitative Data Collection and Postintervention Procedures

The intervention efficacy activities are being complemented by qualitative monitoring of a sample of the intervention arm. A

cohort of adolescent participants from this study arm and a sample of parents of adolescent participants, including 6 parent-child dyads, have been selected to take part in semi-structured interviews (SSIs) once a year during the trial and in FGDs at the end of the trial. These 28 individuals were sampled to represent different SES profiles and types of game players. SES was based on reported food insecurity and house materials on the baseline behavioral survey, while player type was drawn from the behavioral survey (experience with video games) and the game experience survey administered after the first game play period (self-reported time spent on *Tumaini*, degree of completion of the game, and engagement with others about *Tumaini*). In addition to annual SSIs, this cohort will also be invited to share insights into their experiences throughout the study during post-endline FGDs.

A further 8 FGDs will be convened after endline, 4 for additional intervention arm participants and 4 for additional parents, with participants being selected in a way similar to the SSI cohort to represent a range of demographic and intervention-user profiles. Other stakeholders from the community (n=32; eg, teachers, healthcare providers, community leaders) will be invited to participate in endline SSIs or FGDs. These SSIs and FGDs will elicit feedback on the game and study experience that will inform future game dissemination and scale-up and development of future interventions and trials.

Other COVID-19 Adaptations

In order to minimize risks related to COVID-19, study activities are being conducted remotely (eg, video conferencing, phone calls) to the extent possible. As remote data collection relies on participants having a phone, tablet, or computer with internet access in their households, it entails SES bias and hence is only deemed feasible where such bias does not threaten the scientific integrity of the study.

In cases where study activities cannot be conducted remotely, these are conducted at a distance of 2 to 3 meters (about 6-10 feet) from study staff or other participants. Prior to all in-person contact, participants and study staff are screened for symptoms and undergo temperature screenings; these data will not be analyzed as part of this study. Physically distanced study activities (eg, SSIs, informational sessions, consent or assent) involving verbal interaction take place outdoors to the extent possible, and both staff and participants wear masks. If physical distancing cannot be maintained (eg, for blood collection), appropriate PPE (masks and gloves) is used. If they take place in dedicated spaces at KEMRI or community-based organization offices, participants are scheduled to avoid proximity and extend waiting time, and the research area is disinfected between participants. Where electronic devices are handed to participants for survey data collection or intervention delivery, these devices are disinfected prior to distribution and after collection.

So far, all study procedures have been conducted in person following local infection prevention measures, with agreement from the study participants. Options to conduct the procedures remotely remain available to study staff and participants should they be needed.

Safety Monitoring

An independent safety monitor (ISM), appointed for this study, was selected based on prior experience with clinical trials, HIV behavioral interventions, adolescent sexual health, and experience working in Kenya. The ISM's eligibility and lack of conflict of interest were independently confirmed. The ISM reviews study materials including participant safety data and overall study conduct as specified in the protocol. The ISM addresses issues of research participant protection, by examining over time the safety data from the study in order to evaluate safety findings and trends and make recommendations concerning continuation, termination, or other modification of the study based on the observed beneficial or adverse effects and social harms associated with the intervention. In addition, the ISM reviews the general progress and conduct of the study regularly and assists in resolving any problems that may arise.

Sample Size

At age 13 years, according to the most recent available data [45], in Kenya, 4.2% of female and 12.8% of male adolescents have reached sexual debut; by age 17, these figures rise to 40.9% and 42.3% respectively. Condom use at first sex is reported to be 22.8%, 43.3%, and 67.2% for those having sex by age 13 years, before 15 years of age, and after 15 years of age, respectively [45]. Based on these estimates and outcomes of other interventions targeting similar behavior in this age group [14,16], we calculated sample sizes based on a primary binary outcome of "risk" group (experienced sexual debut during the study period without condom use at first sex) versus "low-risk" group (not yet experienced sexual debut by the end of the study or used a condom at sexual debut during the study period). Targeting 1:1 randomization into the 2 study arms, we aimed to consent and enroll a total of 1000 participants. This would be sufficient for 80% power to detect a difference between 18% of control-arm participants "and 11% of intervention-arm participants in the "risk" group for our binary outcome at the α=.05 level, accounting for 25% potential loss to follow-up. This sample size also allows for stratification of participants by HIV status at baseline, when we anticipate around 10 participants to test positive, based on local HIV prevalence among this age group.

Statistical Analysis

The distribution of the primary binary "risk" outcome of this study will be assessed via standard methods for comparing proportions and related 95% CIs. This will be followed by logistic regression modeling to account for covariates, to be determined after analysis of baseline demographic and anticipated confounder data. The crude mean and median age at first sex will be ascertained for all participants who experience their debut over the course of the study, augmenting this analysis to account for gender and other covariates using parametric or semiparametric regression models as appropriate. As with the primary outcome, the binary response of condom use at first sex among those experiencing debut during the study will be analyzed via logistic regression. To determine whether the game has the potential to reduce HSV-2 or HIV infection, incidence rates over the course of the study of HSV-2 and HIV among

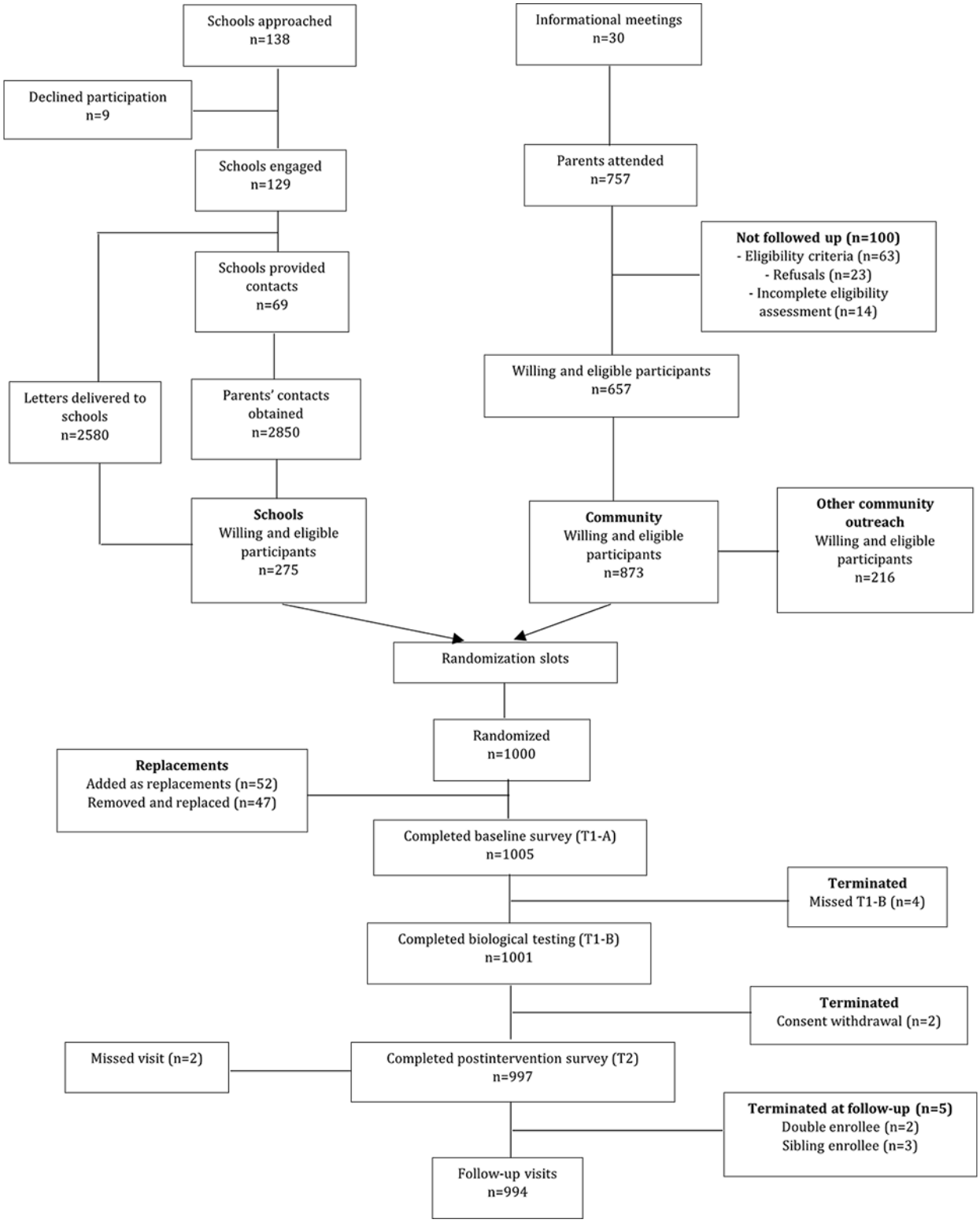
uninfected participants will also be compared by exposure to intervention using Poisson and Cox regression methods.

To determine whether the game-based intervention influences behavioral mediators (knowledge, attitudes, behavioral intentions, and related self-efficacy) of early and unprotected sex, crude and adjusted associations between behavioral mediators and exposure to the intervention will be explored using multivariable regression methods appropriate to each potential mediator, following stipulated principles of mediation analysis [46]. Using baseline behavioral survey data, factor analyses will be conducted to identify theoretical and thematic scales. Survey data will be analyzed both as individual items and as composite scores for these scales. Analyses of outcome data, game experience survey data, and the paradata from *Tumaini* players' smartphones will be used to assess which game components and theoretical constructs mediated intervention effects [46,47].

Results

Recruitment, consenting, and enrollment were carried out between October 1, 2020 and December 3, 2020. In line with our COVID-modified protocol and with schools closed, a total of 138 primary schools were approached within Kisumu County, either directly to distribute recruitment letters or to provide contact details of parents with children in grade 7 (Figure 2). Recruitment letters were randomly distributed to willing schools, but we did not track how many letters were subsequently issued by schools to potential participants. From contact details obtained from schools, the majority of the potential participants were ineligible due to those adolescents not being in grade 7 or within the target age range. There was generally very low yield from the school-based activities due to school closure, which were deemed inefficient and stopped midway through the recruitment period; hence, some contacts were not followed up on.

Figure 2. Trial flow diagram.



Within the same period, 30 informational meetings were held in the community, with 757 parents attending. Of these, a small fraction was ineligible due to their child being older than 14 years, in grade 8, or failing to pass the English proficiency test for the study. The main reasons for parental refusal included concerns with blood draws, concerns with the adolescent being issued a phone, fear of HIV testing, concern with adolescent safety, time constraints associated with study visits, and unwillingness to participate in research. The majority of the participants in this study was enrolled from community meetings.

As noted previously, recruitment randomization slots for 1000 individuals were generated and used to ensure distribution of the sample by gender, age, and residential area prior to randomization by gender. There was a time lag between being recruited and administration of the baseline survey for some participants; hence, a number of eligible and willing adolescents could not be traced for enrollment or had moved out of the study

area. After sampling from the full pool of interested and eligible individuals, 148 willing and eligible adolescents who were not randomized served as a pool for replacement for those who could not initiate study activities. The randomization slot characteristics were used to identify suitable replacements from the pool. During this replacement period, some participants who were identified as needing to be replaced ended up presenting themselves for a baseline survey appointment; hence, the study ended up with 1005 enrolled individuals completing the baseline survey. There were 4 participants who completed the baseline survey only and not the biological testing; they were removed from the study. Baseline behavioral surveys were conducted between November 17, 2020 and December 8, 2020, and biological testing was conducted between November 25, 2020 and December 12, 2020. Study phones with either the intervention or control game were distributed after completion of the biological tests; participants were in possession of the phones for a median period of 37 (IQR 34-40) days.

Postintervention surveys were conducted between January 2, 2021 and January 27, 2021, at which point the study team discovered a participant who was enrolled with 2 different study identification numbers resulting from presenting themselves twice and 3 different cases of siblings living within the same household enrolled in contravention of the protocol. The double enrollee was terminated from the study. In cases where the siblings were in different study arms, the intervention arm participant was retained. In cases where they were both in the same arm, the study criteria for selecting a sibling at recruitment was applied—the first participant based on alphabetical order of their first names was retained. After elimination of individuals later found to be ineligible for these reasons, the study had 1001 participants. The total number who completed baseline and will be included in baseline analyses is 996.

Demographic characteristics of these 996 participants who completed all baseline activities, were allocated to the 2 study arms, and received a phone loaded with either the intervention or control game were evaluated (Table 1). This analysis excludes the double enrollee and 3 siblings who were terminated. The participant sample was 50.1% (499/996) female and had a mean age of 14.0 (SD 0.6) years. At baseline, participants were almost universally enrolled in school, with the majority reporting either grade 6 or 7 as the highest grade they had completed. SES, using housing materials as proxy, was estimated to be almost evenly split between “higher” (507/992, 51.1%) and “lower” SES (485/992, 48.9%), with 4 participants not providing this information. On the combined measure of food insecurity (ie, household hunger either during the day or night or overnight), more than one-half (530/996, 53.2%) of participants reported no food insecurity (“never” responses on both questions), while 33 (33/996, 3.3%) reported high levels thereof (“sometimes” or “often” responses on both questions). Close to 90% (883/996, 88.6%) of participants reported living with at least one parent, with the others, for the most part, in the care of grandparents or other adults, including aunts, uncles, and other guardians. The majority (892/977, 91.3%) of participants identified as Christian (Catholic, Protestant/Anglican, Seventh Day Adventist), while a small number identified their religion as Muslim (31/977, 3.2%) or another local religion (21/977, 2.1%) or reported not being part of a religious group (33/977, 3.4%). Religiosity is high among participants, with 78.0% (777/996) reporting attending services once a week. Household technology access is relatively high, with only 17.9% (178/996) of participants reporting no household smartphone ownership, and 37 (37/996, 3.7%) indicating that they themselves owned a smartphone. We also assessed prebaseline levels of engagement with video games. Although 22.8% (227/996) reported never playing video games, 114 participants (114/996, 11.4%) indicated that they often play. Analysis by arm has not commenced as the study is ongoing; these demographics are therefore presented in aggregate rather than by study arm.

Table 1. Participant demographics (n=996) at baseline.

Characteristics	Results
Sex, n (%)	
Female	499 (50.1)
Male	497 (49.9)
Age (years), mean (SD)	14.0 (0.6)
Attending school (yes), n (%)	992 (99.6)
Highest school grade completed, n (%)	
Grade 5	43 (4.3)
Grade 6	630 (63.3)
Grade 7	280 (28.1)
Grade 8	43 (4.3)
Socioeconomic status[a], n (%)	
High	507 (51.1)
Lower	398 (48.9)
Missing	4 (—)
Household food insecurity[b], n (%)	
None	530 (53.2)
Low	287 (28.8)
Medium	146 (14.7)
High	33 (3.3)
Living situation, n (%)	
Both parents	547 (54.9)
Mother only	276 (27.7)
Father only	60 (6.0)
Grandparents	72 (7.2)
Orphanage or children’s home	3 (0.3)
Other living situation	38 (3.8)
Religion, n (%)	
Catholic	356 (36.4)
Protestant/Anglican	293 (30.0)
Seventh Day Adventist	243 (24.9)
Muslim	31 (3.2)
Other local religion	21 (2.1)
No religion	33 (3.4)
Missing	19 (—)
Attendance at religious services, n (%)	
Once per week	777 (78.0)
1-2 times per month	77 (7.7)
A few times per year	83 (8.3)
Once per year or less	22 (2.2)
Never	37 (3.7)
Household access to smartphone, n (%)	818 (82.1)

Characteristics	Results
Frequency of video game play, n (%)	
Never	227 (22.8)
Rarely	365 (36.7)
Sometimes	290 (29.1)
Often	114 (11.4)

[a]Determined by materials used to construct their homes.

[b]Based on scoring index developed from reported number of days household went hungry or slept hungry.

Discussion

Study Overview

Based on findings from our pilot study of *Tumaini*, this study aims to determine the efficacy of the intervention in delaying sexual debut and increasing condom use at first sex. This study will also determine whether the game-based intervention influences behavioral mediators (knowledge, attitudes, behavioral intentions, and related self-efficacy) of early and unprotected sex and which particular game components and theoretical constructs mediate the desired effects. This study responds to the need for an increased number of behavioral and structural HIV intervention options appropriate for adolescents [48], as adolescents and young people require tailored approaches to meet their needs. The game will also collect data to inform potential implementation on a wider scale, if warranted.

The COVID-19 pandemic has led to substantial changes in routine daily activities and social interactions. Through these and other challenges, the pandemic has affected implementation of ongoing and planned research across the globe. There is a call for research teams to get creative about ways of reaching, engaging, and reimbursing study participants [49] and, in doing so, follow the guidelines provided by local authorities and ethical review committees. The study local ethics and research committee provided guidelines that aim to protect trial integrity and study subjects [50], which informed the proposed changes to the protocol. Researchers have had to identify activities that do not place the study participants at increased risk of COVID-19 while maintaining study rigor, amid debates about how implementing research amid the COVID-19 pandemic may affect the balance between participant risks and benefits [51]. Physical distancing to protect participant and researcher safety has been one of the major concerns with research implementation at this time. For this study, we were able to offer remote consenting; virtual interview options using Zoom software or phone calls; and, whenever physical contact was inevitable, maintaining recommended social distancing, enforced use of appropriate PPE and a shift in planned activities to accommodate the availability of the target population. Online platforms have been shown to be viable options for conducting research during the pandemic [44], with the need for the researcher to be aware of the inherent potential biases [52] and of the limitations of lack of nonverbal cues and privacy and access issues [53]. In our case, despite securing ethics approval, anticipating a need for remote activities, and taking measures to minimize any biases associated with technology access, no participant was consented remotely, nor did any study activities occur via teleconferencing. This is largely because local meetings were permitted with appropriate social distancing and face coverings and were preferred by our participants.

The pandemic has further highlighted some of the advantages of *Tumaini*'s remote delivery platform, and this study is likely to inform creative ways with which we target adolescents with HIV prevention interventions. Many of the countries in sub-Saharan Africa are adopting comprehensive sexuality education amid varying challenges, including unreceptive sociocultural norms, parental attitudes, teacher-related challenges, and economic factors [54,55]. There is considerable interest in mobile health (mHealth) to support the delivery of HIV care and prevention services in the general population, as this offers an opportunity to provide such education at scale and low marginal cost [56]. However, very few mHealth interventions are specifically targeting adolescents with customized messages. If found to be efficacious, *Tumaini* will provide an alternative, appropriate, acceptable, and contextualized mode of passing HIV prevention messages and skills building to adolescents within sub-Saharan Africa.

Limitations

The study design has inherent limitations, which have been compounded by COVID-19–necessitated modifications and implications. First, there is a risk that participants in the intervention arm may discuss or even share the *Tumaini* app with participants in the control arm. To reduce the risk of contamination, study visits will be scheduled such that participants from the 2 arms do not mix during data collection activities. Second, survey activities and game play periods were planned to coincide with normal school holidays, which would be at least 3 weeks, with the game play periods planned during the year-end long holiday period, usually lasting around 6 weeks. With the shift in the academic calendar in Kenya because of forced COVID-19–related school closure, the holidays were subsequently shortened to only 1 week in most cases. This results in a shorter than desired period to conduct the surveys and game play periods. Resulting changes to game play periods may require that the study implement longer than 1-year intervals between subsequent game play periods, with the risk of affecting recall of game contents and key messages and skills in the interim. Third, there are limitations with using self-reported outcomes for condom use and sexual debut. We are supplementing self-report data with biological measures of HIV and HSV-2 at baseline and endline; however, the study is not powered to detect differences in these biological outcomes

due to relatively low rates of HIV or HSV-2 in this age group. Finally, there is an inherent threat of loss-to-follow-up over the 4 years of the study as the adolescents progressively transition through the school system or migrate outside the study area. In addition to accounting for the potential of up to 25%loss-to-follow-up in our sample size calculations, we plan to have annual parental engagement activities to keep parents engaged and inform them of upcoming study activities for the year ahead. Such engagement activities are likely to inform us of planned migration out of the study area and encourage parents to facilitate continued adolescent participation in the study.

Conclusions

This ongoing research demonstrates that, with appropriate revisions to planned activities, incorporation of lessons learned from our previous pilot study, and compliance with regulatory guidelines, recruitment and conduct of research procedures are possible against the background of a global pandemic. This study identified and recruited adolescents for an efficacy trial to evaluate a smartphone game–based intervention aiming to delay sexual debut and increase condom use at first sex. If the trial demonstrates efficacy, it will provide an alternative means of delivering age-appropriate education to adolescents on safer sex, HIV prevention, and effective life skills. It would also identify implementation challenges and how to overcome them in a potential wider scale rollout.

Acknowledgments

Research reported in this publication was supported by the National Institute of Mental Health of the US National Institutes of Health under Award Number 5R01MH118982 (principal investigator: KW). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This research was also supported by the Emory Center for AIDS Research (P30 AI050409). The sponsors played no role in review and approval of the manuscript for publication. The authors thank all of the study staff at Kenya Medical Research Institute (KEMRI) HIV Research Division in Kisumu for their support in operationalizing the study, and in recruiting study participants. This manuscript was approved by Deputy Director, KEMRI Centre for Global Health Research (CGHR).

Authors' Contributions

KW led the design of the study, and GS, KK, RB, RL, VM, and KO contributed to the design of the study. VM, KO, JA, and CM contributed to site-specific modifications and implementation. MM contributed to data management. MM and GS conducted analyses. VM and GS drafted the manuscript, and KO, CM, RL, JA, RN, RB, KK, MM, and KW edited and reviewed the manuscript. KW is overall principal investigator, while VM is the site principal investigator for this study.

Conflicts of Interest

None declared.

References

1. HIV AIDS statistics - 2019 fact sheet. UNAIDS. 2019. URL: https://www.unaids.org/sites/default/files/media_asset/2019-UNAIDS-data_en.pdf [accessed 2022-02-20]
2. Fast-Track - Ending the AIDS epidemic by 2030. UNAIDS. 2014 Nov 18. URL: https://www.unaids.org/en/resources/documents/2014/JC2686_WAD2014report [accessed 2022-02-20]
3. 2020 Global AIDS Update — Seizing the moment — Tackling entrenched inequalities to end epidemics. UNAIDS. 2020 Jul 06. URL: https://www.unaids.org/en/resources/documents/2020/global-aids-report [accessed 2022-02-20]
4. Kenya HIV Estimates: Report 2018. National AIDS Control Council. 2018. URL: https://nacc.or.ke/wp-content/uploads/2018/11/HIV-estimates-report-Kenya-20182.pdf [accessed 2022-02-20]
5. Tenkorang EY, Rajulton F, Maticka-Tyndale E. Perceived risks of HIV/AIDS and first sexual intercourse among youth in Cape Town, South Africa. AIDS Behav 2009 Apr 10;13(2):234-245. [doi: 10.1007/s10461-008-9470-5] [Medline: 18846419]
6. Santelli JS, Speizer IS, Edelstein ZR. Abstinence promotion under PEPFAR: the shifting focus of HIV prevention for youth. Glob Public Health 2013 Jan 18;8(1):1-12 [FREE Full text] [doi: 10.1080/17441692.2012.759609] [Medline: 23327516]
7. Butts SA, Kayukwa A, Langlie J, Rodriguez VJ, Alcaide ML, Chitalu N, et al. HIV knowledge and risk among Zambian adolescent and younger adolescent girls: challenges and solutions. Sex Educ 2018 Aug 31;18(1):1-13 [FREE Full text] [doi: 10.1080/14681811.2017.1370368] [Medline: 31275062]
8. Badru T, Mwaisaka J, Khamofu H, Agbakwuru C, Adedokun O, Pandey SR, et al. HIV comprehensive knowledge and prevalence among young adolescents in Nigeria: evidence from Akwa Ibom AIDS indicator survey, 2017. BMC Public Health 2020 Jan 13;20(1):45 [FREE Full text] [doi: 10.1186/s12889-019-7890-y] [Medline: 31931760]
9. Asaolu IO, Gunn JK, Center KE, Koss MP, Iwelunmor JI, Ehiri JE. Predictors of HIV testing among youth in sub-Saharan Africa: a cross-sectional study. PLoS One 2016 Oct 5;11(10):e0164052 [FREE Full text] [doi: 10.1371/journal.pone.0164052] [Medline: 27706252]

10. Mumtaz GR, Hilmi N, Majed EZ, Abu-Raddad LJ. Characterising HIV/AIDS knowledge and attitudes in the Middle East and North Africa: Systematic review and data synthesis. Glob Public Health 2020 Feb 26;15(2):275-298. [doi: 10.1080/17441692.2019.1668452] [Medline: 31558094]
11. Mandiwa C, Namondwe B, Munthali M. Prevalence and correlates of comprehensive HIV/AIDS knowledge among adolescent girls and young women aged 15-24 years in Malawi: evidence from the 2015-16 Malawi demographic and health survey. BMC Public Health 2021 Aug 04;21(1):1508 [FREE Full text] [doi: 10.1186/s12889-021-11564-4] [Medline: 34348679]
12. Finlay JE, Assefa N, Mwanyika-Sando M, Dessie Y, Harling G, Njau T, et al. Sexual and reproductive health knowledge among adolescents in eight sites across sub-Saharan Africa. Trop Med Int Health 2020 Jan 08;25(1):44-53 [FREE Full text] [doi: 10.1111/tmi.13332] [Medline: 31691455]
13. Poobalan AS, Pitchforth E, Imamura M, Tucker JS, Philip K, Spratt J, et al. Characteristics of effective interventions in improving young people's sexual health: a review of reviews. Sex Education 2009 Aug;9(3):319-336. [doi: 10.1080/14681810903059185]
14. Jemmott JB, Jemmott LS, O'Leary A, Ngwane Z, Lewis DA, Bellamy SL, et al. HIV/STI risk-reduction intervention efficacy with South African adolescents over 54 months. Health Psychol 2015 Jun;34(6):610-621 [FREE Full text] [doi: 10.1037/hea0000140] [Medline: 25110841]
15. Jemmott JB, Jemmott LS, O'Leary A, Ngwane Z, Icard LD, Bellamy SL, et al. School-based randomized controlled trial of an HIV/STD risk-reduction intervention for South African adolescents. Arch Pediatr Adolesc Med 2010 Oct 04;164(10):923-929 [FREE Full text] [doi: 10.1001/archpediatrics.2010.176] [Medline: 20921349]
16. Jemmott LS, Jemmott JB, Ngwane Z, Icard L, O'Leary A, Gueits L, et al. 'Let Us Protect Our Future' a culturally congruent evidenced-based HIV/STD risk-reduction intervention for young South African adolescents. Health Educ Res 2014 Feb 20;29(1):166-181 [FREE Full text] [doi: 10.1093/her/cyt072] [Medline: 23962491]
17. Napierala Mavedzenge SM, Doyle AM, Ross DA. HIV prevention in young people in sub-Saharan Africa: a systematic review. J Adolesc Health 2011 Dec;49(6):568-586. [doi: 10.1016/j.jadohealth.2011.02.007] [Medline: 22098767]
18. DiClemente RJ, Crosby RA. Preventing HIV among adolescents in Sub-Saharan Africa. J Adolesc Health 2009 Feb;44(2):101-102. [doi: 10.1016/j.jadohealth.2008.12.001] [Medline: 19167656]
19. Harrison A, Newell M, Imrie J, Hoddinott G. HIV prevention for South African youth: which interventions work? A systematic review of current evidence. BMC Public Health 2010 Feb 26;10(1):102 [FREE Full text] [doi: 10.1186/1471-2458-10-102] [Medline: 20187957]
20. Michielsen K, Chersich MF, Luchters S, De Koker P, Van Rossem R, Temmerman M. Effectiveness of HIV prevention for youth in sub-Saharan Africa: systematic review and meta-analysis of randomized and nonrandomized trials. AIDS 2010 May 15;24(8):1193-1202. [doi: 10.1097/QAD.0b013e3283384791] [Medline: 20375876]
21. Pettifor A, Bekker LG, Hosek S, DiClemente R, Rosenberg M, Bull SS, HIV Prevention Trials Network (HPTN) Adolescent Scientific Committee. Preventing HIV among young people: research priorities for the future. J Acquir Immune Defic Syndr 2013 Jul;63 Suppl 2(0 2):S155-S160 [FREE Full text] [doi: 10.1097/QAI.0b013e31829871fb] [Medline: 23764629]
22. Visser M, Kotze M, van Rensburg MJ. An mHealth HIV prevention programme for youth: lessons learned from the iloveLife.mobi programme in South Africa. AIDS Care 2020 May 19;32(sup2):148-154. [doi: 10.1080/09540121.2020.1742866] [Medline: 32189510]
23. Rotheram-Borus MJ, Lee S, Swendeman D. Getting to zero HIV among youth: moving beyond medical sites. JAMA Pediatr 2018 Dec 01;172(12):1117-1118 [FREE Full text] [doi: 10.1001/jamapediatrics.2018.3672] [Medline: 30326015]
24. Winskell K, Sabben G, Obong'o C. Interactive narrative in a mobile health behavioral intervention (Tumaini): theoretical grounding and structure of a smartphone game to prevent HIV among young Africans. JMIR Serious Games 2019 May 08;7(2):e13037 [FREE Full text] [doi: 10.2196/13037] [Medline: 31066692]
25. Winskell K, Sabben G, Akelo V, Ondeng'e K, Odero I, Mudhune V. A smartphone game to prevent HIV among young Kenyans: local perceptions of mechanisms of effect. Health Educ Res 2020 Jun 01;35(3):153-164 [FREE Full text] [doi: 10.1093/her/cyaa011] [Medline: 32441759]
26. Sabben G, Akelo V, Mudhune V, Ondeng'e K, Ndivo R, Stephenson R, et al. A smartphone game to prevent HIV among young Africans: protocol for a randomized pilot study of a mobile intervention. JMIR Res Protoc 2019 Mar 27;8(3):e11209 [FREE Full text] [doi: 10.2196/11209] [Medline: 30916661]
27. Winskell K, Sabben G, Akelo V, Ondeng'e K, Obong'o C, Stephenson R, et al. A smartphone game-based intervention (Tumaini) to prevent HIV among young Africans: pilot randomized controlled trial. JMIR Mhealth Uhealth 2018 Aug 01;6(8):e10482 [FREE Full text] [doi: 10.2196/10482] [Medline: 30068501]
28. Sabben G, Mudhune V, Ondeng'e K, Odero I, Ndivo R, Akelo V, et al. A smartphone game to prevent HIV among young Africans (Tumaini): assessing intervention and study acceptability among adolescents and their parents in a randomized controlled trial. JMIR Mhealth Uhealth 2019 May 21;7(5):e13049 [FREE Full text] [doi: 10.2196/13049] [Medline: 31115348]
29. Winskell K, Sabben G, Ondeng'e K, Odero I, Akelo V, Mudhune V. A smartphone game to prevent HIV among young Kenyans: household dynamics of gameplay in a feasibility study. Health Educ J 2019 Aug 01;78(5):595-606 [FREE Full text] [doi: 10.1177/0017896919832344] [Medline: 34219796]

30. Helmy YA, Fawzy M, Elaswad A, Sobieh A, Kenney SP, Shehata AA. The COVID-19 pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and control. J Clin Med 2020 Apr 24;9(4):1225 [FREE Full text] [doi: 10.3390/jcm9041225] [Medline: 32344679]
31. Responding to community spread of COVID-19: interim guidance. World Health Organization. 2020 Mar 07. URL: https://www.who.int/publications-detail-redirect/responding-to-community-spread-of-covid-19 [accessed 2022-02-20]
32. Overview of Public Health and Social Measures in the context of COVID-19. World Health Organization. 2020 May 18. URL: https://www.who.int/publications-detail-redirect/overview-of-public-health-and-social-measures-in-the-context-of-covid-19 [accessed 2022-02-20]
33. Press statement on the update of coronavirus in the country as at 23rd March 2020. Ministry of Health. 2020 Mar 23. URL: https://www.health.go.ke/wp-content/uploads/2020/03/COVID19_Press-Statment-March-23-2020.pdf [accessed 2022-02-20]
34. Aycock DM, Hayat MJ, Helvig A, Dunbar SB, Clark PC. Essential considerations in developing attention control groups in behavioral research. Res Nurs Health 2018 Jun 14;41(3):320-328. [doi: 10.1002/nur.21870] [Medline: 29906317]
35. Willis GB. Cognitive Interviewing. Berlin, Heidelberg: Sage Publications, Inc; 2005.
36. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. International Journal of Qualitative Methods 2019 Sep 11;18:160940691987459. [doi: 10.1177/1609406919874596]
37. Vu A, Tran N, Pham K, Ahmed S. Reliability of the Marlowe-Crowne social desirability scale in Ethiopia, Kenya, Mozambique, and Uganda. BMC Med Res Methodol 2011 Dec 02;11(1):162 [FREE Full text] [doi: 10.1186/1471-2288-11-162] [Medline: 22136357]
38. Clark LF, Miller KS, Nagy SS, Avery J, Roth DL, Liddon N, et al. Adult identity mentoring: reducing sexual risk for African-American seventh grade students. J Adolesc Health 2005 Oct;37(4):337-337.e10. [doi: 10.1016/j.jadohealth.2004.09.024] [Medline: 16182145]
39. Connell C, Nord M, Lofton KL, Yadrick K. Food security of older children can be assessed using a standardized survey instrument. J Nutr 2004 Oct;134(10):2566-2572. [doi: 10.1093/jn/134.10.2566] [Medline: 15465749]
40. Odhiambo FO, Laserson KF, Sewe M, Hamel MJ, Feikin DR, Adazu K, et al. Profile: the KEMRI/CDC health and demographic surveillance system--Western Kenya. Int J Epidemiol 2012 Aug 29;41(4):977-987. [doi: 10.1093/ije/dys108] [Medline: 22933646]
41. Hartung C, Lerer A, Anokwa Y, Tseng C, Brunette W, Borriello G. Open data kit: tools to build information services for developing regions. 2010 Presented at: 4th ACM/IEEE International Conference on Information and Communication Technologies and Development; December 13-16, 2010; London, United Kingdom. [doi: 10.1145/2369220.2369236]
42. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012:276-282. [doi: 10.11613/bm.2012.031]
43. Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials, 3rd edition. Cham, Switzerland: Springer International Publishing; 1998.
44. Deng C, Graz J. Generating Randomization Schedules Using SAS Programming. 2002 Presented at: SAS Users Group International 27; April 14-17, 2002; Orlando, FL.
45. De Cock KM, Rutherford GW, Akhwale W. Kenya AIDS Indicator Survey 2012. J Acquir Immune Defic Syndr 2014 May 01;66 Suppl 1:S1-S2. [doi: 10.1097/QAI.0000000000000152] [Medline: 24732813]
46. 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]
47. Stigler MH, Kugler KC, Komro KA, Leshabari MT, Klepp KI. AIDS education for Tanzanian youth: a mediation analysis. Health Educ Res 2006 Aug 22;21(4):441-451. [doi: 10.1093/her/cyh070] [Medline: 16303784]
48. Hosek S, Pettifor A. HIV prevention interventions for adolescents. Curr HIV/AIDS Rep 2019 Feb 1;16(1):120-128 [FREE Full text] [doi: 10.1007/s11904-019-00431-y] [Medline: 30707399]
49. Sevelius JM, Gutierrez-Mock L, Zamudio-Haas S, McCree B, Ngo A, Jackson A, et al. Research with marginalized communities: challenges to continuity during the COVID-19 pandemic. AIDS Behav 2020 Jul 16;24(7):2009-2012 [FREE Full text] [doi: 10.1007/s10461-020-02920-3] [Medline: 32415617]
50. KEMRI SERU Guidelines for the Conduct of Research During the COVID-19 Pandemic in Kenya. Kenya Medical Research Institute. URL: https://www.kemri.go.ke/wp-content/uploads/2019/11/KEMRI-SERU_GUIDELINES-FOR-THE-CONDUCT-OF-RESEARCH-DURING-THE-COVID_8-June-2020_Final.pdf [accessed 2022-02-20]
51. Hsu NS, Hendriks S, Ramos KM, Grady C. Ethical considerations of COVID-19-related adjustments to clinical research. Nat Med 2021 Feb 25;27(2):191-193. [doi: 10.1038/s41591-020-01216-2] [Medline: 33495603]
52. Lourenco SF, Tasimi A. No participant left behind: conducting science during COVID-19. Trends Cogn Sci 2020 Aug;24(8):583-584 [FREE Full text] [doi: 10.1016/j.tics.2020.05.003] [Medline: 32451239]
53. Dodds S, Hess AC. Adapting research methodology during COVID-19: lessons for transformative service research. JOSM 2020 Aug 10;32(2):203-217. [doi: 10.1108/josm-05-2020-0153]

54. Wekesah FM, Nyakangi V, Onguss M, Njagi J, Bangha M. Comprehensive Sexuality Education in Sub-Saharan Africa. African Population and Health Research Center. 2019. URL: https://aphrc.org/wp-content/uploads/2019/12/COMPREHENSIVE-SEXUALITY-EDUCATION-IN-SUB-SAHARAN-AFRICA-1.pdf [accessed 2022-02-20]
55. Wangamati CK. Comprehensive sexuality education in sub-Saharan Africa: adaptation and implementation challenges in universal access for children and adolescents. Sex Reprod Health Matters 2020 Dec 09;28(2):1851346 [FREE Full text] [doi: 10.1080/26410397.2020.1851346] [Medline: 33295853]
56. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013 Jan 04;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]

Abbreviations

CAB: community advisory board
ELISA: enzyme-linked immunosorbent assay
FGD: focus group discussion
HSV: herpes simplex virus
IRB: institutional review board
ISM: independent safety monitor
KEMRI: Kenya Medical Research Institute
mHealth: mobile health
ODK: Open Data Kit
PPE: personal protective equipment
RCT: randomized controlled trial
SERU: Scientific and Ethics Review Unit
SES: socioeconomic status
SSI: semistructured interviews
WHO: World Health Organization

Edited by G Eysenbach; submitted 22.11.21; peer-reviewed by H Engelbrecht, K Wrona; comments to author 24.12.21; revised version received 10.01.22; accepted 11.01.22; published 03.03.22.

Please cite as:
Mudhune V, Sabben G, Ondenge K, Mbeda C, Morales M, Lyles RH, Arego J, Ndivo R, Bednarczyk RA, Komro K, Winskell K
The Efficacy of a Smartphone Game to Prevent HIV Among Young Africans: Protocol for a Randomized Controlled Trial in the Context of COVID-19
JMIR Res Protoc 2022;11(3):e35117
URL: https://www.researchprotocols.org/2022/3/e35117
doi:10.2196/35117
PMID:35030090

Protocol

Designing and Evaluating a Personalized, Human-Centered Dietary Decision Support System for Use Among People With Diabetes in an Indian Setting: Protocol for a Quasi-Experimental Study

Dinesh Kumar[1*], MD; Ashok Bhardwaj[2*], MD; Shruti Sharma[3*], MSc; Bhavya Malhotra[3*], MSc; Chioma Amadi-Mgbenka[4*], MPH, PhD; Ashoo Grover[5*], MD; Ashish Joshi[4*], MBBS, MPH, PhD

[1]Department of Community Medicine, Himachal Pradesh, India

[2]Dr. Radhakrishnan Government Medical College, Himachal Pradesh, India

[3]Foundation of Healthcare Technologies Society, New Delhi, India

[4]Graduate School of Public Health and Health Policy, City University of New York, New York, NY, United States

[5]Indian Council of Medical Research, New Delhi, India

[*]all authors contributed equally

Corresponding Author:
Ashish Joshi, MBBS, MPH, PhD
Graduate School of Public Health and Health Policy
City University of New York
55W 125th street
New York, NY, 10027
United States
Phone: 1 4435706018
Email: ashish.joshi@sph.cuny.edu

Abstract

Background: Human-centered dietary decision support systems are fundamental to diabetes management, and they address the limitations of existing diet management systems.

Objective: The objective of the proposed study is to evaluate the use of an interactive, telephone-linked, personalized, human-centered decision support system for facilitating the delivery of personalized nutrition care for patients with diabetes.

Methods: A quasi-experimental trial was conducted between the period of June and December 2018. Study participants were recruited from Community Health Center, Dharamshala, Kangra (urban population), and Model Rural Health Unit, Haroli Block, Una (rural population). Eligible participants included adults aged ≥30 years with controlled or uncontrolled diabetes, those who agreed to participate in the study, those who were available for follow-up interviews, and those with a telephone or computer at home. Diabetic status was determined via a physician's diagnosis. Individuals with mental or physical challenges that affected their ability to use an electronic diet record, those who were not available for a telephone follow-up, and those who were involved in other protocols related to dietary assessments were excluded. The study participants were randomized into the following two groups: the intervention group (telephone-linked dietary decision support system) and the control group (paper-based diet record). Study participants in the intervention group recorded their daily dietary intake by using a telephone-linked, personalized, human-centered dietary decision support system and received personalized feedback and diet education via SMS text messaging. Study participants in the control group were provided with only a paper-based diet record for documenting their daily dietary intake. Follow-up visits were conducted at 3 and 6 months from the baseline in both groups. Differences in diabetes knowledge, attitudes, and practices will be measured across groups.

Results: The collection of baseline data from 800 study participants in both the intervention (n=400) and control groups (n=400), which were stratified by urban (control group: n=200; intervention group: n=200) and rural settings (control group: n=200; intervention group: n=200), has been completed. Follow-up data collection for months 3 and 6 is ongoing and is expected to be completed by October 2019.

Conclusions: We anticipate that the intervention group will show significant changes in nutrition knowledge, attitudes, and practices; satisfaction with care; and overall diabetes management. We also expect to see urban-rural differences across the

groups. The uniqueness of our nutrient data capture process is demonstrated by its cultural and contextually relevant features—diet capture in both English and Hindi, diet conversion into caloric components, sustained diet data collection and participant adherence through telephone-linked care, and auto-generated reminders.

International Registered Report Identifier (IRRID): DERR1-10.2196/13635

(JMIR Res Protoc 2022;11(3):e13635) doi:10.2196/13635

KEYWORDS

type 2 diabetes; diabetes management; dietary decision support; diet record; India

Introduction

Type 2 diabetes (T2D) is a rapidly growing chronic health problem, and the complications of T2D cause significant morbidity and mortality. Globally, 415 million people are living with T2D mellitus, and this is estimated to increase to 642 million people by 2040 [1]. The proliferation of T2D is the most notable in low- and middle-income countries, and this has been attributed to a number of factors, including an aging population, population growth, urbanization, the increasing prevalence of obesity, and physical inactivity [2-5]. India, China, and the United States have the highest diabetes incidence rates globally. According to the International Diabetes Federation estimates, the number of patients with diabetes in India is projected to increase to approximately 70 million by 2025—almost double the amount from 2007 [1].

A key outcome of proper diabetes management is the development of healthy eating patterns [3]. In addition, daily physical activity sessions are recommended to regulate blood glycemic levels. [4]. Proper glycemic control is critical to the prevention or delay of the onset of acute and chronic complications and to the improvement of quality of life among people with diabetes. Providing self-management education and support tools to improve diabetes knowledge, foster treatment adherence, promote lifestyle changes, and enable the self-monitoring of blood glucose is fundamental to this process [5,6].

Dietary intake assessment is a fundamental step in developing interventions for people living with T2D [6]. Dietary intake has been measured by using a variety of methods in the existing literature, including dietary recalls, weighed diet records, diet history reports, and food frequency questionnaires [6-8]. None of these methods have demonstrated adequate accuracy and reliability in determining food intake. A variety of technology-based applications have also been developed for estimating and recording food consumption, but significant limitations have been reported, including small sample sizes, insufficient outcome evaluation periods, high costs, and inappropriate message framing [9,10]. Future studies using randomized controlled trial designs, larger sample sizes, and longer durations of evaluation are recommended to establish the intervention capabilities of technology-based diet interventions [10-14].

The objective of the proposed study is to evaluate the use of an interactive, telephone-linked, personalized, human-centered decision support system for facilitating the delivery of personalized nutrition care for patients with diabetes. To the best of our knowledge, this is the first study to design and develop an intervention model after assessing the needs of users from urban and rural settings in an Indian context. The proposed study has three specific aims, as follows:

- Examine and compare the factors that influence the intake of a healthy diet among people with diabetes living in the urban and rural settings of Himachal Pradesh
- Identify the necessary components of the personalized, human-centered dietary decision support system that can gather, analyze, and provide individualized dietary feedback
- Compare the effectiveness of the personalized, human-centered dietary decision support system to that of other paper-based methods of diet recording for documenting nutrient intake among people with diabetes

Methods

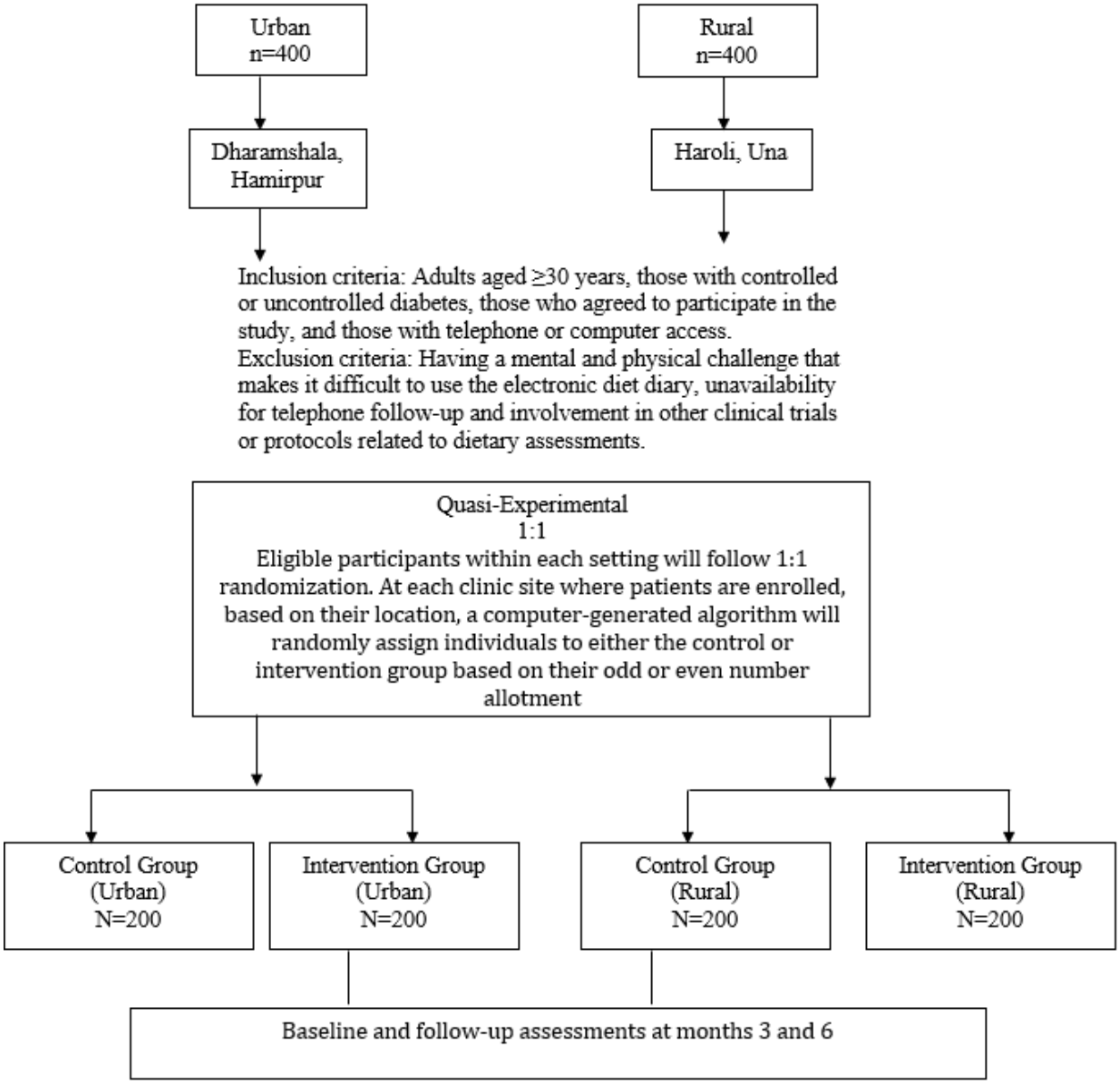
Ethics Approval

Ethical approval was obtained from the Institutional Review Board of Dr. Rajendra Prasad Medical College Kangra (institutional review board number: 116/2016).

Study Design

A quasi-experimental trial was conducted between the period of June and December 2018 (Figure 1). Study participants were recruited from the following two urban and rural locations: Community Health Center, Dharamshala, Kangra (urban population), and Model Rural Health Unit, Haroli Block, Una (rural population). Eligible study participants included (1) adults aged ≥30 years, (2) those with controlled or uncontrolled diabetes, (3) those who agreed to participate in the study, (4) those who were available for follow-up interviews, and (5) those with a telephone or computer at home. Diabetic status was based on a physician-confirmed diagnosis of diabetes and the prescription of medications for diabetes management. The exclusion criteria comprised the following: (1) the presence of any mental or physical challenge affecting the study participants' ability to use an electronic diet record, (2) unavailability for a telephone follow-up, and (3) involvement in other clinical trials or protocols related to dietary assessments. The study was funded by the Indian Council of Medical Research.

Figure 1. Study participant recruitment.



Study Groups

The study participants were randomly assigned into the following two groups: the intervention group (telephone-linked dietary decision support system) and the control group (paper-based diet record; Figure 1). The randomization was conducted following the assessment of participants' eligibility for inclusion in the study. Study participants in the intervention group (group 1) recorded their daily dietary intake by using a telephone-linked, personalized, human-centered dietary decision support system and received personalized feedback and diet education via SMS text messaging. The personalized, human-centered dietary decision support system is a telephone-linked decision support system designed to provide tailored dietary education to patients with diabetes. The study participants were able to access the personalized, human-centered dietary decision support system through their computers, cell phones, or a telephone-linked service, depending on the technology platform available to them. The telephone-linked service enabled the participants (including those who did not have a technology platform to access the personalized, human-centered dietary decision support system) to receive phone calls from our research team in order to record their daily dietary intake. The personalized, human-centered dietary decision support system also generated automatic alerts that served as reminders to the participants who had not reported their daily diet intake. Study participants in the control group were provided with a paper-based diet record for documenting their daily dietary intake. Both the intervention group and control group were provided with a diabetes educational booklet at baseline. However, the study participants in the control group were also provided with a booklet to document their diet log on a daily basis. Follow-up visits were conducted at 3 and 6 months from the baseline in both groups. Baseline data were gathered on sociodemographics; health literacy; physical activity; anthropometric measurements; blood sugar testing;

and diabetes knowledge, attitudes, and practices (KAPs). Follow-up data were gathered on anthropometric measurements, diabetes KAPs, blood sugar testing, and satisfaction with medical care.

Intervention Development

A telephone-based, personalized, human-centered dietary decision support system for populations with diabetes was designed to provide tailored dietary education to patients with diabetes (Figure 2). A human-centered approach was used in the design process. Human-centered design principles require end users to be prioritized in intervention design and development [15]. User characteristics, needs, and preferences are vital for ensuring the optimal use of technology-enabled interventions [15] (Figure 3). The failure to fully implement human-centered design in dietary intervention designs has been a limitation in prior studies [11]. The personalized, human-centered dietary decision support system is comprised of the following components: (1) the electronic monitoring of diet records, (2) personalized dietary and disease feedback, (3) the nutrient calculation of daily diet intake per meal, and (4) alerts and reminders (Figure 4).

The personalized, human-centered dietary decision support system can be used on cell phones, PDAs, and computers, depending on the technology platforms available to users across various Indian settings. Dietary information from the prospective users was gathered on a daily basis. This information included the various food sources consumed and the quantity and timing of food consumption. A nutrient database of food choices and their caloric values was subsequently created [16]. This nutrient database included the following components: food type, food class, quantity, and calorie estimation. A message library in both the English and Hindi languages (a local Indian dialect) was prepared, so that individuals could be sent weekly messages on how best to manage their diet. Based on the study participants' dietary intake, interactive tailored feedback was provided (Figure 5). Study participants also received auto-generated reminders and alerts if their dietary information was not received as scheduled.

Figure 2. Dietary decision support system workflow.

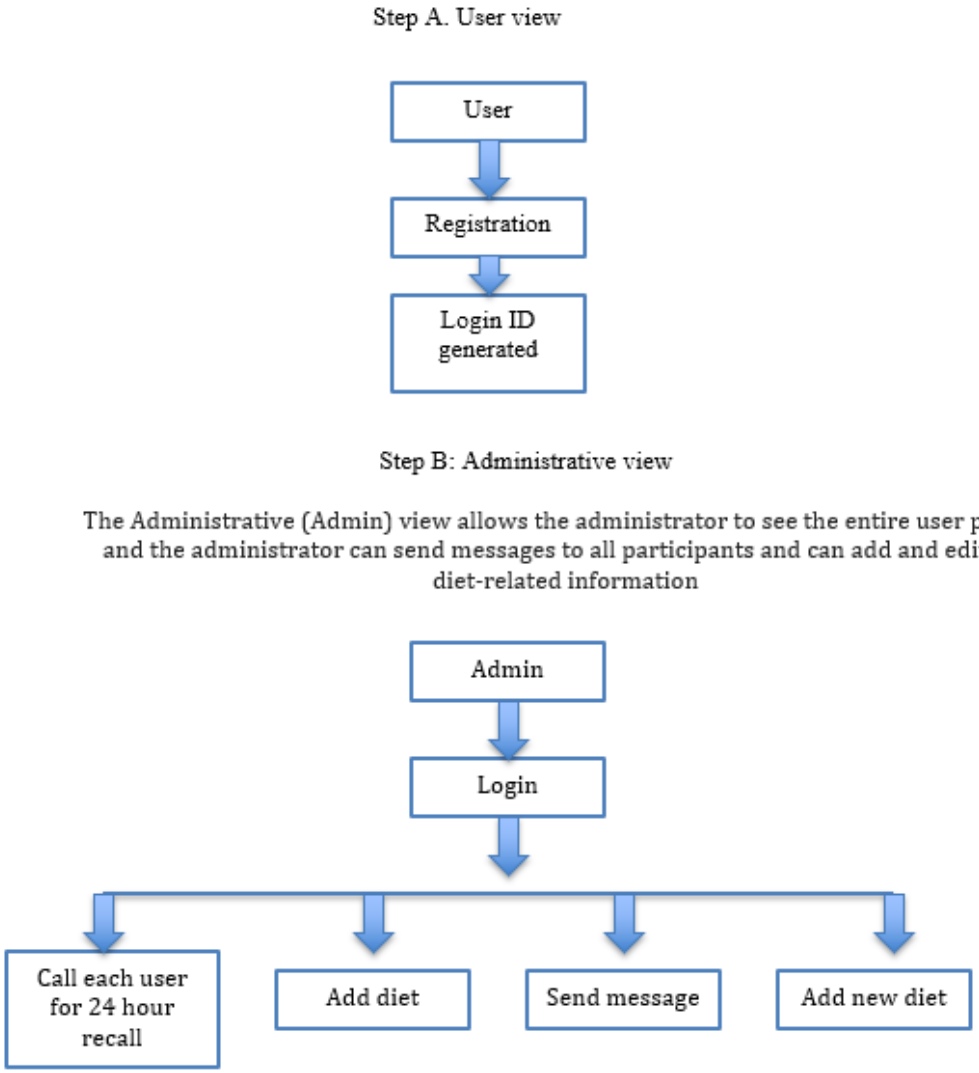


Figure 3. Administrative control panel: entry of study participant demographics.

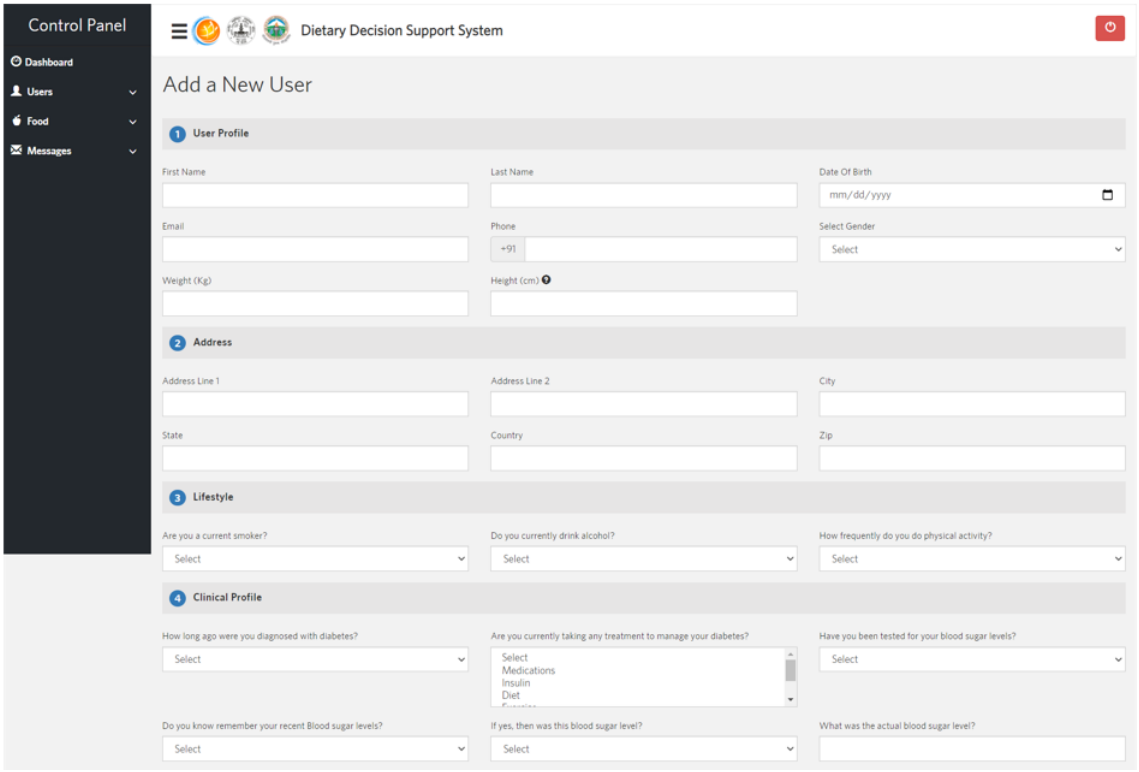


Figure 4. Administrative control panel: addition of food categories.

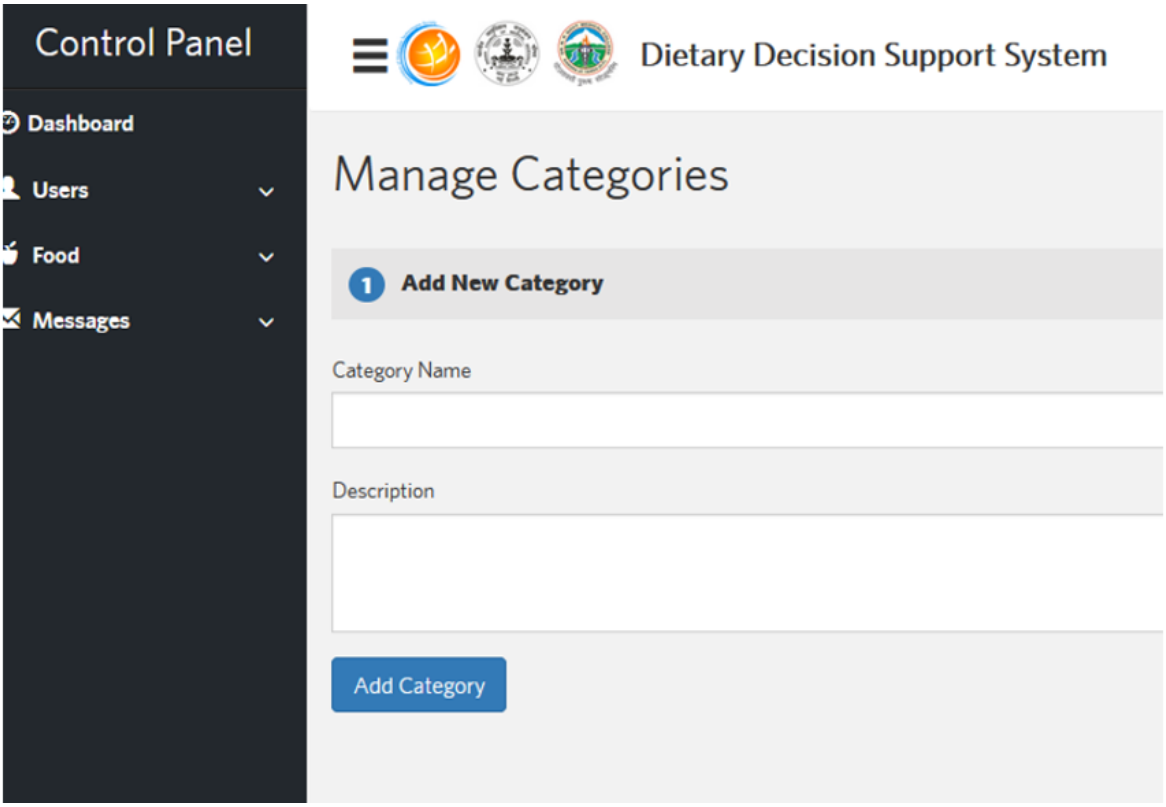
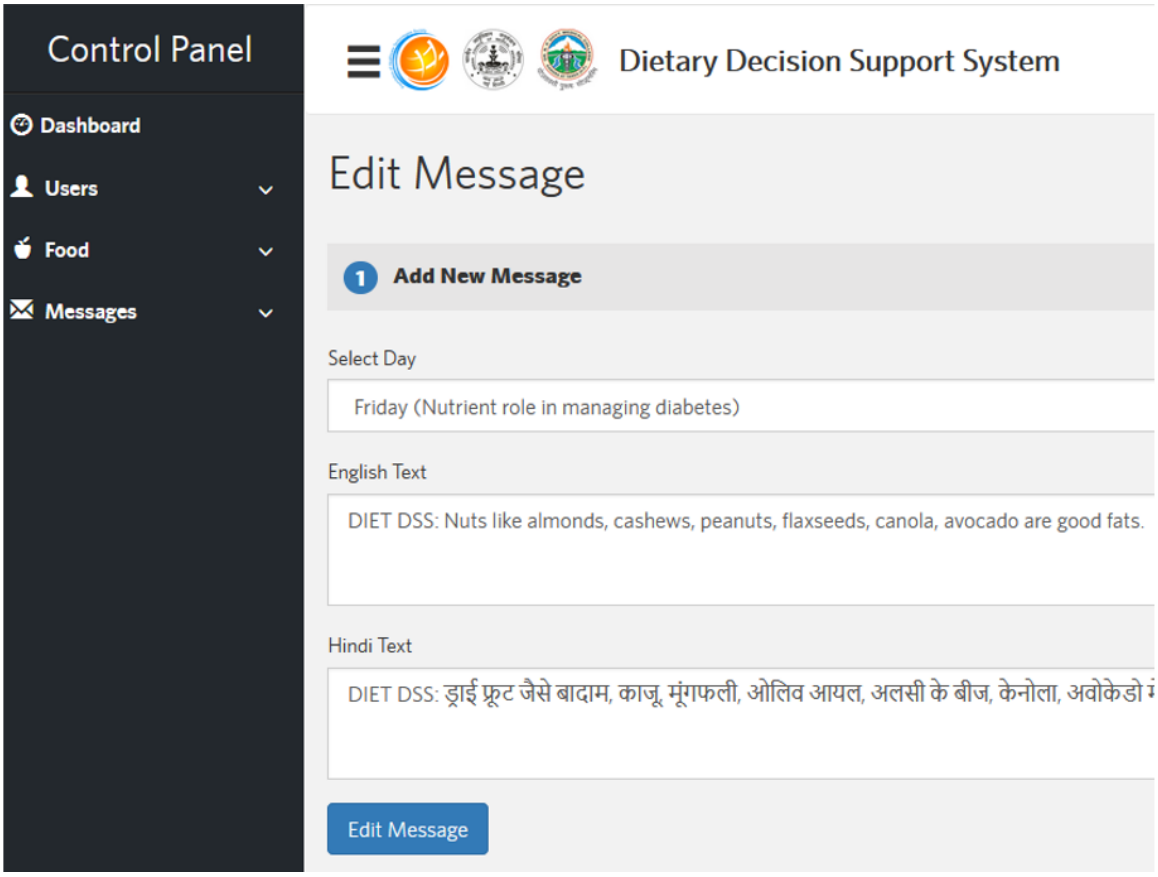


Figure 5. Administrative control panel: generation of personalized dietary feedback.



Study Enrollment and Data Collection Procedure

Data were collected at baseline, and follow-ups were conducted at months 3 and 6, resulting in 4 data collection time points. After each time point, the field-workers scheduled follow-up interviews to administer the study questionnaires. The study staff who were collecting data were blinded to the group assignments.

Informed Consent

Informed consent forms were administered to the eligible individuals, and those who consented were enrolled in the study. The institutional review board–approved consent forms were administered by members of our research team to the eligible individuals. These forms described the study, the measures used by the researchers to protect the confidentiality of the responses, and the voluntary nature of the study. The consent forms were countersigned by members of the research team, and copies were provided to the study participants.

Data Entry and Quality Assurance

Data entry was performed by a team of field-workers and data management personnel. To ensure efficiency and high-quality data collection and processing, we (1) used a well-trained team of field-workers, (2) used a clearly defined study manual, (3) conducted weekly meetings with the research team, and (4) maintained logs of all patient contacts. To ensure efficient and accurate data management, we maintained (1) logs of all of the data instruments that were filled during each visit for every patient, (2) central data processing, and (3) weekly data checks. Data security was ensured through regular backups, password protection, and storage in a locked file cabinet.

Variables Assessed

Sociodemographics

Baseline data were gathered on study participants' age, income level, employment status, education level, smoking status, and alcohol consumption. Information was also collected on computer usage, internet usage, the frequency of computer and internet usage, and sources of health information.

Health Literacy

Health literacy is defined as "the ability to perform basic reading and numerical tasks required to function in the health care environment" [17]. Health literacy was assessed by using a 3-item health literacy screening questionnaire [17]. The questions included the following: (1) "How often do you have someone (like a family member, friend, hospital/clinic worker or caregiver) help you read hospital materials," (2) "How often do you have problems learning about your medical condition because of difficulty understanding written information" (problems reading), and (3) "How confident are you filling out forms by yourself" (confident with forms)? Responses were rated on a Likert scale ranging from 0 to 4 and included the following options: "all of the time," "most of the time," "some of the time," "a little of the time," and "none of the time" [17].

Anthropometry

Height, weight, and waist circumference were measured by using a standard technique [18]. BMIs were computed from the height and weight measurements.

Physical Activity Assessment

All participants completed a validated short form of the International Physical Activity Questionnaire in order to calculate the total time that they spent on performing various forms of physical activity—recreational activities, occupational activities, household work, and transportation-related activities—in the last 7 days [19]. Total weekly physical activity—metabolic equivalents of task (METs; MET hours per week)—will be calculated as the weighted sum of the reported time spent at each intensity by using a MET value specific to each category (walking: 3.3 METs; moderate: 4 METs; vigorous: 7 METs) [19].

Outcomes

The study outcomes include dietary KAPs. These outcomes were assessed by using the following tools. Satisfaction with medical care was also assessed.

Nutrition Knowledge Scale

The Nutrition Knowledge Scale comprises the following four types of items: (1) 10 items on the relationship between diet and disease, (2) 10 items on food comparisons in terms of their nutrient content (eg, fat, fiber, calcium, calories, and sodium); (3) 6 items on the daily serving requirements of different food groups; and (4) 5 items on weight and weight loss. The scale is in a multiple-choice format, and 1 point is awarded for correct answers; otherwise, 0 points are awarded. The interitem reliability (Cronbach α coefficient) of the Nutrition Knowledge Scale is .78 [20].

Nutrition Attitude Scale

The Nutrition Attitude Scale consists of 19 items and the following three sections: (1) care about nutrition (8 items), (2) emotional eating (6 items), and (3) the importance of nutrition (5 items). The response options are rated on a 5-point Likert scale and include “strongly disagree,” “disagree,” “neutral,” “agree,” and “strongly agree.” Reverse sentence items are scored reversely. The Cronbach α coefficient of the Nutrition Attitude Scale is .73 [20].

Nutrition Behavior Scale

The Nutrition Behavior Scale is a 24-item scale with responses on a 5-point Likert scale. It comprises the following two sections: (1) food selection and care about nutrition (15 items) and (2) emotionally and externally cued eating (9 items). The response items include “never,” “seldom,” “sometimes,” “often,” and “always,” ranging in score from 1 to 5. Reverse sentence items are scored reversely. The Cronbach α coefficient of the Nutrition Behavior Scale is .76 [20].

Satisfaction With Medical Care

Satisfaction with medical care will be measured by using the Client Satisfaction Questionnaire-8 (CSQ-8). The CSQ-8 is an 8-item questionnaire in which items are rated by using a 4-point Likert scale [21].

Blood Sugar Testing

We will also assess participants’ blood sugar levels at baseline and at 6 months to assess the pattern of blood sugar control.

Data Analysis Plan

In the first step of the exploratory analysis, we will present a table to summarize the overall characteristics of all variables, including outcome, confounding, and process variables. This table will also serve as quality control for the original data and be used to mine missing data patterns and outliers. The distributions of the continuous outcomes will be explored for normality, and transformations will be used if necessary. Continuous variables will be summarized using means, medians, SDs, and ranges, while categorical variables will be examined by using frequencies and percentages. An exploratory analysis on process variables will be used for a post hoc analysis. As a third step, we will cross-tabulate the group assignments against outcome variables to assess the association in the initial stage of the study.

Sample Size Justification

The sample size calculation was based on knowledge, attitude, and behavior outcomes, as reported in previous studies [22]. To detect a mean change of 0.2 with an SD of 0.7 at a 2-sided .05 α level of significance and a power of 85%, a sample size of 200 individuals in each group is needed. The total participant sample size will be 800.

Statistical Analysis

Baseline characteristics will be summarized and stratified by intervention group. The primary outcomes—adherence to dietary guidelines, KAPs, and patient satisfaction with medical care—will be compared between groups by using a log-rank test. A Cox proportional hazards regression model will be used to adjust for confounders that differ significantly between the two groups at the .20 level of significance. A generalized linear mixed model using a binomial distribution and logit link function will be used to examine the binary secondary objective—subjects’ satisfaction with medical care. The group types (ie, the intervention and control groups) will be used as a fixed effect, and subject data will be used as random effects. This will be done by using a variance components covariance structure. To adjust for confounders that differ significantly between the two groups at the .20 level of significance, covariates will be added to the model, and a linear regression model will be used.

Results

Data collection was conducted between the period of June and December 2018. The collection of baseline data from 800 study participants in both the intervention (n=400) and control groups (n=400), which were stratified by urban (control group: n=200; intervention group: n=200) and rural settings (control group: n=200; intervention group n=200), has been completed. Follow-up data collection for months 3 and 6 is ongoing. We expect the follow-up data collection to be completed by October 2019. We are currently analyzing the baseline data and

generating reports, which will be presented in our upcoming manuscript.

Discussion

We anticipate that our results will show a significant difference in nutrition KAPs between the study groups. The study participants receiving tailored dietary education will likely show better improvements in their KAPs and overall diabetes management. We also anticipate that satisfaction with care (measured by the CSQ-8) will be higher in the intervention group, owing to the unique features of our interactive, telephone-linked, personalized, human-centered decision support system (eg, personalized dietary education and reminder alerts for completing the food log). We also expect to see urban-rural differences across the intervention and control groups. Our dietary decision support system is the first of its kind to be successfully implemented in a lower-middle–income country such as India, and it has the potential to inform dietary data management on a wider scope. Key gaps in diet-related interventions are centered on the tailoring, sustainability, and evaluation of such interventions [15,23]. Studies that were conducted to examine the effectiveness of existing diet apps highlighted the need for tailoring approaches that can address the personal needs of the users [23]. Another crucial gap in the design approaches of existing diet apps is the challenge of identifying which specific components are associated with intervention effectiveness. The task of identifying specific components associated with effectiveness becomes an arduous one when human-centered approaches are not effectively incorporated into the design process. A systematic review investigating the effectiveness of eHealth and mobile health interventions that promote physical activity and healthy diets across 13 low- and middle-income countries showed that it was not possible to identify the specific components associated with app effectiveness among the included studies [23]. A human-centered design approach is needed to address these challenges. The uniqueness of our nutrient data capture process is demonstrated by its incorporation of human-centered, cultural, and contextually relevant features—diet capture in both English and Hindi, diet conversion into caloric components, sustained diet data collection and participant adherence through telephone-linked care, and auto-generated reminders. We conducted focus groups with the study population in addition to literature reviews, with the aim of identifying user characteristics, needs, and preferences that potentially influence users' satisfaction with and use of the interactive, telephone-linked, personalized, human-centered decision support system [24-26].

Conflicts of Interest

None declared.

References

1. International Diabetes Federation. IDF Diabetes Atlas, Eighth Edition 2017. Brussels, Belgium: International Diabetes Federation; 2017.
2. Dehghan M, Mente A, Zhang X, Swaminathan S, Li W, Mohan V, Prospective Urban Rural Epidemiology (PURE) study investigators. Associations of fats and carbohydrate intake with cardiovascular disease and mortality in 18 countries from five continents (PURE): a prospective cohort study. Lancet 2017 Nov 04;390(10107):2050-2062. [doi: 10.1016/S0140-6736(17)32252-3] [Medline: 28864332]
3. Muralidharan S, Mohan V, Anjana RM, Jena S, Tandon N, Allender S, et al. Mobile health technology (mDiab) for the prevention of type 2 diabetes: Protocol for a randomized controlled trial. JMIR Res Protoc 2017 Dec 12;6(12):e242 [FREE Full text] [doi: 10.2196/resprot.8644] [Medline: 29233806]
4. Haas K, Hayoz S, Maurer-Wiesner S. Effectiveness and feasibility of a remote lifestyle intervention by dietitians for overweight and obese adults: Pilot study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e12289 [FREE Full text] [doi: 10.2196/12289] [Medline: 30973338]
5. Little M, Humphries S, Patel K, Dewey C. Decoding the type 2 diabetes epidemic in rural India. Med Anthropol 2017;36(2):96-110 [FREE Full text] [doi: 10.1080/01459740.2016.1231676] [Medline: 27644458]
6. Schembre SM, Liao Y, O'Connor SG, Hingle MD, Shen SE, Hamoy KG, et al. Mobile ecological momentary diet assessment methods for behavioral research: Systematic review. JMIR Mhealth Uhealth 2018 Nov 20;6(11):e11170 [FREE Full text] [doi: 10.2196/11170] [Medline: 30459148]
7. Béjar LM, Reyes OA, García-Perea MD. Electronic 12-hour dietary recall (e-12HR): Comparison of a mobile phone app for dietary intake assessment with a food frequency questionnaire and four dietary records. JMIR Mhealth Uhealth 2018 Jun 15;6(6):e10409 [FREE Full text] [doi: 10.2196/10409] [Medline: 29907555]
8. Affret A, El Fatouhi D, Dow C, Correia E, Boutron-Ruault MC, Fagherazzi G. Relative validity and reproducibility of a new 44-item diet and food frequency questionnaire among adults: Online assessment. J Med Internet Res 2018 Jul 05;20(7):e227 [FREE Full text] [doi: 10.2196/jmir.9113] [Medline: 29980502]
9. Ernsting C, Dombrowski SU, Oedekoven M, Sullivan JLO, Kanzler M, Kuhlmey A, et al. Using smartphones and health apps to change and manage health behaviors: A population-based survey. J Med Internet Res 2017 Apr 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]

10. Burrows TL, Rollo ME, Williams R, Wood LG, Garg ML, Jensen M, et al. A systematic review of technology-based dietary intake assessment validation studies that include carotenoid biomarkers. Nutrients 2017 Feb 14;9(2):140 [FREE Full text] [doi: 10.3390/nu9020140] [Medline: 28216582]
11. Coughlin S, Thind H, Liu B, Champagne N, Jacobs M, Massey RI. Mobile phone apps for preventing cancer through educational and behavioral interventions: State of the art and remaining challenges. JMIR Mhealth Uhealth 2016 May 30;4(2):e69 [FREE Full text] [doi: 10.2196/mhealth.5361] [Medline: 27242162]
12. Willmott TJ, Pang B, Rundle-Thiele S, Badejo A. Weight management in young adults: Systematic review of electronic health intervention components and outcomes. J Med Internet Res 2019 Feb 06;21(2):e10265 [FREE Full text] [doi: 10.2196/10265] [Medline: 30724736]
13. Jimoh F, Lund EK, Harvey LJ, Frost C, Lay WJ, Roe MA, et al. Comparing diet and exercise monitoring using smartphone app and paper diary: A two-phase intervention study. JMIR Mhealth Uhealth 2018 Jan 15;6(1):e17 [FREE Full text] [doi: 10.2196/mhealth.7702] [Medline: 29335239]
14. Paulsen MM, Hagen MLL, Frøyen MH, Foss-Pedersen RJ, Bergsager D, Tangvik RJ, et al. A dietary assessment app for hospitalized patients at nutritional risk: Development and evaluation of the MyFood app. JMIR Mhealth Uhealth 2018 Sep 07;6(9):e175 [FREE Full text] [doi: 10.2196/mhealth.9953] [Medline: 30194059]
15. Joshi A, Amadi C, Katz B, Kulkarni S, Nash D. A human-centered platform for HIV infection reduction in New York: Development and usage analysis of the Ending the Epidemic (ETE) Dashboard. JMIR Public Health Surveill 2017 Dec 11;3(4):e95 [FREE Full text] [doi: 10.2196/publichealth.8312] [Medline: 29229592]
16. Longvah T, Ananthan R, Bhaskarachary K, Venkaiah K. Indian food composition tables. Biology Education. URL: https://tinyurl.com/ypkurzwt [accessed 2022-02-16]
17. 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]
18. Sánchez-García S, García-Peña C, Duque-López MX, Juárez-Cedillo T, Cortés-Núñez AR, Reyes-Beaman S. Anthropometric measures and nutritional status in a healthy elderly population. BMC Public Health 2007 Jan 03;7:2 [FREE Full text] [doi: 10.1186/1471-2458-7-2] [Medline: 17201919]
19. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
20. Lin W, Hang CM, Yang HC, Hung MH. 2005-2008 Nutrition and Health Survey in Taiwan: the nutrition knowledge, attitude and behavior of 19-64 year old adults. Asia Pac J Clin Nutr 2011;20(2):309-318 [FREE Full text] [Medline: 21669600]
21. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann 1982;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-x] [Medline: 10259963]
22. Fearon E, Chabata ST, Thompson JA, Cowan FM, Hargreaves JR. Sample size calculations for population size estimation studies using multiplier methods with respondent-driven sampling surveys. JMIR Public Health Surveill 2017 Sep 14;3(3):e59 [FREE Full text] [doi: 10.2196/publichealth.7909] [Medline: 28912117]
23. West JH, Hall PC, Arredondo V, Berrett B, Guerra B, Farrell J. Health behavior theories in diet apps. J Consum Health Internet 2013 Feb 15;17(1):10-24. [doi: 10.1080/15398285.2013.756343]
24. Wang Q, Egelandsdal B, Amdam GV, Almli VL, Oostindjer M. Diet and physical activity apps: Perceived effectiveness by app users. JMIR Mhealth Uhealth 2016 Apr 07;4(2):e33 [FREE Full text] [doi: 10.2196/mhealth.5114] [Medline: 27056639]
25. Müller AM, Alley S, Schoeppe S, Vandelanotte C. Int J Behav Nutr Phys Act 2016 Oct 10;13(1):109 [FREE Full text] [doi: 10.1186/s12966-016-0434-2] [Medline: 27724911]
26. Valdovinos MG, Schieber E, McMahon M, Beard L, Wilkinson A, Carpenter J. Adverse side effects of psychotropic medication and challenging behavior: Pilot work assessing impact. J Dev Phys Disabil 2017 Dec;29(6):969-982 [FREE Full text] [doi: 10.1007/s10882-017-9570-0] [Medline: 29398883]

Abbreviations

CSQ-8: Client Satisfaction Questionnaire-8
KAP: knowledge, attitude, and practice
MET: metabolic equivalent of task
T2D: type 2 diabetes

Edited by G Eysenbach; submitted 05.02.19; peer-reviewed by N Fischer, V Mohan; comments to author 08.05.19; revised version received 28.05.19; accepted 29.05.19; published 08.03.22.

Please cite as:
Kumar D, Bhardwaj A, Sharma S, Malhotra B, Amadi-Mgbenka C, Grover A, Joshi A
Designing and Evaluating a Personalized, Human-Centered Dietary Decision Support System for Use Among People With Diabetes in an Indian Setting: Protocol for a Quasi-Experimental Study
JMIR Res Protoc 2022;11(3):e13635
URL: https://www.researchprotocols.org/2022/3/e13635
doi:10.2196/13635
PMID:35258472

Protocol

Viability of an Early Sleep Intervention to Mitigate Poor Sleep and Improve Well-being in the COVID-19 Pandemic: Protocol for a Feasibility Randomized Controlled Trial

Kathleen Patricia O'Hora[1], BSc; Raquel A Osorno[1], PsyD; Dena Sadeghi-Bahmani[1,2], PhD; Mateo Lopez[1], BA; Allison Morehouse[1], BA; Jane P Kim[1], PhD; Rachel Manber[1], PhD; Andrea N Goldstein-Piekarski[1,3], PhD

[1]Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, United States

[2]Department of Psychology, Stanford University, Stanford, CA, United States

[3]Sierra-Pacific Mental Illness Research, Education, and Clinical Center (MIRECC), Veterans Affairs Palo Alto Health Care System, Palo Alto, CA, United States

Corresponding Author:
Andrea N Goldstein-Piekarski, PhD
Psychiatry and Behavioral Sciences
Stanford University
401 Quarry Road
Stanford, CA, 94304
United States
Phone: 1 (650) 721 4780
Email: agoldpie@stanford.edu

Abstract

Background: The COVID-19 pandemic has led to drastic increases in the prevalence and severity of insomnia symptoms. These increases in insomnia complaints have been paralleled by significant decreases in well-being, including increased symptoms of depression, anxiety, and suicidality and decreased quality of life. However, the efficacy and impact of early treatment of insomnia symptoms on future sleep and well-being remain unknown.

Objective: Here, we present the framework and protocol for a novel feasibility, pilot study that aims to investigate whether a brief telehealth insomnia intervention targeting new insomnia that developed during the pandemic prevents deterioration of well-being, including symptoms of insomnia, depression, anxiety, suicidality, and quality of life.

Methods: The protocol details a 2-arm randomized controlled feasibility trial to investigate the efficacy of a brief, telehealth-delivered, early treatment of insomnia and evaluate its potential to prevent deterioration of well-being. Participants with clinically significant insomnia symptoms that began during the pandemic were randomized to either a treatment group or a 28-week waitlist control group. Treatment consists of 4 telehealth sessions of cognitive behavioral therapy for insomnia (CBT-I) delivered over 5 weeks. All participants will complete assessments of insomnia symptom severity, well-being, and daily habits checklist at baseline (week 0) and at weeks 1-6, 12, 28, and 56.

Results: The trial began enrollment on June 3, 2020 and closed enrollment on June 17, 2021. As of October 2021, 49 participants had been randomized to either immediate treatment or a 28-week waitlist; 23 participants were still active in the protocol.

Conclusions: To our knowledge, this protocol would represent the first study to test an early sleep intervention for improving insomnia that emerged during the COVID-19 pandemic. The findings of this feasibility study could provide information about the utility of CBT-I for symptoms that emerge in the context of other stressors before they develop a chronic course and deepen understanding of the relationship between sleep and well-being.

Trial Registration: ClinicalTrials.gov NCT04409743; https://clinicaltrials.gov/ct2/show/NCT04409743

International Registered Report Identifier (IRRID): DERR1-10.2196/34409

(JMIR Res Protoc 2022;11(3):e34409) doi:10.2196/34409

KEYWORDS

insomnia; COVID-19; pandemic; telehealth; cognitive behavioral therapy; CBT-I; sleep; depression; well-being; telemedicine; impact; mental health; therapy

Introduction

Background

The COVID-19 pandemic and resulting mass home confinement have led to a significant increase in insomnia complaints [1-3]. Recent studies from around the world report the prevalence of moderate to severe insomnia symptoms during the pandemic to range from 17.4% to 33.7% [4-6], compared with 10% to 15% prepandemic [7,8]. The prevalence rate is even higher among health care workers, individuals with medical comorbidities, and individuals in close contact with the virus [9,10]. Many of these insomnia symptoms may have developed as a direct result of the pandemic and home confinement [11]. For example, increases in sleep disturbances were directly influenced by the increase in COVID-19–related deaths, supporting the relationship between pandemic severity and insomnia severity [12]. Additionally, the increase in insomnia complaints may also be in part due to the exacerbation of several established risk factors for insomnia emerging from the COVID-19 pandemic including increased loneliness, perceived stress, and screen time, together with decreased social connection and physical activity [13-16].

The observed increases in insomnia complaints are paralleled by significant decreases in well-being, including increased depression, anxiety, and suicidality and decreased quality of life [17-21]. Critically, prior research in a nonpandemic environment found that insomnia is not only highly comorbid with depression and anxiety but also a strong predictive factor in their development and prognosis [22,23]. Given these findings, it is possible that the increase in insomnia complaints following the pandemic may be directly contributing to the increased prevalence of depression and anxiety during the pandemic [17,18]. Thus, treating insomnia symptoms early may be one approach to improve well-being and prevent future depression during the COVID-19 pandemic.

Cognitive behavioral therapy for insomnia (CBT-I) is the gold standard, first-line, nonpharmacological treatment for chronic insomnia recommended by the American College of Physicians [24]. It has been proven safe and effective in adults across the lifespan, in individual [25] and group formats [26] and when delivered in person or via telehealth [27]. A recently published article [28] from the European Academy for CBT-I provided evidence supporting the efficacy of CBT-I to treat sudden-onset insomnia and the validity of telehealth-delivered CBT-I. There is evidence that CBT-I not only reduces sleep complaints in those with chronic insomnia but can also reduce depression, anxiety, and suicidality and improve quality of life [29-33]. There is further evidence demonstrating that the resulting improvements in insomnia symptoms mediate the changes in depression symptoms but not the reverse [34]. Although the literature on this is limited, taken together, this evidence suggests that providing a brief, telehealth-delivered CBT-I to individuals with newly developed, pandemic-onset insomnia complaints may not only improve sleep but also improve well-being. However, to our knowledge, there are no studies investigating whether CBT-I would be an effective early treatment for insomnia arising from a highly disruptive and stressful event, such as a global pandemic, or whether intervening early in insomnia symptom onset could help mitigate other negative mental health outcomes. Additionally, it remains unknown whether common risk factors (eg, loneliness, perceived stress, and screen time) for insomnia and poor well-being that have been exacerbated by the pandemic may impact the effectiveness of an early intervention for insomnia symptoms.

Typically, individuals with sleep disturbance do not seek treatment unless their condition develops into chronic insomnia. This delay in seeking insomnia treatment makes parsing temporal, mechanistic relationships between insomnia and well-being nearly impossible. However, lifestyle changes and stress associated with the COVID-19 pandemic created large-scale disturbances in psychological well-being and sleep. These circumstances provided a novel opportunity to study the relationship between sleep and well-being by deploying an early sleep intervention to treat insomnia symptoms that have not yet developed into chronic insomnia. These circumstances have thus provided a unique opportunity through a pilot study to respond to a public health crisis and explore the temporal interrelationship between new sleep disturbances and deterioration in well-being as well as to assess whether intervening early in sleep disturbances is enough to alter these trajectories. The findings of the pilot study, the protocol of which is detailed in this paper, will be integral for guiding larger-scale trials in nonpandemic settings.

Objectives

Our feasibility, pilot study investigates the viability of an early treatment for insomnia symptoms to treat insomnia symptoms arising during the COVID-19 pandemic and determine pandemic-related risk factors for worsening well-being and sleep outcomes. We also assess the impact of the intervention on insomnia severity and well-being across 28 weeks.

We will accomplish these objectives by conducting a waitlist-controlled trial across 28 weeks to address 3 aims.

Aim 1 is to determine whether a brief, telehealth CBT-I reduces insomnia symptoms arising during the COVID-19 pandemic. We hypothesize that CBT-I will lead to improvements in insomnia severity, as measured by the Insomnia Severity Index (ISI) [35,36], across 28 weeks. We also hypothesize that fewer participants in the CBT-I group will meet the Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-5) criteria for insomnia disorder, compared with the waitlist control group at weeks 12 and 28.

Aim 2 is to determine whether brief, telehealth CBT-I mitigates negative mental health outcomes arising during the COVID-19 pandemic. We hypothesize that, compared with the waitlist control group, the CBT-I group will have an improved trajectory of well-being across 28 weeks and will have better well-being at weeks 7, 12, and 28. Additionally, we hypothesize improvements in insomnia symptoms will mediate improvements in well-being from baseline to weeks 6, 12, and 28.

Aim 3 is to determine whether risk factors for insomnia that might be aggravated during the COVID-19 pandemic predict worse insomnia and negative mental health outcomes at

follow-up. We hypothesize that self-reported high levels of social isolation, perceived stress, sleep reactivity, and screen time and low physical activity at baseline will predict worse long-term outcomes at 12 and 28 weeks across both study groups.

Methods

Trial Design

Overall Design

We designed a 2-arm randomized controlled feasibility trial to investigate the efficacy of an early, brief, telehealth-delivered insomnia treatment to prevent adverse sleep and well-being outcomes. Participants with clinically significant insomnia symptoms (current ISI total score ≥10) that began during the pandemic were randomized to either a treatment group or a waitlist control group. Treatment consists of 4 telehealth sessions of CBT-I delivered over 5 weeks. Participants in the waitlist control group do not receive any study interventions during the 28-week primary assessment period. All participants complete assessments of insomnia symptom severity, depressive symptom severity, anxiety symptom severity, quality of life, and pandemic-related risk factors at baseline (week 0), weeks 1-6, week 12, week 28, and week 56 (Figure 1). Primary outcomes focus on weeks 6, 12, and 28.

This study design creates 2 study phases: a 28-week waitlist-controlled (primary assessment) period, which allows the assessment of the therapy compared with a treatment-naive group and a delayed-start (secondary assessment) period, during which participants originally assigned to the waitlist control group receive the study therapy and participants assigned to the treatment group no longer receive study treatment (Figure 2). This allows assessment of long-term changes in sleep and well-being between individuals who underwent an early behavioral intervention for insomnia and those who did not. Our innovative approach will help elucidate mechanistic pathways between sleep and well-being as well as provide a needed clinical response to the COVID-19 pandemic and resulting mental health crisis.

Figure 1. Study flow for each treatment group from prescreening through the 56-week follow-up, with primary outcome time points occurring at weeks 6, 12, and 28. CBT-I: cognitive behavioral therapy for insomnia.

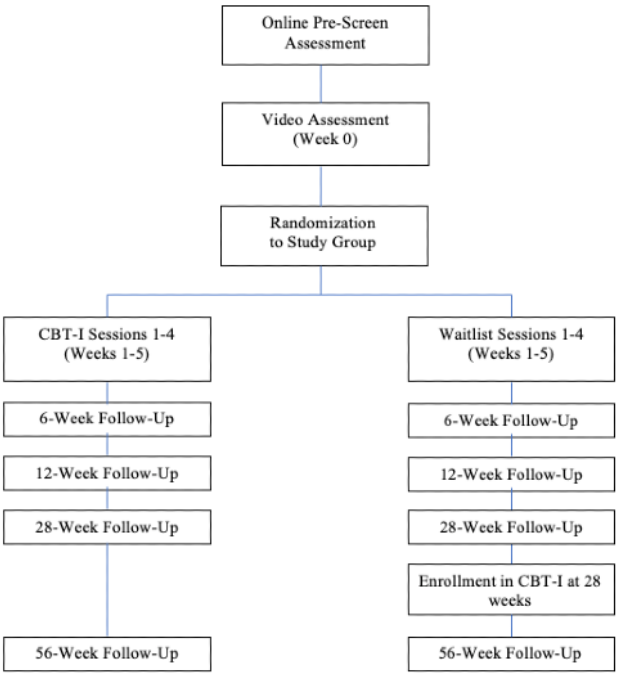
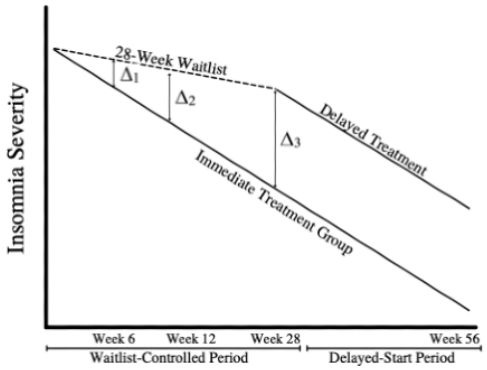


Figure 2. Study design in which the period between week 0 to week 28 is the waitlist-controlled period in which the immediate treatment group (cognitive behavioral therapy for insomnia [CBT-I]) can be directly compared with the waitlist control group (no CBT-I). At 28 weeks, participants in the waitlist group begin therapy and become the delayed treatment group.



Randomization

Participants were randomized into either CBT-I or a 28-week waitlist control condition using stratification by biological sex at birth (male, female) and baseline insomnia severity (ISI ≥10, ISI <10) in a 1:1 ratio. Participants who dropped out of the study after randomization, but before the week 1 visit, were replaced in the randomization matrix. Our sample size estimates (see the Power Calculation section) accounted for this replacement.

Power Calculation

Due to the anticipated difficulties in recruitment that will likely arise from the complexities of running a study during a pandemic and the time-sensitive nature of implementing an early sleep intervention relative to both the start of the pandemic and insomnia symptoms, we view this trial as a feasibility study. Therefore, for this feasibility trial, we aimed to recruit a total of 50 subjects, which would result in 25 subjects in each study group. With this sample size, our power calculations were derived using the 2-group *t* test of equal means based on the difference between CBT-I and waitlist control group at 28 weeks, at a 5% level of significance. With 25 per group, the study is 80% powered to detect a large effect size (Cohen d=0.80).

Intervention

CBT-I is a comprehensive, multimodal approach that addresses maladaptive cognitions and behaviors that contribute to and maintain sleep difficulties. Treatment consists of education about the 2-process model of sleep (the homeostatic and circadian processes [37]) and their interaction with hyperarousal. Behavioral components of treatment include time-in-bed restriction (also known as sleep restriction [38]), stimulus control [39], and relaxation techniques. The cognitive component of treatment includes identification and modification of maladaptive beliefs and thought patterns about sleep to reduce sleep-related anxiety.

The study treatment protocol was adapted from Edinger's 4-session, open-source CBT-I manual [40]. Treatment was "front-loaded," in that Session 1 includes sleep education, stimulus control, and an initial time-in-bed restriction, based on sleep diary data collected in the 1 to 2 weeks prior to the first session. Sessions 2 through 4 are dedicated to adjusting time-in-bed based on ongoing sleep diary data, addressing treatment adherence issues, cognitive therapy, and relaxation skills. Session 4 also includes information about relapse prevention. Treatment is administered by a licensed psychologist or a doctorate-level graduate student trained and supervised by a licensed psychologist. The treatment is outlined in Table 1.

Table 1. Session by session outline of brief, telehealth cognitive behavioral therapy for insomnia (CBT-I).

Week	Session	Time	Content
1	1	60 minutes	• Review the sleep log and answers provided on the brief sleep assessment. • Educate about sleep and basic sleep hygiene instructions. • Introduce two-process model of sleep (circadian rhythm and sleep drive) and interfering role of arousal. • Determine standard wake time and initial time in bed (TIB) prescription. • Provide stimulus control instructions. • Answer questions and address concerns. • Assign homework.
2	2	30-45 minutes	• Review sleep log and adjust TIB prescriptions. • Encourage/reinforce adherence. • Identify/troubleshoot participant's problems in adhering to recommended changes in sleep behaviors. • Address sleep effort and sleep-related anxiety. • Review role of arousal and teach relaxation technique. • Answer questions and address concerns. • Assign homework.
3 or 4	3	30-45 minutes	• Review sleep log and adjust TIB prescriptions. • Encourage/reinforce adherence. • Identify and troubleshoot the participant's problems in adhering to prescribed interventions (TIB, relaxation). • Address sleep effort and sleep-related anxiety. • Answer questions and address concerns. • Assign homework.
4 or 5	4	30-45 minutes	• Review sleep log and adjust TIB prescriptions. • Encourage/reinforce adherence. • Identify and troubleshoot the participant's problems in adhering to prescribed interventions (TIB, relaxation). • Address sleep effort and sleep-related anxiety. • Discuss relapse prevention. • Answer questions and address concerns. • Provide instruction on how to continue increasing TIB if desired total sleep time is not yet achieved.

Control Condition

We will compare the CBT-I group to a 28-week waitlist control group. Participants assigned to the waitlist control group do not engage in any study interventions while on the waitlist (during the primary assessment period) but complete study assessments at weeks 1-6, 12, and 28. After completion of all primary study time points (28-week follow-up), those assigned to the waitlist control group receive the same 4 telehealth CBT-I sessions delivered over 5 weeks as did the CBT-I group (secondary assessment period). The waitlist control group also completes additional questionnaires during their 4 treatment sessions, but the collected data will not be included in primary analyses.

Data Collection

This study is conducted through Stanford Zoom and Stanford RedCap. The Stanford RedCap platform is developed and operated by Stanford Medicine Research Information Technology team. The RedCap platform services at Stanford are subsidized by (1) Stanford School of Medicine Research Office and (2) the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through grant UL1 TR001085. All self-reported data are collected through questionnaires built on Stanford RedCap. Data from clinical interviews are entered into RedCap during the interview and then reviewed by the interviewer after the session. Once a participant is deemed eligible by the study team, they are randomized and scheduled for a second Zoom meeting with study staff. At this second Zoom meeting, they are notified of their study arm assignment by a research coordinator and complete week 1 questionnaires. If they are assigned to the treatment group, they meet with the therapist after completing week 1 questionnaires. For each following treatment session (weeks 2-5), participants meet with a research coordinator before meeting with the therapist. At the posttreatment time point (week 6), participants are sent a survey link to complete online questionnaires. If the participant is assigned to the waitlist control group, the week 1 session ends after completing the questionnaires. Participants in the waitlist control group are sent an email with a link to weekly surveys for each subsequent weekly session (weeks 2-6). Once the participant completes the week 1 questionnaires, they are considered enrolled in the study.

Recruitment, Enrollment Criteria, Screening, and Retention

Participants are adults in the United States aged 18 years or older who experienced new sleep disturbances after the start of the COVID-19 pandemic. Participants were recruited nationally through online postings, newsletters, and social media. A subset of participants was recruited from an ongoing survey-based observational study about sleep and well-being during the COVID-19 pandemic. All participants met the inclusion and exclusion criteria outlined in Textbox 1.

Interested participants completed online prescreening questions (Table 2) assessing insomnia symptoms before and after the pandemic began (March 1, 2020), insomnia symptom duration, any unstable medications, and seizure history. Participants who met the prescreening criteria (Table 2) were invited to a video call with study staff to further assess eligibility. During this call, a research coordinator explained details of the study procedures and obtained informed consent. Consenting participants and research staff obtaining consent signed the informed consent form via a RedCap survey with e-signature capabilities. Participants were emailed a copy of the signed consent form. After informed consent was obtained, a trained member of the study staff then administered clinical interviews to assess eligibility. Before ending the call, a research coordinator sent the participant a link to complete remaining online screening questionnaires at their own pace. After the participants completed all online screening questionnaires and screening interviews (described in the Measures section), the study team collectively determined if the inclusion and exclusion criteria were met.

During the screening session, sleep disturbance was assessed using the Duke Structured Interview for Sleep Disorders (hereafter referred to as Duke) [41] and ISI. Participants were excluded from the study if their reported sleep disturbance duration on the Duke began prior to the start of the pandemic or if another sleep disorder was primarily responsible for their symptoms. Two versions of the ISI were administered at screening: one assessing symptoms during the week before the pandemic began (past) and one assessing symptoms in the past 2 weeks (current). Participants were deemed eligible for the study if past ISI score was <10 and current ISI score was ≥10, as a score of 10 was found to be optimal in detecting insomnia in a community sample [36].

Participants with current or past psychosis, bipolar disorder, or epilepsy were excluded from the study due to safety concerns. Manipulating sleep increases the risk of seizures [42], mania [43], and psychosis [44] in individuals with a history of these conditions. Further, the brief version of CBT-I utilized in this study has not been validated as a reliable therapy in these populations. We administered the Mini-International Neuropsychiatric Interview (MINI) [45] to screen for psychosis and bipolar disorder and took a basic medical history to screen for history of seizures.

Participants currently abusing substances or taking over-the-counter or prescribed medications for sleep were not permitted in the study. To participate, all other medications must have been stable for at least 3 weeks, and medical conditions must have been deemed stable for at least 3 months by the study clinicians. However, hypnotics and other medications or supplements used to treat sleep disturbance were not permitted to be used during participation in the study. We collected information about current medication use as part of a basic medical history and assessed substance abuse or dependence using the MINI.

For eligible participants, data collected at the screening time point (including surveys after the session) are used as baseline measures.

Textbox 1. Criteria for study participation.

Inclusion Criteria

1. Age 18 years or older
2. Having access to the internet and an email address
3. Acute subjective complaint of sleep disturbance (Insomnia Severity Index [ISI] before the pandemic <10 and current ISI ≥10) that began after March 1, 2020 or the COVID-19 pandemic (as reported during interview)
4. Living in the United States
5. Literate and fluent in English
6. Willingness to participate in the study, sign the consent, and complete majority of questionnaires

Exclusion Criteria

1. Presence of suicidal ideation representing high risk as measured by Sheehan-Suicide Tracking Scale (S-STS)
2. Use of medication specifically prescribed for sleep disturbance and unwilling or unable to discontinue more than 1 week prior to baseline data collection
3. Current or lifetime history of bipolar disorder or psychosis
4. Current substance abuse or dependence
5. Not able to verbalize understanding of involvement in research and provide written, informed consent
6. Unstable pharmacotherapy for other mental health disorders (<3 weeks since beginning new medication)
7. Severe impediment to vision, hearing, or hand movement likely to interfere with the ability to complete assessments or are unable or unlikely to follow study protocols
8. Working rotating shift that overlaps with midnight

Table 2. Prescreening questions and responses indicating eligibility for a screening session.

Prescreening questions	Response criteria for screening session
Are you currently taking any prescribed or over-the-counter sleep medication?	No or willing to discontinue medication prior to enrollment
Have you started a new medication within the last four weeks?	No
How many months have you had trouble sleeping?	Duration indicates symptoms started after the start of the COVID-19 pandemic (March 1, 2020)
Do you have a personal history of epilepsy, convulsions, or seizures?	No
Current Insomnia Severity Index	Total score ≥10
Past Insomnia Severity Index	Total score <10

We employ several precautionary measures to reduce attrition and retain participants. First, at the beginning of the study, participants met with a member of the study team to discuss study procedures and answer questions. Participants are encouraged to ask questions throughout their involvement in the study. Participants were assigned an assessor who administers their clinical interviews at every time point, to encourage familiarity and build rapport with study staff. Participants voluntarily provided multiple different types of contact information (eg, email, phone numbers) for appointment reminders, and sessions are scheduled based on participants' time preferences. Study staff is persistent in attempting to recontact and engage noncompliant participants. Lastly, study assessments and data collection were carefully designed to minimize barriers to participation. We carefully curated the RedCap database so that participants do not have to re-enter information previously provided, and we eliminated long, superfluous questionnaires. We also added automated features to reduce the need for technical knowledge to navigate the questionnaires.

Protection Against Risks

All enrolled participants' depressive and anxiety symptoms are monitored by a trained clinical psychologist who reviews all adverse events and any significant changes in well-being. Unlike many research studies that recruit locally and conduct study visits in person, the present study enrolls participants across the nation who may suffer from severe depressive symptoms such as suicidal ideation. To address this unique challenge, we developed a robust distressed patient protocol adapted from the 2009 model developed by Draucker et al [46]. Our protocol includes several precautionary measures and multiple levels of assessment and risk classification. Each participant was asked to provide the phone number and email address of an emergency contact during the consenting process. At the start of the first video treatment session, the visit provider confirms the address of the participant's physical location and has the participant's

local emergency contact personnel readily available at each session. Study staff are trained to recognize signs of distress and indications that the participant may be having thoughts of harm to self or others. If such signs present over the course of the study visit, study staff conduct a formal risk assessment (using the Sheehan-Suicidality Tracking Scale [S-STS] if suicidal ideation is present or a homicide risk assessment if homicidal ideation is present) and determine if the participant is at imminent risk to self or others.

Study staff follow detailed instructions on how to proceed after a determination is made. If imminent threat is not determined, study participants are encouraged to follow up with their own mental health providers and are given contact information for their local emergency room, the National Suicide Prevention Lifeline, and the study psychologist. If imminent risk is determined, a warm transfer is provided to the National Suicide Prevention Lifeline. The visit provider then contacts the study psychologist, and together, they determine if contact of additional parties, including the participant's emergency contact or local sheriff's department, is warranted. At each stage, the study psychologist and principal investigator are apprised of steps taken, appropriate documentation is completed, and the institutional review board (IRB) is notified of any adverse events.

In addition to the aforementioned risk mitigation protocol, additional risk management strategies are utilized that leverage built-in, automated systems in RedCap if risk is detected during online survey completion. In RedCap, we established branching logic that automatically classifies subjects as low, moderate, or high suicide risk and sends precomposed emails based on risk level. Participants flagged as low, moderate, or high risk are sent direct emails with study staff contact information, as well as the National Suicide Prevention Lifeline. Those flagged as high risk are also contacted directly by the study psychologist who conducts a full risk assessment via telephone to ensure participant safety to self and others. Emergency contacts or local sheriff's departments are contacted if a participant is deemed to be high-risk and cannot be directly reached.

Measures

Screening

During the screening session, sleep disturbance was assessed using the Duke [41] and ISI. The Duke is a structured clinical interview that screens for sleep disorders in accordance with criteria of both the DSM-IV and the International Classification of Sleep Disorders, 2nd Edition, (ICSD-2). It is composed of 4 modules that assess sleep disorder symptoms associated with insomnia, hypersomnia, circadian rhythm sleep disorders, and sleep disorders associated with parasomnias.

The ISI is a 7-item, self-report measure of insomnia severity. The items consist of severity of early, middle, or late insomnia; sleep dissatisfaction; interference with daytime functioning; perception of sleep problems by others; and distress caused by sleep difficulties. Items are scored from 0 to 4, with 0 indicating no problem and 4 indicating a very severe problem. Score ranges for insomnia are as follows: 0-7, absent; 8-14, subthreshold; 15-21, moderate; and 22-28, severe.

We administered the MINI [45] to screen for psychosis, bipolar disorder, and substance abuse or dependence. The MINI is a structured, diagnostic interview for DSM-5 psychiatric disorders that was administered by clinically trained study staff.

Suicidal ideation and behaviors are assessed using the self-report version of the S-STS [47]. The S-STS is a 15-item questionnaire assessing the risk of suicidality using a 5-point Likert scale ranging from 0 to 4, with 0 indicating no problem and 4 indicating a very severe problem. For this study, to minimize participant burden, we omitted item 15, which asks about suicide attempts, because high-risk participants would already be identified through the distressed patient protocol (described in previous sections).

Primary Outcomes

Primary outcomes of insomnia, well-being, and predictors of treatment response are collected at baseline (week 0), weeks 6, 12, 28, and 56. Primary outcomes of insomnia (ISI) and well-being were also collected at weeks 1-5. The primary time points are weeks 6, 12, and 28.

Insomnia

The change in clinically significant insomnia symptoms (meeting criteria for insomnia disorder diagnosis) and subjective ratings of current insomnia symptoms are primary measures of insomnia. Insomnia disorder diagnosis is assessed using the Duke insomnia disorder module. Subjective sleep complaints are assessed using ISI report of current symptoms over the past 2 weeks. The Duke insomnia disorder module is only collected at weeks 0, 12, 28, and 56.

Well-being

We measure change in depressive symptoms as primary outcomes of well-being. Depressive symptoms are evaluated using the Patient Health Questionnaire-9 (PHQ-9) [48]. The PHQ-9 is a self-administered, 9-item questionnaire that assesses each of the 9 DSM-IV depression criteria. The total score ranges from 1 to 27. Answers are given on a 4-point Likert scale ranging from 0 (Not at all) to 3 (Nearly every day), with higher scores reflecting increased severity of depressive symptoms.

Predictors of Treatment Response

Baseline levels of sleep reactivity and pandemic-related risk factors, including loneliness, perceived stress, screen time, social connection, and physical activity, are measured as predictors of long-term treatment outcomes.

Loneliness is measured using the University of California, Los Angeles (UCLA) Loneliness Scale [49]. This self-report measure assesses participant's subjective feelings of loneliness and social isolation. The UCLA Loneliness Scale asks participants to rate how often each of the 20 items is descriptive of them, rated from 1 (never) to 4 (often). The responses are for an overall score range of 20 to 80, with higher scores indicating greater degrees of reported loneliness.

Sleep reactivity is measured by the Ford Insomnia Response to Stress Test (FIRST) [50]. This 9-item self-report tool measures risk of participants experiencing situational insomnia due to common stressful conditions. A higher score on the FIRST

indicates higher sleep reactivity. The FIRST has good reliability and validity and has demonstrated high internal consistency across multiple demographic groups in clinical and population-based samples [51].

Perceived stress is measured using the Perceived Stress Scale (PSS) [52]. The PSS is a widely used 14-item self-report questionnaire that assesses how stressful participants believe their lives to be. Items are generalized and measure the degree to which participants judge their lives to have been uncontrollable and unpredictable over the course of the previous month. Items are scored on a 5-point Likert scale, with total scores ranging from 0 to 56 and higher scores indicating higher levels of perceived stress.

Screen time is measured through modified self-report questions in the Coronavirus Health Impact Survey (CRISIS) [53]. Participants are asked to estimate the number of hours spent per day, over the course of the 2 most recent weeks, watching television or digital media (eg. Netflix, YouTube, web surfing) or using social media (eg. FaceTime, Facebook, Instagram, Snapchat, Twitter, TikTok).

Social connection is measured by the Social Network Index (SNI) [54]. The SNI is a 12-item questionnaire that assesses participation in different types of social relationships. The 12 types of relationships (eg, friend, children, spouse, religious group member) are scored by the number of network members with which they communicate at least every 2 weeks. A higher score on the SNI indicates a larger social network.

Physical activity is evaluated using the International Physical Activity Questionnaire (IPAQ) [55]. The IPAQ assesses the time spent on an individual's physical activity across 5 life domains over the previous 7 days. The activity domains consist of physical activity related to work, transportation, housework and caring for family, and recreation and sports, as well as the amount of time spent sitting each day. Minutes of sitting and walking, as well as moderate-intensity (walking not included) and vigorous-intensity activities, were calculated for each domain and for the entire past 7 days. The IPAQ has high reliability and validity and has been widely used to measure comparable estimates of physical activity in large populations. A higher score on the IPAQ indicates an increased physical activity.

Secondary Outcomes

All secondary outcomes of insomnia and well-being are collected at weeks 0, 12, 28, and 56. Secondary measures of insomnia and the Generalized Anxiety Disorder-7 (GAD-7) are also collected at weeks 1-6.

Insomnia

Changes in sleep onset latency (SOL), number of awakenings, wake after sleep onset (WASO), total sleep time (TST), and sleep efficiency (SE) over time are measured as secondary measures of insomnia symptoms using sleep diaries; 7 days of sleep diaries are collected at baseline and weeks 1-5, 12, 28, and 56. Sleep diaries collect information about sleep and rise times, time in and out of bed, number of middle-of-the-night awakenings, duration of these awakenings, sleep quality, nap frequency and duration, and caffeine and alcohol consumption. SOL is the time in minutes from “lights out” to sleep onset. WASO is the sum of the total number of minutes of wakefulness occurring after sleep onset and before final awakening (sleep offset). TST is the total time spent asleep, from the start of sleep onset to sleep offset, subtracting any periods of wakefulness. SE is calculated as TST divided by total time spent in bed, multiplied by 100.

Well-being

Secondary measures of well-being include measures of anxiety symptoms, suicidal ideation, and quality of life, as well as an additional measure of depressive symptoms.

Anxiety symptoms over time are assessed using the GAD-7 [56]. The GAD-7 is a widely used diagnostic self-report scale that assesses severity of anxiety symptomology. The GAD-7 is a 7-item, 4-point Likert scale ranging from 0 to 3 that measures severity of anxiety symptoms over the previous 2 weeks, with a total score ranging from 0 to 21. Higher scores indicate more severe anxiety symptomology.

Suicidal ideation and behaviors over time are measured by the S-STS [47]. For this outcome measure, 14 items (excluding item 15) are summed for an overall score ranging from 0 to 56. Again, higher scores indicate more severe difficulties.

Quality of life is assessed using the 36-Item Short Form Health Survey (SF-36) [57]. The SF-36 is a 36-item self-report survey to assess comprehensive quality-of-life measures. It consists of 8 subscales: vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, and mental health. This survey is widely used and has been proven to be a reliable indicator of quality-of-life measures. The global score range is 0 to 100, with higher scores indicating better health conditions.

The Beck Anxiety Inventory (BAI) [58] is used as a measure of anxiety symptom change over time. The BAI is a 21-item self-report scale that assesses the severity of anxiety symptoms. Items are scored from 0 to 3. Higher scores indicate greater levels of severity, and the ranges for anxiety levels are as follows: 0-9, normal to minimal; 10-18, mild to moderate; 19-29, moderate to severe; and 30-63, severe. The BAI consists of 2 factors: somatic and cognitive.

The Beck Depression Inventory-II (BDI-II) [59,60] is used as a secondary measure of depressive symptom change over time. We will sum all items except one sleep item, and the average item score for the remaining 20 items will be multiplied by 21 (the original number of items) to create a modified depression scale that maintains the original range (ranges: 0-13, minimal; 14-19, mild; 20-28, moderate; and 29-63, severe). The BDI-II is a 21-item self-report scale with high validity and reliability that assesses the severity of depressive symptoms. The depressive items consist of sadness, pessimism, past failure, loss of pleasure, guilty feelings, punishment feelings, self-dislike, self-criticalness, suicidal thoughts or wishes, crying, agitation, loss of interest, indecision, worthlessness, loss of energy, changes in sleep pattern, irritability, changes in appetite, concentration difficulty, tiredness or fatigue, and loss of interest

in sex. Items are scored from 0 to 3, and higher scores indicate greater level of severity.

Statistical Procedure and Data Analysis

Given the nature of this feasibility study, for all our analyses and interpretation, we will place a primary emphasis on estimation of effect sizes and confidence intervals rather than on testing statistical significance.

Primary Analyses

Statistical analyses aim 1 is to determine whether a brief, telehealth CBT-I reduces insomnia symptoms arising during the COVID-19 pandemic. All primary analyses will be performed using the intention-to-treat principle. We will test whether CBT-I is superior to a waitlist control in reducing insomnia symptoms by using a mixed effects linear model with autoregressive error structure using intention-to-treat analysis with outcomes at time points 6, 12, and 28. Insomnia severity as measured by the ISI will be entered as the dependent variable with randomization group (2 levels), time point (weeks 0, 6, 12, and 28), and group-by-time interaction included as fixed effects. The model will also include a random slope. Several hypotheses (1.1a-d) will be tested using the described mixed effects model.

Hypothesis 1.1a is that individuals assigned to the CBT-I group will experience an improved trajectory of insomnia symptoms during the waitlist-controlled 28-week period relative to those assigned to the waitlist control group. This hypothesis will be tested using a likelihood ratio test of the coefficients of the time and group-by-time interaction of the aforementioned mixed effects linear model.

Hypotheses 1.1b-1.1d are that individuals assigned to the CBT-I group during the waitlist-controlled period will have lower insomnia symptoms compared with those assigned to the waitlist-control group immediately posttreatment (hypothesis 1.1b; week 6) and at the short-term (hypothesis 1.1c; week 12) and long-term (hypothesis 1.1d; week 28) follow-ups. These hypotheses will be tested using the aforementioned mixed effects linear model with independent likelihood ratio tests using the treatment coefficient for the posttreatment (hypothesis 1.1b; Δ_1 in Figure 2), short-term (hypothesis 1.1c; Δ_2 in Figure 2), and long-term (hypothesis 1.1c; Δ_3 in Figure 2) follow-up time points.

We will also test whether CBT-I is superior to the waitlist control in preventing an insomnia diagnosis following treatment by using multiple logistic regression analysis.

Hypotheses 1.2a and 1.2b are that individuals who were assigned to the CBT-I group will be less likely to have an insomnia diagnosis than those who were assigned to the waitlist control group at weeks 12 and 28. These hypotheses will be tested using 2 separate logistic regression models, one using the short-term (hypothesis 1.2a; 12 weeks) and one using long-term (hypothesis 1.2b; 28 weeks) follow-up insomnia diagnosis, as defined by the Duke, as the dependent variable, with treatment (CBT-I, waitlist) as the independent variable.

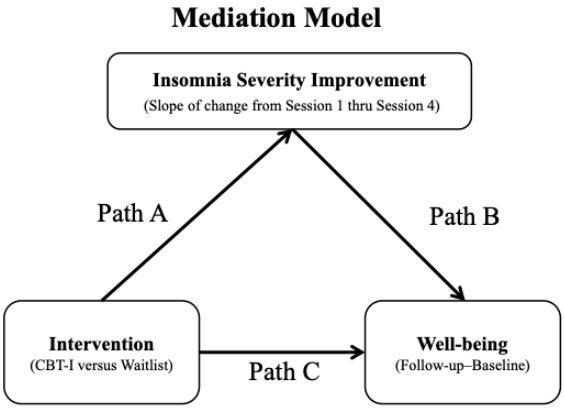
Statistical analyses aim 2 is to determine whether a brief, telehealth CBT-I improves well-being during the COVID-19 pandemic. We will test whether immediate CBT-I is superior to a waitlist control in improving depressive symptoms as the primary measure of well-being during the waitlist-controlled period (28-week duration) by using a mixed effects linear model with autoregressive error structure using intention-to-treat analysis. Depression severity, as assessed by the BDI-II, will be entered as the dependent variable with baseline covariates used for stratified randomization (ie, baseline insomnia severity and sex), group (immediate CBT-I or waitlist control), time point (weeks 0-6, 12, and 28), and group-by-time interaction included as fixed effects. The model will also include a random slope. Separate, secondary analyses will be conducted on the secondary measures relating to well-being described in the Secondary Outcomes section (eg, anxiety, suicidality, quality of life). Several hypotheses will be tested using this model.

Hypothesis 2.2a is that individuals assigned to the CBT-I group will have an improved trajectory of depressive symptoms across the 28 weeks relative to individuals assigned to the waitlist control group. This hypothesis will be tested using a likelihood ratio test of the coefficients of the time and group-by-time interaction terms of the aforementioned mixed effects linear model.

Hypotheses 2.2b-2.2d are that individuals assigned to the CBT-I group will have lower depressive symptoms compared who those who were assigned to the waitlist control group immediately posttreatment (hypothesis 2.2b; week 6) and at the short-term (hypothesis 2.2c; week 12) and long-term (hypothesis 2.2d; week 28) follow-ups. These hypotheses will be tested using the aforementioned mixed effects linear model with independent likelihood ratio tests using the treatment coefficient for the posttreatment (hypothesis 1.2b; week 6), short-term (hypothesis 1.2c; week 12), and long-term (hypothesis 1.2c; week 28) follow-up time points.

Hypothesis 2.3 is that improvement in insomnia symptom severity (measured by the ISI) will mediate subsequent improvement in well-being from baseline (week 0) to posttreatment (week 6), short-term follow-up (week 12), and long-term follow-up (week 28). Using the approach described by Kraemer at al [61] and Manber et al [34], linear regression analyses will be used to determine whether change in the ISI score from Session 1 (week 1) to Session 4 (week 5) of treatment mediates the subsequent depressive symptom improvement (primary outcome measure) after the completion of treatment (week 6). Change in ISI score at Session 4 will be estimated as the slope of the regression line of ISI scores from Session 1 (week 1) through the treatment phase to Session 4 (week 6) using all available data from each participant. The model will include group (CBT-I or waitlist control), change in ISI score from Session 1 to Session 4 (participant specific slopes), and their interaction with depression severity, as measured by the PHQ-9, as the dependent variables. Mediation will be tested using the MacArthur approach as described by Kraemer et al [62]. Specifically, a mediation effect requires a significant effect of the intervention on the mediator (Path A in Figure 3) and a significant association between the mediator and the outcome (Path B in Figure 3).

Figure 3. Mediation Model of Insomnia Severity Improvements mediating the change in well-being associated with the intervention.



Statistical analyses aim 3 is to determine whether risk factors for insomnia that are aggravated during the COVID-19 pandemic predict worse insomnia and well-being outcomes at follow-up.

Hypothesis 3 is that high levels of social isolation, perceived stress, sleep reactivity, and screen time and low levels of physical activity caused by the COVID-19 pandemic will collectively predict a worse long-term outcome across both intervention arms. Two separate linear mixed models will be conducted for each outcome variable (insomnia severity and depressive symptoms). In each model, the outcome variable at 12 weeks and 28 weeks will be entered as the dependent variable with social isolation, perceived stress, sleep reactivity, screen time, and physical activity measures at baseline as well as experimental arm entered as predictors.

Secondary Analyses

Secondary analyses will be conducted. (1) Analytic methods described in the primary aims will be repeated but applied to secondary measures of sleep disturbance and well-being outcomes as described in the Secondary Outcomes section. (2) Analytic methods described in the primary aims will be repeated but applied to outcome measures collected at week 56. (3) Sparse unsupervised clustering and principal component analysis analyses will be used to identify cohesive factors of dysfunction in sleep complaints and patterns of mental health outcomes at baseline. Regression models will be used to quantify the relationships within and between sleep complaints and patterns of mental health outcomes. (4) Age and sex differences in treatment response and as moderators of relationships between sleep complaints, insomnia risk factors, and mental health outcomes will be explored.

Ethics Approval and Trial Registration

The IRB of Stanford University (Stanford, CA) approved the study, which is performed following the rules of the seventh edition (2013) of the Declaration of Helsinki (IRB-55940). It received initial approval from the IRB on April 30, 2020.

Results

Trial Status and Timeline

The trial began recruitment on June 4, 2020, and the first participant was consented on June 10, 2020. As of October 2021, 794 subjects had completed prescreening, and 96 subjects had completed a Zoom screening session. A total of 49 participants were randomized to a study group (26 to the CBT-I group and 23 to the waitlist control group). Overall, 38 participants had completed the week 6 follow-up, 37 had completed the week 12 follow-up, 34 had completed the week 28 follow-up, and 15 had completed the week 56 follow-up. At the time of this writing, there were an additional 3 participants awaiting the 28-week follow-up and a total of 23 participants awaiting the week 56 follow-up (13 in CBT-I and 10 in waitlist control group). The study closed to enrollment of new subjects on June 17, 2021. As of the writing of this paper, due to the outstanding data collection of the remaining 28- and 56-week data, we had conducted interim analyses for the week 6 follow-up but had not tested the results of our primary hypotheses in full. We expect primary results of the study to be published in 2022.

Discussion

Here, we outline the protocol of an innovative research project responding to the mental health crisis related to the COVID-19 pandemic. We aim to address several gaps in the literature by investigating the use of an early, brief, nonpharmacological insomnia intervention delivered via telehealth to treat insomnia symptoms arising during a stressful life event and prevent worsening insomnia or mental health outcomes. This study represents multiple levels of innovation. First, to the best of our knowledge, this is the only project clinically responding to the large-scale increase in sleep disturbance during the COVID-19 pandemic. Although many studies are documenting the robust changes in sleep during the pandemic, this project is the first to test an intervention and examine pandemic-related predictors of intervention response. Second, on a broader scale, this is one of only a few studies testing the prospective temporal relationship between sleep disturbance and well-being. Lastly, to our knowledge, this is the only study of early deployment of CBT-I to treat insomnia symptoms arising from a stressful global event.

This study was conceived and launched in response to the global pandemic of COVID-19, which, while rendering it novel, presents unique challenges. First, to address our scientific questions, we sought to recruit participants early in the course of their sleep disturbance, causing a significant urgency to

launch the study. This time pressure led us to conduct this study with limited resources, which impacted our recruitment efforts. We were unable to offer compensation for participation. We relied heavily on recruiting participants through free online platforms, such as social media posts, online message boards, and electronic newsletters. Due to these recruitment concerns and the resulting decreased statistical power for the planned analysis, we view this as a feasibility study. Therefore, the planned analyses and publications resulting from this protocol will utilize effect size estimates, rather than statistical significance, to provide pilot and feasibility data to inform future hypotheses. A second potential limitation is the unequal attrition rate in the treatment versus the waitlist control group. Since participants are not financially compensated, the primary motivation of participation is meeting with a therapist on a weekly basis to address their sleep concerns. Thus, it is possible that individuals in the waitlist control group were more likely to decline to participate in follow-up time points due to loss of interest or spontaneous improvement in symptoms over time. Although this could potentially affect the ability to detect significant differences in between-group comparisons at the long-term follow-up time points due to inadequate power, our analytic approach using regression models has the advantage of increased efficiency and power over unadjusted analyses [63,64]. We will address this issue by assessing whether missingness is related to any observed covariates. If this is the case, an issue of greater concern is the potential bias in the estimate of the treatment effect if neither the missing completely at random (MCAR) nor the missing at random (MAR) assumption holds. Although mixed effects linear models lead to unbiased estimates of the treatment effect under the MCAR and MAR assumptions, estimates may be biased under nonignorable missingness. Further sensitivity analyses will be performed if this is the case.

Although the unique conditions created by the COVID-19 pandemic posed many novel challenges, it also raises the potential significance of this protocol and feasibility study: It is unlikely that we or another group will be able to replicate this study in the future. Therefore, although the sample will be smaller than nonpandemic trials, it will likely be the largest sample to report on a sleep intervention deployed during the acute stage of a global pandemic. Supporting our hypothesis, the findings of a recent, large-scale clinical trial indicate that treatment of insomnia with CBT-I has an overall benefit in the prevention of incidence and recurrence of major depression in older adults with insomnia disorder. However, this trial utilized an in-person intervention in subjects with chronic insomnia [65]. Our study remains important by extending these findings to address the public health need for an effective, remote intervention to treat insomnia early and potentially prevent negative well-being outcomes. This project will lay the groundwork for future studies investigating early deployment of CBT-I to treat insomnia precipitated by a stressful event and whether this prevents adverse mental health outcomes for which poor sleep is a risk factor. Further, our investigation of pandemic-related risk factors for insomnia could guide future hypotheses for studies conducted in nonpandemic conditions.

Acknowledgments

We thank Theresa Brown, Andrea Cordero, Maria Ocampo, Margot Paul, Kelly Showen, Alix Simonson, and Jennifer Stephens for their help with data collection protocols.

Conflicts of Interest

None declared.

References

1. Marelli S, Castelnuovo A, Somma A, Castronovo V, Mombelli S, Bottoni D, et al. Impact of COVID-19 lockdown on sleep quality in university students and administration staff. J Neurol 2021 Jan;268(1):8-15 [FREE Full text] [doi: 10.1007/s00415-020-10056-6] [Medline: 32654065]
2. Li Y, Qin Q, Sun Q, Sanford LD, Vgontzas AN, Tang X. Insomnia and psychological reactions during the COVID-19 outbreak in China. J Clin Sleep Med 2020 Aug 15;16(8):1417-1418 [FREE Full text] [doi: 10.5664/jcsm.8524] [Medline: 32351206]
3. Hyun S, Hahm HC, Wong GTF, Zhang E, Liu CH. Psychological correlates of poor sleep quality among U.S. young adults during the COVID-19 pandemic. Sleep Med 2021 Feb;78:51-56 [FREE Full text] [doi: 10.1016/j.sleep.2020.12.009] [Medline: 33385779]
4. Gualano MR, Lo Moro G, Voglino G, Bert F, Siliquini R. Effects of Covid-19 lockdown on mental health and sleep disturbances in Italy. Int J Environ Res Public Health 2020 Jul 02;17(13):1 [FREE Full text] [doi: 10.3390/ijerph17134779] [Medline: 32630821]
5. Yu BY, Yeung W, Lam JC, Yuen SC, Lam SC, Chung VC, et al. Prevalence of sleep disturbances during COVID-19 outbreak in an urban Chinese population: a cross-sectional study. Sleep Med 2020 Oct;74:18-24 [FREE Full text] [doi: 10.1016/j.sleep.2020.07.009] [Medline: 32836181]
6. Kokou-Kpolou CK, Megalakaki O, Laimou D, Kousouri M. Insomnia during COVID-19 pandemic and lockdown: Prevalence, severity, and associated risk factors in French population. Psychiatry Res 2020 Aug;290:113128 [FREE Full text] [doi: 10.1016/j.psychres.2020.113128] [Medline: 32563951]

7. LeBlanc M, Mérette C, Savard J, Ivers H, Baillargeon L, Morin CM. Incidence and risk factors of insomnia in a population-based sample. Sleep 2009 Aug;32(8):1027-1037 [FREE Full text] [doi: 10.1093/sleep/32.8.1027] [Medline: 19725254]
8. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. Sleep Med Rev 2002 Apr;6(2):97-111. [doi: 10.1053/smrv.2002.0186] [Medline: 12531146]
9. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. JAMA Netw Open 2020 Mar 02;3(3):e203976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]
10. Pinto J, van Zeller M, Amorim P, Pimentel A, Dantas P, Eusébio E, et al. Sleep quality in times of Covid-19 pandemic. Sleep Med 2020 Oct;74:81-85 [FREE Full text] [doi: 10.1016/j.sleep.2020.07.012] [Medline: 32841849]
11. Stanton R, To QG, Khalesi S, Williams SL, Alley SJ, Thwaite TL, et al. Depression, anxiety and stress during COVID-19: associations with changes in physical activity, sleep, tobacco and alcohol use in Australian adults. Int J Environ Res Public Health 2020 Jun 07;17(11):1 [FREE Full text] [doi: 10.3390/ijerph17114065] [Medline: 32517294]
12. Zhang Y, Zhang H, Ma X, Di Q. Mental health problems during the COVID-19 pandemics and the mitigation effects of exercise: a longitudinal study of college students in China. Int J Environ Res Public Health 2020 May 25;17(10):1 [FREE Full text] [doi: 10.3390/ijerph17103722] [Medline: 32466163]
13. Grossman ES, Hoffman YSG, Palgi Y, Shrira A. COVID-19 related loneliness and sleep problems in older adults: Worries and resilience as potential moderators. Pers Individ Dif 2021 Jan 01;168:110371 [FREE Full text] [doi: 10.1016/j.paid.2020.110371] [Medline: 32904342]
14. Cellini N, Canale N, Mioni G, Costa S. Changes in sleep pattern, sense of time and digital media use during COVID-19 lockdown in Italy. J Sleep Res 2020 Aug;29(4):e13074 [FREE Full text] [doi: 10.1111/jsr.13074] [Medline: 32410272]
15. Hale L, Guan S. Screen time and sleep among school-aged children and adolescents: a systematic literature review. Sleep Med Rev 2015 Jun;21:50-58 [FREE Full text] [doi: 10.1016/j.smrv.2014.07.007] [Medline: 25193149]
16. Gubelmann C, Heinzer R, Haba-Rubio J, Vollenweider P, Marques-Vidal P. Physical activity is associated with higher sleep efficiency in the general population: the CoLaus study. Sleep 2018 Jul 01;41(7):1. [doi: 10.1093/sleep/zsy070] [Medline: 29617980]
17. Deng J, Zhou F, Hou W, Silver Z, Wong CY, Chang O, et al. The prevalence of depression, anxiety, and sleep disturbances in COVID-19 patients: a meta-analysis. Ann N Y Acad Sci 2021 Feb;1486(1):90-111 [FREE Full text] [doi: 10.1111/nyas.14506] [Medline: 33009668]
18. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the general population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
19. Qiu J, Shen B, Zhao M, Wang Z, Xie B, Xu Y. A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations. Gen Psychiatr 2020;33(2):e100213 [FREE Full text] [doi: 10.1136/gpsych-2020-100213] [Medline: 32215365]
20. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and ecological momentary assessment study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
21. Fernández-Abascal EG, Martín-Díaz MD. Longitudinal study on affect, psychological well-being, depression, mental and physical health, prior to and during the COVID-19 pandemic in Spain. Pers Individ Dif 2021 Apr;172:110591 [FREE Full text] [doi: 10.1016/j.paid.2020.110591] [Medline: 33518870]
22. Romero-Gonzalez B, Puertas-Gonzalez JA, Mariño-Narvaez C, Peralta-Ramirez MI. [Confinement variables by COVID-19 predictors of anxious and depressive symptoms in pregnant women]. Med Clin (Barc) 2021 Feb 26;156(4):172-176 [FREE Full text] [doi: 10.1016/j.medcli.2020.10.002] [Medline: 33243419]
23. Hertenstein E, Feige B, Gmeiner T, Kienzler C, Spiegelhalder K, Johann A, et al. Insomnia as a predictor of mental disorders: A systematic review and meta-analysis. Sleep Med Rev 2019 Feb;43:96-105. [doi: 10.1016/j.smrv.2018.10.006] [Medline: 30537570]
24. Qaseem A, Kansagara D, Forciea MA, Cooke M, Denberg TD, Clinical Guidelines Committee of the American College of Physicians. Management of Chronic Insomnia Disorder in Adults: A Clinical Practice Guideline From the American College of Physicians. Ann Intern Med 2016 Jul 19;165(2):125-133 [FREE Full text] [doi: 10.7326/M15-2175] [Medline: 27136449]
25. Trauer JM, Qian MY, Doyle JS, Rajaratnam SMW, Cunnington D. Cognitive behavioral therapy for chronic insomnia: a systematic review and meta-analysis. Ann Intern Med 2015 Aug 04;163(3):191-204. [doi: 10.7326/M14-2841] [Medline: 26054060]
26. Koffel EA, Koffel JB, Gehrman PR. A meta-analysis of group cognitive behavioral therapy for insomnia. Sleep Med Rev 2015 Feb;19:6-16 [FREE Full text] [doi: 10.1016/j.smrv.2014.05.001] [Medline: 24931811]
27. Arnedt JT, Conroy DA, Mooney A, Furgal A, Sen A, Eisenberg D. Telemedicine versus face-to-face delivery of cognitive behavioral therapy for insomnia: a randomized controlled noninferiority trial. Sleep 2021 Jan 21;44(1):1. [doi: 10.1093/sleep/zsaa136] [Medline: 32658298]

28. Altena E, Baglioni C, Espie CA, Ellis J, Gavriloff D, Holzinger B, et al. Dealing with sleep problems during home confinement due to the COVID-19 outbreak: Practical recommendations from a task force of the European CBT-I Academy. J Sleep Res 2020 Aug;29(4):e13052 [FREE Full text] [doi: 10.1111/jsr.13052] [Medline: 32246787]
29. Gebara MA, Siripong N, DiNapoli EA, Maree RD, Germain A, Reynolds CF, et al. Effect of insomnia treatments on depression: A systematic review and meta-analysis. Depress Anxiety 2018 Aug;35(8):717-731. [doi: 10.1002/da.22776] [Medline: 29782076]
30. Bélanger L, Harvey AG, Fortier-Brochu E, Beaulieu-Bonneau S, Eidelman P, Talbot L, et al. Impact of comorbid anxiety and depressive disorders on treatment response to cognitive behavior therapy for insomnia. J Consult Clin Psychol 2016 Aug;84(8):659-667 [FREE Full text] [doi: 10.1037/ccp0000084] [Medline: 26963600]
31. Luik AI, Kyle SD, Espie CA. Digital cognitive behavioral therapy (dCBT) for insomnia: a state-of-the-science review. Curr Sleep Med Rep 2017;3(2):48-56 [FREE Full text] [doi: 10.1007/s40675-017-0065-4] [Medline: 28553574]
32. Trockel M, Karlin BE, Taylor CB, Brown GK, Manber R. Effects of cognitive behavioral therapy for insomnia on suicidal ideation in veterans. Sleep 2015 Feb 01;38(2):259-265 [FREE Full text] [doi: 10.5665/sleep.4410] [Medline: 25515115]
33. Scott A, Webb T, Martyn-St James M, Rowse G, Weich S. Improving sleep quality leads to better mental health: A meta-analysis of randomised controlled trials. Sleep Med Rev 2021 Dec;60:101556 [FREE Full text] [doi: 10.1016/j.smrv.2021.101556] [Medline: 34607184]
34. Manber R, Buysse DJ, Edinger J, Krystal A, Luther JF, Wisniewski SR, et al. Efficacy of cognitive-behavioral therapy for insomnia combined with antidepressant pharmacotherapy in patients with comorbid depression and insomnia: a randomized controlled trial. J Clin Psychiatry 2016 Oct;77(10):e1316-e1323. [doi: 10.4088/JCP.15m10244] [Medline: 27788313]
35. Morin CM. Insomnia: Psychological assessment and management. New York, NY: Guilford Press; 1993.
36. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 2011 May 01;34(5):601-608 [FREE Full text] [doi: 10.1093/sleep/34.5.601] [Medline: 21532953]
37. Borbely AA. Concepts and models of sleep regulation: a preface. Journal of Sleep Research 1992;1(2):61. [doi: 10.1111/j.1365-2869.1992.tb00012.x]
38. Spielman AJ, Saskin P, Thorpy MJ. Treatment of chronic insomnia by restriction of time in bed. Sleep 1987 Feb;10(1):45-56. [Medline: 3563247]
39. Bootzin RR, Epstein D, Wood JM. Stimulus Control Instructions. In: Hauri PJ, editor. Case Studies in Insomnia. Critical Issues in Psychiatry (An Educational Series for Residents and Clinicians). Boston, MA: Springer; 1991:19-28.
40. Edinger J. Treatment Manual Cognitive-Behavioral Insomnia Therapy. URL: https://www.med.unc.edu/neurology/wp-content/uploads/sites/716/2018/05/jdedingrCBTManual.pdf [accessed 2022-02-13]
41. Edinger J, Kirby A, Lineberger M, Loiselle M, Wohlgemuth W, Means M. The Duke Structured Interview Schedule for DSM-IV-TR and International Classification of Sleep Disorders. ICSD-2 Sleep Disord Diagnoses 2009:1-78.
42. Bazil CW. Seizure modulation by sleep and sleep state. Brain Res 2019 Jan 15;1703:13-17. [doi: 10.1016/j.brainres.2018.05.003] [Medline: 29782849]
43. Kaplan KA, Harvey AG. Behavioral treatment of insomnia in bipolar disorder. Am J Psychiatry 2013 Jul;170(7):716-720 [FREE Full text] [doi: 10.1176/appi.ajp.2013.12050708] [Medline: 23820830]
44. Reeve S, Sheaves B, Freeman D. The role of sleep dysfunction in the occurrence of delusions and hallucinations: A systematic review. Clin Psychol Rev 2015 Dec;42:96-115 [FREE Full text] [doi: 10.1016/j.cpr.2015.09.001] [Medline: 26407540]
45. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34. [Medline: 9881538]
46. Draucker CB, Martsolf DS, Poole C. Developing distress protocols for research on sensitive topics. Arch Psychiatr Nurs 2009 Oct;23(5):343-350. [doi: 10.1016/j.apnu.2008.10.008] [Medline: 19766925]
47. Sheehan DV, Giddens JM, Sheehan IS. Status update on the Sheehan-Suicidality Tracking Scale (S-STS) 2014. Innov Clin Neurosci 2014 Sep;11(9-10):93-140 [FREE Full text] [Medline: 25520892]
48. 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]
49. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
50. Kalmbach DA, Pillai V, Arnedt JT, Drake CL. Identifying at-risk individuals for insomnia using the Ford Insomnia Response to Stress Test. Sleep 2016 Feb 01;39(2):449-456 [FREE Full text] [doi: 10.5665/sleep.5462] [Medline: 26446111]
51. Jarrin DC, Chen IY, Ivers H, Drake CL, Morin CM. Temporal stability of the Ford Insomnia Response to Stress Test (FIRST). J Clin Sleep Med 2016 Oct 15;12(10):1373-1378 [FREE Full text] [doi: 10.5664/jcsm.6192] [Medline: 27568895]
52. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. Journal of Health and Social Behavior 1983 Dec;24(4):385. [doi: 10.2307/2136404]
53. Nikolaidis A, Paksarian D, Alexander L, Derosa J, Dunn J, Nielson DM, et al. The Coronavirus Health and Impact Survey (CRISIS) reveals reproducible correlates of pandemic-related mood states across the Atlantic. Sci Rep 2021 Apr 14;11(1):8139 [FREE Full text] [doi: 10.1038/s41598-021-87270-3] [Medline: 33854103]

54. Cohen S, Doyle WJ, Skoner DP, Rabin BS, Gwaltney JM. Social ties and susceptibility to the common cold. JAMA 1997 Jun 25;277(24):1940-1944. [Medline: 9200634]
55. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
56. Spitzer RL, Kroenke K, Williams JBW, 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]
57. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. Health Econ 1993 Oct;2(3):217-227. [doi: 10.1002/hec.4730020305] [Medline: 8275167]
58. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: Psychometric properties. Journal of Consulting and Clinical Psychology 1988;56(6):893-897. [doi: 10.1037/0022-006X.56.6.893]
59. Beck AT, Steer RA, Ball R, Ranieri W. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. J Pers Assess 1996 Dec;67(3):588-597. [doi: 10.1207/s15327752jpa6703_13] [Medline: 8991972]
60. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation; 1996.
61. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. Arch Gen Psychiatry 2002 Oct;59(10):877-883. [doi: 10.1001/archpsyc.59.10.877] [Medline: 12365874]
62. Kraemer HC, Kiernan M, Essex M, Kupfer DJ. Health Psychol 2008 Mar;27(2S):S101-S108 [FREE Full text] [doi: 10.1037/0278-6133.27.2(Suppl.).S101] [Medline: 18377151]
63. Moore KL, Neugebauer R, Valappil T, Laan MJ. Robust extraction of covariate information to improve estimation efficiency in randomized trials. Stat Med 2011 Aug 30;30(19):2389-2408 [FREE Full text] [doi: 10.1002/sim.4301] [Medline: 21751231]
64. Benkeser D, Díaz I, Luedtke A, Segal J, Scharfstein D, Rosenblum M. Improving precision and power in randomized trials for COVID-19 treatments using covariate adjustment, for binary, ordinal, and time-to-event outcomes. Biometrics 2021 Dec;77(4):1467-1481 [FREE Full text] [doi: 10.1111/biom.13377] [Medline: 32978962]
65. Irwin MR, Carrillo C, Sadeghi N, Bjurstrom MF, Breen EC, Olmstead R. Prevention of incident and recurrent major depression in older adults with insomnia: a randomized clinical trial. JAMA Psychiatry 2022 Jan 01;79(1):33-41 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.3422] [Medline: 34817561]

Abbreviations

BAI: Beck Anxiety Inventory
BDI-II: Beck Depression Inventory-II
CBT-I: cognitive behavioral therapy for insomnia
CRISIS: Coronavirus Health Impact Survey
DSM-5: Diagnostic and Statistical Manual of Mental Disorders 5th Edition
Duke: Duke Structured Interview for Sleep Disorders
FIRST: Ford Insomnia Response to Stress Test
GAD-7: Generalized Anxiety Disorder-7
ICSD-2: International Classification of Sleep Disorders, 2nd Edition
IPAQ: International Physical Activity Questionnaire
IRB: institutional review board
ISI: Insomnia Severity Index
MAR: missing at random
MCAR: missing completely at random
MINI: Mini-International Neuropsychiatric Interview
PHQ-9: Patient Health Questionnaire-9
PSS: Perceived Stress Scale
SE: sleep efficiency
SF-36: 36-Item Short Form Health Survey
SNI: Social Network Index
SOL: sleep onset latency
S-STS: Sheehan Suicidality Tracking Scale
TST: total sleep time
UCLA: University of California, Los Angeles
WASO: wake after sleep onset

Edited by G Eysenbach; submitted 22.10.21; peer-reviewed by Z Butt, H Ding, R Rastmanesh; comments to author 16.11.21; revised version received 07.12.21; accepted 18.12.21; published 14.03.22.

Please cite as:
O'Hora KP, Osorno RA, Sadeghi-Bahmani D, Lopez M, Morehouse A, Kim JP, Manber R, Goldstein-Piekarski AN
Viability of an Early Sleep Intervention to Mitigate Poor Sleep and Improve Well-being in the COVID-19 Pandemic: Protocol for a Feasibility Randomized Controlled Trial
JMIR Res Protoc 2022;11(3):e34409
URL: https://www.researchprotocols.org/2022/3/e34409
doi:10.2196/34409
PMID:34995204

Protocol

A Quality Improvement Emergency Department Surge Management Platform (SurgeCon): Protocol for a Stepped Wedge Cluster Randomized Trial

Hensley H Mariathas[1], MPhil, PhD; Oliver Hurley[1], BSc, MEnvSc; Nahid Rahimipour Anaraki[1], PhD; Christina Young[1], PhD; Christopher Patey[2,3], MD; Paul Norman[3], BN; Kris Aubrey-Bassler[2], MD, MSc; Peizhong Peter Wang[2,4], MD, MPH, PhD; Veeresh Gadag[5], MSc, MPhil, PhD; Hai V Nguyen[6], PhD; Holly Etchegary[2], PhD; Farah McCrate[7], PhD; John C Knight[2,8], PhD; Shabnam Asghari[1], MD, MPH, PhD

[1]Centre for Rural Health Studies, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada

[2]Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada

[3]Eastern Health, Carbonear Institute for Rural Reach and Innovation by the Sea, Carbonear General Hospital, Carbonear, NL, Canada

[4]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[5]Division of Community Health and Humanities, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada

[6]School of Pharmacy, Memorial University of Newfoundland, St. John's, NL, Canada

[7]Department of Research and Innovation, Eastern Health, St. John's, NL, Canada

[8]Newfoundland and Labrador Centre for Health Information, St. John's, NL, Canada

Corresponding Author:
Shabnam Asghari, MD, MPH, PhD
Centre for Rural Health Studies, Faculty of Medicine
Memorial University of Newfoundland
300 Prince Philip Drive
St. John's, NL, A1B3V6
Canada
Phone: 1 709 777 2142
Fax: 1 709 777 6118
Email: shabnam.asghari@med.mun.ca

Abstract

Background: Despite many efforts, long wait times and overcrowding in emergency departments (EDs) have remained a significant health service issue in Canada. For several years, Canada has had one of the longest wait times among the Organisation for Economic Co-operation and Development countries. From a patient's perspective, this challenge has been described as "patients wait in pain or discomfort for hours before being seen at EDs." To overcome the challenge of increased wait times, we developed an innovative ED management platform called *SurgeCon* that was designed based on continuous quality improvement principles to maintain patient flow and mitigate the impact of patient surge on ED efficiency. The SurgeCon quality improvement intervention includes a protocol-driven software platform, restructures ED organization and workflow, and aims to establish a more patient-centric environment. We piloted SurgeCon at an ED in Carbonear, Newfoundland and Labrador, and found that there was a 32% reduction in ED wait times.

Objective: The primary objective of this trial is to determine the effects of SurgeCon on ED performance by assessing its impact on length of stay, the time to a physician's initial assessment, and the number of patients leaving the ED without being seen by a physician. The secondary objectives of this study are to evaluate SurgeCon's effects on patient satisfaction and patient-reported experiences with ED wait times and its ability to create better-value care by reducing the per-patient cost of delivering ED services.

Methods: The implementation of the intervention will be assessed using a comparative effectiveness-implementation hybrid design. This type of hybrid design is known to shorten the amount of time associated with transitioning interventions from being the focus of research to being used for practice and health care services. All EDs with 24/7 on-site physician support (category *A* hospitals) will be enrolled in a 31-month, pragmatic, stepped wedge cluster randomized trial. All clusters (hospitals) will start

with a baseline period of usual care and will be randomized to determine the order and timing of transitioning to intervention care until all hospitals are using the intervention to manage and operationalize their EDs.

Results: Data collection for this study is continuing. As of February 2022, a total of 570 randomly selected patients have participated in telephone interviews concerning patient-reported experiences and patient satisfaction with ED wait times. The first of the 4 EDs was randomly selected, and it is currently using SurgeCon's eHealth platform and applying efficiency principles that have been learned through training since September 2021. The second randomly selected site will begin intervention implementation in winter 2022.

Conclusions: By assessing the impact of SurgeCon on ED services, we hope to be able to improve wait times and create better-value ED care in this health care context.

Trial Registration: ClinicalTrials.gov NCT04789902; https://clinicaltrials.gov/ct2/show/NCT04789902

International Registered Report Identifier (IRRID): DERR1-10.2196/30454

(JMIR Res Protoc 2022;11(3):e30454) doi:10.2196/30454

KEYWORDS

SurgeCon; emergency department; stepped wedge design; cluster randomized trials; wait time

Introduction

Background

Long wait times and overcrowding are challenging emergency departments (EDs) around the world [1-4]. Several other countries with advanced health care systems cannot keep pace with patient demand. In particular, Canada ranks among countries with the longest wait times compared with those of peer-industrialized countries [5]. The Canadian Institute for Health Information (CIHI) reported an 11% increase in ED wait times from 2015-2016 to 2016-2017 [2]. This translates to long wait times and deters patients from pursuing the necessary care they need and increases the likelihood of patients leaving the ED without being seen (LWBS) by a physician [6,7]. In Newfoundland and Labrador, Canada's easternmost province, long wait times are plaguing the province much like the rest of Canada [8-10].

The Newfoundland and Labrador provincial government and the province's 4 regional health authorities (RHAs) have the option of expanding the health care workforce [11,12] at a time of historic fiscal restraint or finding effective interventions to improve the efficiency of their current ED service [13]. Overtime [14-16], expanding the ED [17,18] and redirecting patients to primary care [19,20] have not been shown to be effective. According to an October 2020 report from the CIHI, ED services are making up a larger percentage of total hospital spending with a 4% annual growth rate, which was observed between 2005 and 2019 [21]. The same report states that ED staff are twice as likely to work overtime compared with staff in other departments [21].

We have created a quality improvement intervention called SurgeCon. As a pragmatic ED management platform, SurgeCon includes 3 separate components (described below) that together act to decrease wait times and improve the sustainability of Newfoundland and Labrador's ED services without significant workforce changes. These interventions include restructuring the ED organization and workflow, fostering a patient-centric environment, and quantifying ED demands and available resources in real time. SurgeCon is designed to enable frontline health care workers to anticipate and mitigate surges in patient volume through a series of proactive steps and decision-making tools. SurgeCon attracted the attention of the Newfoundland and Labrador Eastern Health (EH) RHA after they missed their own ED wait time benchmarks in 2016 [22]. The initial development of the SurgeCon intervention came about after an external review was completed by an independent third party to determine which areas of ED operations could be adjusted to improve wait times and departmental efficiency. The external review was one of many components included in a provincial wait time reduction initiative [23].

We piloted SurgeCon at the ED located in Carbonear, Newfoundland and Labrador, an EH administered hospital, over 45 months from July 1, 2013, to March 31, 2017. Data from the pilot study were analyzed using an interrupted time series analysis to assess its effect on ED performance. The resulting change in indicators was noteworthy, despite a 25.7% increase in patient volume. Over the course of the 45-month pilot study, average time to physician's initial assessment (PIA) decreased from 104.3 (SD 9.9) minutes to 42.2 (SD 8.1) minutes, length of stay (LOS) in the ED decreased from 199.4 (SD 16.8) minutes to 134.4 (SD 14.5) minutes, and the number of patients LWBS decreased from 12.1% (SD 2.2%) to 4.6% (SD 1.7%). All of these changes were statistically significant. The marked and sustained impact of SurgeCon on ED performance in Carbonear supports the case for its extension to other EDs.

The proposed innovative clinical trial and the implementation of SurgeCon (see the implementation paper) [24] will generate practical information on its effectiveness in a range of urban and rural ED settings and generate data to support its more comprehensive implementation. Given the successful results from the pilot study, and if SurgeCon proves to address ED patient flow issues in this study, the rest of Canada and other countries could significantly benefit from its implementation.

Given the scope of the SurgeCon research program, a separate protocol related to evaluating the intervention's implementation was published in a separate article [24]. This research protocol focuses on the innovative clinical trial stepped wedge cluster randomized trial (SW-CRT) design used to assess the effectiveness of the intervention. In this protocol paper, we will provide a brief description of the intervention and associated

research activities that will be carried out at each of the 4 selected ED sites.

Study Aim and Objectives

In this study, we present the protocol for a 31-month SW-CRT trial. Our aim is to evaluate the performance of the SurgeCon platform in improving important patient and service process outcomes in EDs in an RHA in Newfoundland and Labrador and develop strategies to promote its scalability, sustainability, and successful implementation across the Canadian health system. The primary objective of the trial is to evaluate the effects of SurgeCon on ED performance by assessing its impact on LOS, PIA, and LWBS. The secondary objective of the trial is to assess the intervention's impact on patient satisfaction and patient-reported experience with ED wait times and its ability to create better-value ED services by reducing ED costs.

Methods

Ethics Approval

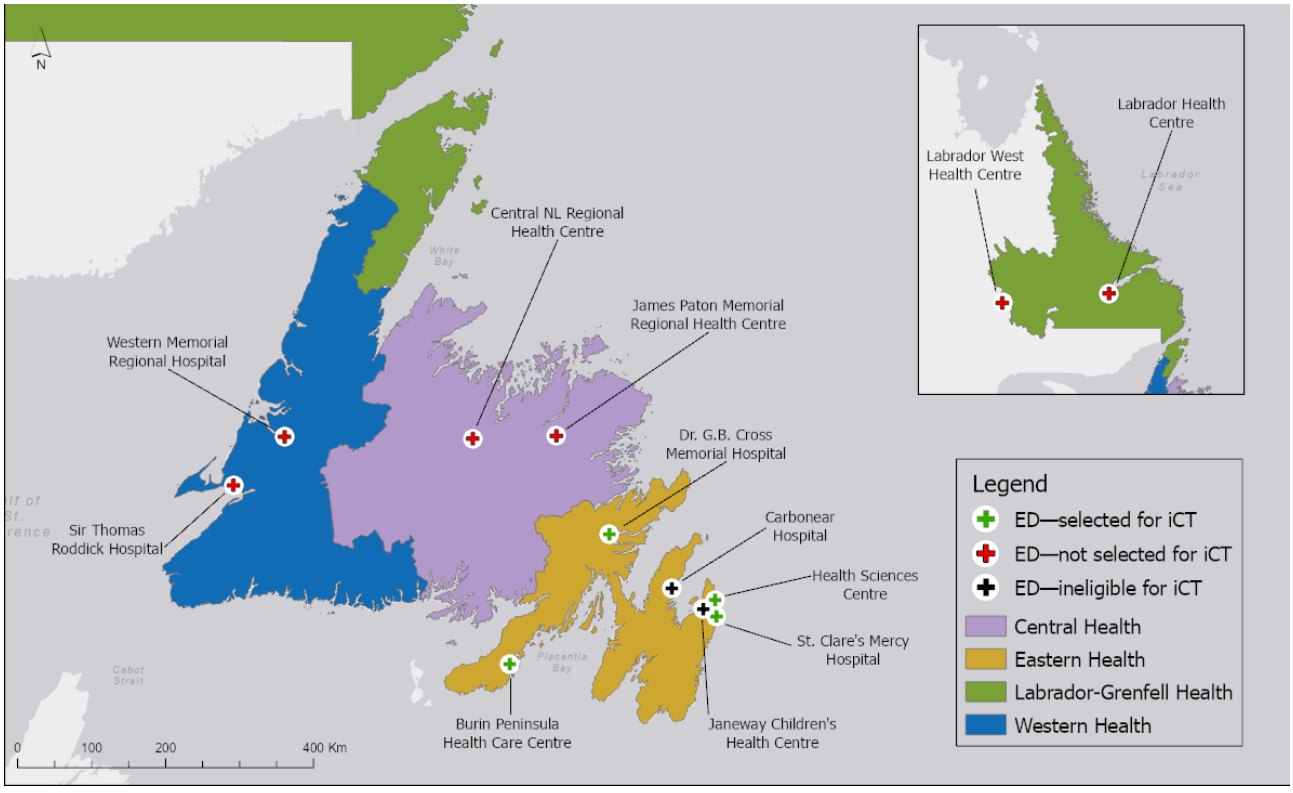
This study has been approved by the Newfoundland Labrador Health Research Ethics Board with researcher portal file 20201482.

Study Setting

Newfoundland and Labrador's health care system is delivered through four RHAs: EH, Central Health, Western Health, and Labrador-Grenfell Health. The eastern RHA will be participating in this research initiative as a collaborative research partner and will be the only RHA participating in the study. The other 3 health authorities will act as knowledge users and will not be included in the SurgeCon trial. However, the research team will provide interim reports to the other health authorities to allow them to monitor and learn from research findings over the course of the study period to guide future implementation in their own centers.

This is a multisite study, including 2 urban and 2 rural EDs with 24/7 on-site physician support in the EH region of Newfoundland (Figure 1). The remaining 2 EDs with 24/7 on-site physician support within the study area that are not receiving the intervention are a pediatric ED, which operates differently and has wait times that differ greatly from adult or general ED, and the Carbonear General Hospital, which was chosen for the pilot study. The 4 sites receiving the intervention include 2 urban sites (Health Sciences Centre and St. Clare's Mercy Hospital) and 2 rural sites (Dr. G.B. Cross Memorial Hospital and Burin Peninsula Health Care Centre).

Figure 1. Emergency department (ED) locations in 4 regional health authorities of Newfoundland and Labrador (NL). iCT: innovative clinical trial.



Study Population

All individuals who visited any of the 4 selected EDs during the study period will be included in the collection of deidentified health administrative data. We will also collect and monitor ED-level key performance indicator (KPI) data such as LOS, PIA, and LWBS because they are impacted by patient volume. Patients who receive care at these EDs will be randomly selected for subsequent follow-up after they are discharged to collect information related to their experience in the ED and to determine their level of satisfaction with the care that they received.

Site Randomization

As we will be evaluating the effectiveness of SurgeCon using an SW-CRT, the 4 different hospitals will be randomized, with 1 ED site starting at the first sequence. The next site will start 6 months later in the second sequence and will continue until

all EDs are allocated to sequences. A simple random sample will be performed by the study statistician, who will generate a randomization list using statistical software that will determine the order of intervention implementation. The researchers and participants will not be blinded to whether they are in the intervention or the control cohort.

Patient Randomization

To assess patient-reported experiences and satisfaction, discharged patients who are subsequently contacted to complete satisfaction and patient experience surveys will be randomly selected using a random time and date generator program.

Study Outcomes

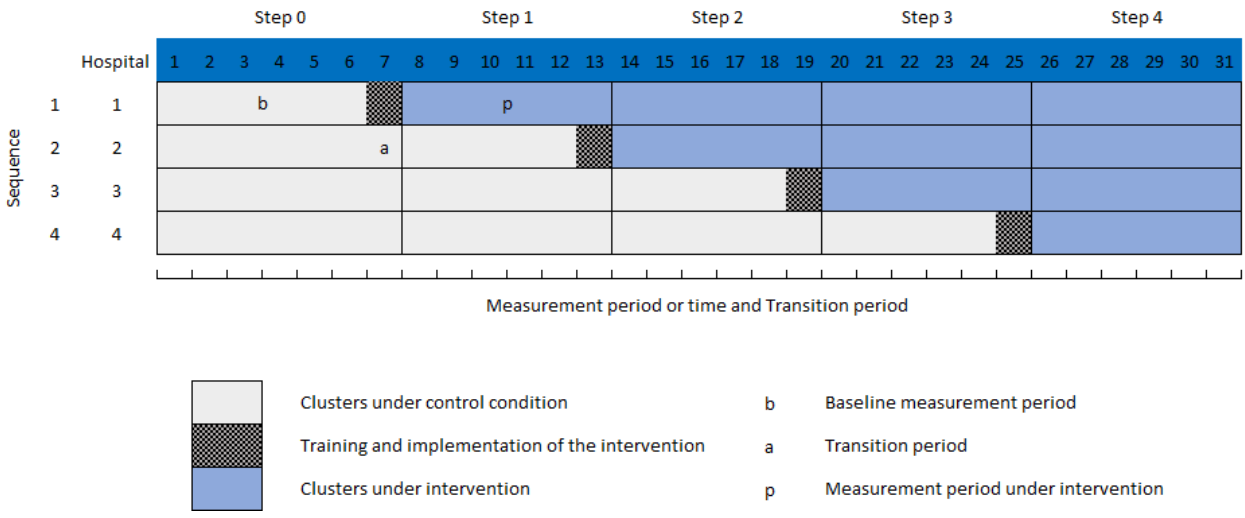
The SurgeCon study aims to measure both ED KPIs (LOS, PIA, and LWBS) and patient perceptions and satisfaction related to the care they received in the ED. Patient health outcomes, use of health care resources, and overall cost will also be assessed based on a patient's arrival time. We will also collect information on the potential adverse effects of the intervention such as 24-hour readmission and mortality. These data will help examine trends in mortality and readmissions before and after intervention implementation.

The study uses a comparative effectiveness-implementation hybrid design and includes outcomes for effectiveness and implementation. Two measurement levels will be considered for the study, ED- or service-level and patient-level outcomes will be used to determine the effectiveness of SurgeCon. To guide our choice of outcomes and to determine how best to evaluate the intervention's implementation, we applied a combination of 2 frameworks, that is, RE-AIM (reach, effectiveness, adoption, implementation, and maintenance) and Consolidated Framework for Implementation Research [25,26]. To guide our implementation strategy, we will collect data related to organizational climate, fidelity of staff training, fidelity of intervention delivery, implementation cost, barriers and enablers to adoption, implementation, institutionalization, intervention acceptability, intervention appropriateness, feasibility of maintaining SurgeCon, sustainability of SurgeCon, and scalability of the SurgeCon intervention. Detailed information on implementation assessment and outcomes is available via our other work [24].

Study Design

The SW-CRT design is a novel, robust, and flexible 1-way crossover cluster randomized trial design increasingly being used in trial arms with varying time delays in which all clusters start from the control condition to the active intervention condition state. In particular, this longitudinal stepped wedge study design includes a repeated cross-sectional design, as illustrated in Figure 2, where each hospital will eventually receive the intervention. All participating sites will begin the trial in a control condition where they will continue to use a usual care model or the model of care provided before the beginning of the trial. Each site will switch from providing usual care to providing care using the intervention that will take place at predetermined time periods during the study. At the end of the trial experiment, each of the sites will have implemented all of the intervention's components. The stepped wedge randomized trial design used in this study is normally carried out at the cluster level rather than at the individual level. Therefore, a clustered randomized stepped wedge design will be the focus of this protocol.

Figure 2. Schematic representation of a SW-CRT with 4 steps for a 31-month study period. SW-CRT: stepped wedge cluster randomized trial.



The advantages of the SW-CRT include logistical flexibilities, efficiencies in terms of power, and sample size compared with traditional (clustered) parallel-group designs. Furthermore, the ethical advantages in longitudinal and open cohort studies have also been recognized [27-29]. A simple random sampling technique will be used to determine the order of intervention implementation across the 4 selected sites. The assignment of 1 cluster per sequence will maximize statistical power. The study period can be subdivided by four 6-month steps. Each step starts with the implementation of the intervention at one of the 4 selected sites. Observations are collected repeatedly from each cluster in multiple periods. The key parameters such as the number of clusters or clusters per sequence, steps, and measurement periods (both control and intervention conditions), intracluster correlation coefficient (ICC), and effect size are required for sample size calculation for the SW-CRT design.

The total study period available for the intervention to remain active is 31 months, and it is expected that at least one month will be needed for training and implementation at each ED site. A 1-month intervention adoption period was considered to be sufficient for the intervention to be fully operationalized for a single cluster. During the first 6 months of the baseline period, all patients who visit any of the 4 hospitals will receive usual care. After the baseline period, a randomly allocated hospital will begin SurgeCon intervention implementation every 6 months and will continue to use the intervention for operations and management until the end of the study. Staff at each of the randomly selected hospitals will undergo training, establish processes and guidelines that are consistent with efficiency principles covered during training, and begin routine data entry. During the last 6 months of the study period, all 4 hospitals will be operating using the SurgeCon ED management platform exclusively. We will be monitoring the characteristics of each hospital and the composition of their associated frontline teams to determine if significant changes have occurred since the control period and whether an adjustment to our analysis plan is required.

Sample Size

On the basis of the results of a pilot study conducted at a rural ED site in EH's jurisdiction, we established a common framework for optimal sample size calculations when the number of clusters available for randomization is limited to 4 sites for the SW-CRT design. We chose 1 cluster, which will cross over to an intervention state at a randomly assigned step. To keep things relatively simple, we assume that an equal number of observations is sampled in each cross section of each cluster. Alternatively, we expect an equal amount of observations per period per cluster. Results from our pilot study show a 15% to 30% decrease in the LOS or wait times between 6 and 30 months after SurgeCon implementation [30-32]. This trial is powered to detect a 15% change in LOS at 5% of type I error and 80% power with an ICC [33] of 0.1 for repeated measures (6 measures per step). A 10% reduction in ED wait time in the study sites could result in 10-minute reduction in LOS. To detect this change, a minimum sample size of N=20,280 (169/month/hospital for 30 months and 4 hospitals) is required. To be able to conduct age, sex, and patient acuity subgroup analyses for all ED visits across the 4 intervention sites, we will include all ED visits for this portion of the analysis as there is a low cost associated with extracting data from existing ED repositories, which routinely capture patient record-level information. For patient satisfaction, the study is powered to detect a 30% change in patient-reported experience measures or patient satisfaction [34-37] at a 5% of type I error and 80% power with an ICC of 0.1. Therefore, a sample size of 1320 (11/month/hospital for 30 months and 4 hospitals) would be sufficient for this study. Considering a 50% response rate, we will conduct 25 surveys per month per hospital. In particular, we demonstrated the required sample size and power calculation procedure with illustration for different combinations of the ICC to detect the standard effect size. For this illustrative purpose, we consider a clinical trial powered to detect a 10% to 30% reduction in wait times on a continuous scale at the 4 different EDs with 5% of type I error rate and 80% power.

Data Collection and Monitoring

Aggregate- and individual record-level data will be analyzed over the course of the study period to assess the effect of the intervention on patient outcomes, health service efficiency, and the cost of providing emergency care. Our analysis will include aggregate KPI data, record-level health administrative data, and aggregate financial data and will be provided to the research team by the Newfoundland and Labrador Centre for Health Information. The ED KPIs (ie, LOS, PIA, and LWBS) used as the primary outcomes for this study are further described in Table S1 in Multimedia Appendix 1 alongside other variables that will be used for statistical modeling. Each of the primary outcomes will be calculated and assessed monthly over the course of the study period. The total number of monthly ED visits will be used as the denominator for our cluster-level summaries.

Individual record-level health administrative data will include demographic variables (eg, sex, date of birth, and postal code), mortality data, wait times data, diagnosis data, and triage acuity scores, among other variables. These data will be used to create aggregate KPI and financial data provided by Newfoundland and Labrador Centre for Health Information and to assess the intervention's impact on the type and volume of patients who visit one of the 4 selected ED sites during the study period. Financial data will capture expenditures for staff, supplies, and procedures originating from the ED and will include data related to pharmacy, diagnostic imaging, laboratory testing, surgical day care, operating room procedures, physician salaries, physician fees for service claims, nurse salaries, administrative staff salaries, ambulance services, and other related ED costs.

Patient-reported experiences and patient satisfaction survey data will be collected 3 to 5 days after ED discharge via a telephone interview conducted by a research assistant who is also an EH employee. No identifiable information related to their ED visit will be collected during the survey, and patient consent will be obtained before the interview through an implied consent process that does not require a signed consent form. We provide all study participants with the opportunity to request an informed consent document that can be provided by email or post mail and contains contact information for the principal investigator and project manager. To minimize the loss to follow-up, the research assistant will attempt to reach patients up to 3 times within 2 weeks of the initial attempt. The interview will take <30 minutes, using a questionnaire we created based on CIHI's patient-reported experiences [38] and patient satisfaction surveys for EDs [39]. Although most of Newfoundland and Labrador's population is English speaking (97%) [40], language was not considered an exclusion criterion. Patient interview responses will be stored via Qualtrics (a web-based survey program), where research team members will be able to analyze the data while maintaining patient anonymity. An important consideration for this approach is that the only ethically approved means of contacting patients is by telephone using patient contact information at EH. Other options were explored such as surveying patients directly in the ED, but it was deemed unsuitable, as many patients are not likely to be in a physical or mental state conducive to participating in a study.

Beyond the data provided by traditional provincial data custodians, we will also consider the SurgeCon platform's routinely collected aggregate data. These data include calculated SurgeCon levels, ED beds, bed availability, patient acuity, and patient process tracking. The SurgeCon action-based protocol is subdivided into 5 levels of escalation and is used to indicate the level of demand, availability of resources, and capacity in the ED. The levels range from 1 (optimal operating conditions) to 5 (very busy or patient surge). They are calculated using several variables such as ED and inpatient unit bed availability, resource shortages, and number of patients in the waiting room left to be triaged or seen by a physician, among many other options. A charge nurse or other frontline staff manually enter variables used by SurgeCon's algorithm to determine SurgeCon levels via a data entry portal. Any data entered into the data entry portal will be made accessible to the research team through the creation of a special user role that provides access to all site-specific information and the ability to export collected data.

Our implementation strategy includes the operationalization of routine data capturing by charge nurses. The research team will regularly monitor data quality and completeness across all study sites. We will be working with frontline health care staff at each of the sites to find solutions when sites are found to be missing data entry intervals or if data quality is low. We are exploring opportunities to automate data collection for certain variables if they are found to be feasible and appropriate.

Statistical Analysis

Analyses

The characteristics of hospitals and patients will be recorded. We will describe the clinical and demographic characteristics of hospitals and patients for each period. We will report the response for patient satisfaction surveys and report the duration of intervention adherence at each study site and, if applicable, the reasons for noncompliance. For each outcome, we will report the results for each period, including the effect size and its precision. The 30 months of cumulatively collected data from EH and provincial health administrative databases in both the intervention and usual care periods will be modeled as a linear mixed model. We plan to analyze research data using a generalized linear mixed model (GLMM) [41] or generalized estimating equation [42]. Hussey and Hughes [43] have suggested a model-based approach for analyzing data using a repeated cross-sectional design where outcome measurements will be measured from different individuals at each measurement interval. This approach was proposed for continuous outcome variables and has been commonly used at the design stage of these studies [44]. Individuals within the same cluster are likely to be positively correlated, and the strength of the correlation can be measured by the ICC under this model and is assumed to be constant over time. However, the model suggested by Hussey and Hughes [43] has been extended to allow a more general correlation structure between individuals within the same cluster [45-47]. We will adhere to the intention-to-treat analysis; however, sensitivity analyses for comparing the results under the intention-to-treat assumption with the complier and per protocol will be conducted. Moreover, interim analyses will be conducted on a monthly basis to inform us about the findings in a timely manner and allow us to make any modifications if required [48].

The primary method of analysis used in this study is GLMM, which will consider monthly cluster-level summary measures such as means of LOS, PIA, or proportion of LWBS. Our GLMM models will include fixed effects for time, intervention effects, random effects, and random time effects for each cluster. In addition, ED setting (rural or urban), ED volume, and size of ED administrative resource will be added as cluster-level covariates. For secondary outcomes such as patient-reported experience and satisfaction, our individual-level GLMM models will include fixed effects for time, intervention effect, and random effects and random time effects for each cluster, with the covariates consisting of age, gender, the reason for ED visit, and cluster-level LOS, PIA, and proportion of LWBS.

Analytical Considerations

In the basic model suggested by Hussey and Hughes [43], a homogeneous secular trend is assumed across all clusters. However, this SW-CRT has numerous methodological difficulties such as confounding with time, time-varying correlation structure, change in treatment effect over time, within-cluster contamination, and change in design variation. These complexities differ according to the way SW-CRTs are designed. A summary of key methodological issues that need extra consideration when reporting SW-CRT is presented in Table S2 in Multimedia Appendix 1. We will be continuously monitoring and intervening when necessary to manage analytical challenges associated with this kind of design. Any unexpected event during the study period will be recorded, and methodological approaches to overcome the issue will be described in the reports.

Patient Engagement

As a patient-oriented study, the inclusion of patients in the research team is critical. An example of this is our lead patient research partner who has been advising the research team since the pilot study. The research team is also advised by the patient engagement working group, which comprises the lead patient partner, scientific patient engagement lead, clinicians, researchers, and students. We designed the protocol to give patients a variety of opportunities to participate in research activities. Applicable research activities vary in terms of the level of engagement required (eg, surveys vs full-team membership). We are currently recruiting additional patient research partners for the study's patient engagement working group, but patients can also be involved in other committees and working groups suited to their interests and needs. Overall, the level of engagement will vary from receiving information to consultation to full collaboration with patient partners to inform all project activities. We used 4 essential pillars, as suggested by Shippee et al [49], to inform our patient engagement strategy, which is further described in our implementation protocol [24]. We are committed to upholding the guiding principles of inclusiveness, support, mutual respect, and cobuilding inherent in these pillars.

Impact of COVID-19

The impact of the COVID-19 pandemic on the study and its protocol is ongoing. Public health care measures such as stay-at-home orders, handwashing, and physical and social distancing compounded with COVID-19 concerns have caused rapid disruptions in daily life and a delay in the implementation of the intervention. A notable effect may be a significant change in the number of patients visiting EDs. The fear that the public is experiencing because of COVID-19 is likely exacerbated by measures related to stay-at-home orders, self-isolation protocols, including quarantines, travel restrictions, and closures of nonessential businesses. With most of Newfoundland and Labrador's residents practicing social distancing and self-isolating and an increase in the number of workers who are now working remotely, the potential for injuries such as trauma due to motor vehicle collisions may decline considerably [50]. Given the potential reduction in injury and the climate of fear at the beginning of the COVID-19 pandemic, patients were found to be less likely to use hospital ED services [51]. Another impact of the pandemic is the temporary suspension of essential and nonessential medical procedures. The suspension of certain health care services may have downstream effects on wait times because delays in surgeries and other important medical procedures could increase the number of higher acuity patients who develop complications because of the delay. These higher acuity patients require additional resources and time and can have a significant impact on patient flow and wait times. Other pandemic-related considerations include added wait times because of public health measures such as increased handwashing, social distancing, donning and doffing of personal protective equipment, and sanitizing high-touch surfaces. More recently, a transition from in-person to web-based or telephone family physician consultations may be increasing ED patient volumes [52,53]. The pandemic has also exacerbated issues related to physician and nursing shortages in the province [54-56]. As a result, achieving optimal sample size from each of the ED sites might be compromised because of pandemic-related ED service use trends. The research team will look to explore the effects of the COVID-19 pandemic on patient flow and wait times in EDs.

Description of Intervention

Our intervention comprises several components that will be implemented sequentially. The SurgeCon intervention is a pragmatic ED management platform that includes 3 distinct intervention components that together act to improve ED efficiency, patient satisfaction, and the value of emergency service costs (Table 1). SurgeCon's intervention process starts with a site assessment during the transition period that aims to clarify key performance issues, collect information related to the ED's organizational and workflow structures, and prepare ED staff and management for upcoming operational changes, while also establishing a patient-centered ED environment and action-based ED management (Figure 3). Due to the high degree of variability that exists between EDs, the information collected during the site assessment allows for the customization of SurgeCon's underlying protocols and determines whether certain components of the intervention are appropriate or applicable for implementation. A site assessment will be conducted by members of our working group who developed SurgeCon in Carbonear. The implementation working group will follow a 4-step protocol for these site assessment visits; for further details please see the implementation protocol [24].

Figure 3. The SurgeCon intervention logic model. ED: emergency department; LOS: length of stay; LWBS: leaving the emergency department without being seen; PIA: physician's initial assessment; PREM: patient-reported experience measure.

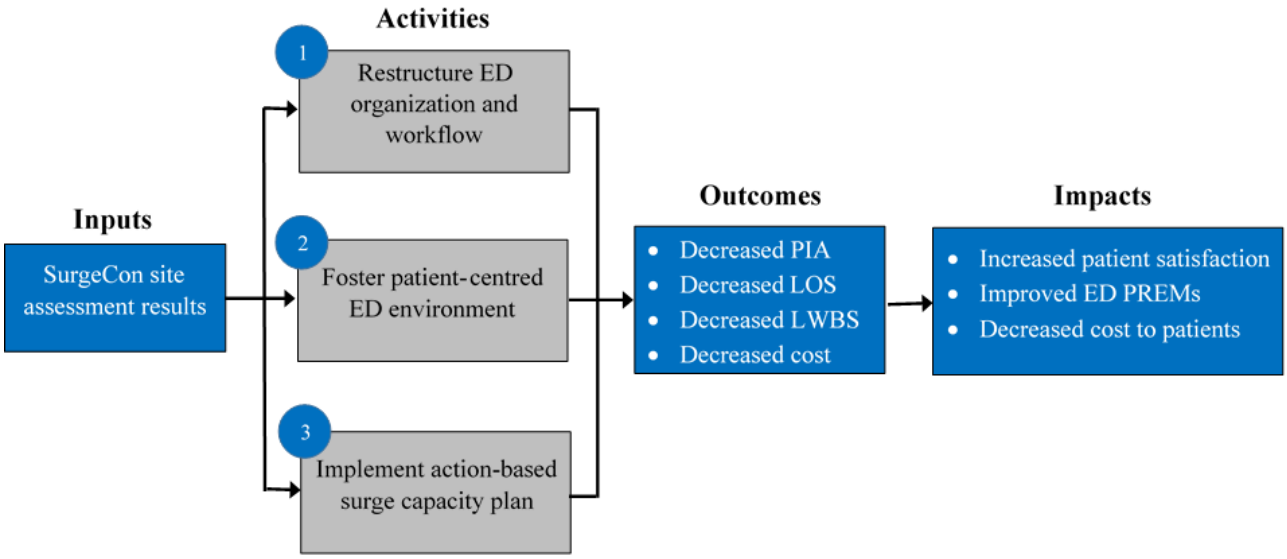


Table 1. Intervention components and their associated action plans.

Intervention components and action plans	Description and strategy
Restructure ED[a] organization and workflow	
Stable patient priority setting	Canadian EDs have been using the CTAS[b] since 1999 [57]. What the SurgeCon intervention proposes is to augment CTAS by combining CTAS scores and waiting times to best meet the suggested time frame set out by the CTAS organization. Patients who are CTAS level 1 (resuscitation) or 2 (emergent) who are not stable will still be treated immediately in this model, but individuals who are less urgent (CTAS 3-5) will be treated when the CTAS 1-2 individuals become more stable, but before they are discharged. This allows the physician to quickly treat and discharge less urgent individuals before returning to spend more time with the urgent cases. This prioritization method can significantly improve patient flow without having to compromise patient safety.
Door-to-door focus	A number of studies have found a strong correlation between patient satisfaction and PIA[c]; the shorter the PIA, the more satisfied the patient [58,59]. To reduce the time to PIA, we will use the following strategies: 1. ED physicians and frontline providers will triage with nursing staff with the goal of increased patient discharge from the fast-track area or triage room without waiting. 2. Triage nurse–driven orders (eg, symptom management, laboratory testing, and diagnostic imaging) will only be applied on patients who would be waiting longer than 1 hour to see a physician. If the patient can be seen by a physician within an hour, waiting for potentially unnecessary test results could delay the PIA. ED physicians will review patients arriving on ambulance stretchers in the hallway if there are no available beds instead of waiting for a bed to be free.
Nurse practitioner–physician communication	SurgeCon reorganizes the traditional Canadian ED communication structure by promoting communication between NPs[d], ERPs[e], and RNs[f] to work collaboratively to improve patient flow through appropriate allocation of patients. By opening lines of communication between NPs, ERPs, and RNs, the entire ED can work in unison to move patients further along their path to being discharged.
Establish a patient-centered ED environment	
Improve the overall appearance of physical spaces in the ED (eg, waiting room, fast-track zone, examination rooms, and treatment space) to improve patient satisfaction	This will be conducted in collaboration with Eastern Health. In consultation with our local patient partners, we will renovate, redecorate, and declutter ED spaces, removing outdated or irrelevant wall postings. All subsequent wall postings will require departmental approval and will be placed in a central location.
eHealth action–based ED management	
eHealth ED management solution	Using automated extraction of data from HISs[g] where possible and manual entry otherwise, SurgeCon's digital component will be able to perform real time analysis on extracted data in a routine and timely manner to give ED staff a sense of overall demand and available resources at any given time. SurgeCon's eHealth component will be installed and tested in each hospital during the adoption phase of the study. In situations where data elements might not be captured by existing ED repositories, an ED staff member will address this issue by manually entering specific variables (eg, number of ambulances waiting to be off-loaded) as part of their regular duties. Manual entry occurs via a web-based SurgeCon portal and is subsequently reported at a frequency that is determined by ED staff and management. SurgeCon's eHealth component is currently available on desktops and mobile devices and has been deployed in Eastern Health's secure network with the assistance of the Newfoundland and Labrador Centre for Health Information. The data entry portal and dashboard that provides real time data is normally displayed on a large digital whiteboard in close proximity to the nursing station. This allows all team members to have a clear understanding of the level of demand, current capacity, and available resources. The Carbonear pilot site has since operationalized the task of a 2-hour data entry interval and have done so without a significant workforce change. We have found that a 2-hour interval for data entry is a feasible target and can be quickly performed once certain reporting processes have been established [30,31]. We have further developed and tested the digital application at our pilot site. The development of the application will continue throughout the study to ensure feedback and information are incorporated into iterative software updates.

Intervention components and action plans	Description and strategy

eHealth action–based protocol	We have created a unique frontline, action-based protocol that helps ED staff (paramedics, nurses, and physicians) manage their actions to actively reduce patient surges and wait times and increase patients' access to emergency medical care. The protocol is delivered via a digital whiteboard app, which will be installed in the nursing station in the ED. The app uses algorithms (adjusted to meet the needs of each hospital) to advise when to use volume-based staffing (shifting staff between areas of the hospital based on workload), appropriate and timely involvement of hospital management, and overcapacity protocols, which may otherwise be overlooked by distracted frontline ED staff. All intervention sites will be routinely collecting data related to staff availability, ED and inpatient bed availability, aggregate patient acuity, and process tracking, among other important variables through SurgeCon's eHealth component. A SurgeCon level is calculated via an algorithm that uses variable data to determine the level of demand in the department and resource availability to meet the demand. The action-based protocol included in SurgeCon's eHealth platform assigns actions based on the SurgeCon level calculated.
	The following list includes examples of actions that may be assigned once a threshold for a specific variable has been exceeded:
	• Observation: Patients admitted in ED
	  • Action: Notify charge nurse on accepting unit to create a plan for timely transfer of admissions
	• Observation: Critical patients (1:1 nursing care)
	  • Action: Notify ICU[h] to plan for help
	• Observation: Pending transfer out
	  • Action: If the flight team requires it, make appropriate arrangements

[a]ED: emergency department.

[b]CTAS: Canadian Triage and Acuity Scale.

[c]PIA: physician's initial assessment.

[d]NP: nurse practitioner.

[e]ERP: emergency room physician.

[f]RN: registered nurse.

[g]HIS: hospital information system.

[h]ICU: intensive care unit.

Intervention Components

The SurgeCon intervention is guided by continuous process improvement principles that look to ultimately improve quality of services, reduce waste (low-value care), reduce time (ED wait times), and reduce cost (eg, cost per patient, cost of overtime, and cost to patients) [60-63]. The following components of the SurgeCon intervention will follow these principles closely and will be implemented using other continuous improvement methods such as *Kaizen* events [64]. The SurgeCon intervention includes 3 components (Figure 3).

The major intervention components and their action plans with description and strategy are given in Table 1.

Results

This study was funded in April 2019, was approved by the Newfoundland and Labrador Health Research Ethics Board on March 19, 2020, and is now registered with ClinicalTrials.gov. Data collection for this study is ongoing. As of February 2022, a total of 570 randomly selected patients have participated in telephone interviews focusing on patient-reported experiences and patient satisfaction with ED wait times. The first intervention site was randomly selected and began intervention implementation in September 2021. Since that time, SurgeCon's eHealth component has been configured, and deployed and health care staff have received ED patient flow training. The second intervention site is scheduled to be randomly selected and begin intervention implementation during the winter of 2022.

Discussion

Principal Findings

The SurgeCon intervention has the potential to be a scalable solution that can address ED wait times. In this paper, we have described the protocol for a stepped wedge design, which is a relatively new type of study design that is progressively being used to evaluate the efficiency of public health services. This study design is known for its application in assessing the implementation of evidence-based quality improvement initiatives in health care settings. We chose the stepped wedge design as an informative, efficient, and valid design to examine the SurgeCon platform's effectiveness in improving important patient and health service processes and outcomes in EDs located in a single region of Newfoundland and Labrador.

Conclusions

By assessing the impact of the SurgeCon intervention on the efficiency of ED services, we hope to be able to improve wait times and patient experiences and produce better-value care.

Acknowledgments

The work summarized in this protocol is supported by the Canadian Institutes of Health Research and provincial sponsors, including the Newfoundland and Labrador Provincial Government (Department of Tourism, Culture, Industry, and Innovation), Newfoundland and Labrador Eastern Regional Health Authority, and the Trinity Conception Placentia Health Foundation. The authors would like to thank Dr James Rourke for his input related to the initial conceptualization of the study and Newfoundland and Labrador Support for People and Patient Oriented Research and Trials for advising the research team on matters related to patient engagement and knowledge translation.

Authors' Contributions

All authors contributed to the concept of the study. HHM and SA developed the study content and drafted the manuscript. All authors reviewed the manuscript, provided feedback, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Definitions, sources, and purpose of study variables; methodological issues encountered; and applicable strategies to overcome them.
[DOCX File , 23 KB - resprot_v11i3e30454_app1.docx]

References

1. Torjesen I. Latest waiting time figures for emergency departments in England are worse on record. BMJ 2018;361:k1658. [doi: 10.1136/bmj.k1658] [Medline: 29650511]
2. Emergency department wait times in Canada continuing to rise. Canadian Institute for Health Information. 2018. URL: https://www.cihi.ca/en/emergency-room-waits-continue-to-increase [accessed 2018-07-04]
3. Dramatic rise in waiting times at emergency departments. BBC News. 2018. URL: https://www.bbc.com/news/uk-northern-ireland-44562129 [accessed 2018-11-28]
4. Katz A, Enns J. How long is too long to wait in an emergency room? Huffington Post. 2018. URL: https://www.huffingtonpost.ca/alan-katz/emergency-room-waits-canada_b_16132058.html [accessed 2018-12-14]
5. The commonwealth fund 2010 international health policy survey in eleven countries. The Commonwealth Fund. 2010. URL: https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_chartbook_2010_pdf_2010_ihp_survey_chartpack_full_12022010.pdf [accessed 2012-06-26]
6. Vezyridis P, Timmons S. National targets, process transformation and local consequences in an NHS emergency department (ED): a qualitative study. BMC Emerg Med 2014;14:12 [FREE Full text] [doi: 10.1186/1471-227X-14-12] [Medline: 24927819]
7. Chang AM, Lin A, Fu R, McConnell KJ, Sun B. Associations of emergency department length of stay with publicly reported quality-of-care measures. Acad Emerg Med 2017;24(2):246-250 [FREE Full text] [doi: 10.1111/acem.13102] [Medline: 27640877]
8. Barker J. Labrador senior waits 10 hours in emergency room without care for broken arm. CBC News. 2018. URL: https://www.cbc.ca/news/canada/newfoundland-labrador/goose-bay-emergency-wait-time-1.4582333 [accessed 2018-11-30]
9. Hutton F. Picture of senior waiting for hours in St.John's emergency room sparks outrage. CBC News. 2018. URL: https://www.cbc.ca/news/canada/newfoundland-labrador/senior-hsc-waiting-social-media-1.4587954 [accessed 2018-11-30]
10. Breen K. St. Anthony patients kept on stretchers while beds remained in storage. CBC News. 2018. URL: https://www.cbc.ca/news/canada/newfoundland-labrador/hospital-bed-closure-st-anthony-1.4673014 [accessed 2018-11-30]
11. Carter AJ, Chochinov AH. A systematic review of the impact of nurse practitioners on cost, quality of care, satisfaction and wait times in the emergency department. CJEM 2007;9(4):286-295. [doi: 10.1017/s1481803500015189] [Medline: 17626694]
12. Maggio PM, Brundage SI, Hernandez-Boussard T, Spain DA. Commitment to COT verification improves patient outcomes and financial performance. J Trauma 2009;67(1):190-195. [doi: 10.1097/TA.0b013e3181a51b2f] [Medline: 19590334]
13. Jarvis PR. Improving emergency department patient flow. Clin Exp Emerg Med 2016;3(2):63-68 [FREE Full text] [doi: 10.15441/ceem.16.127] [Medline: 27752619]
14. Schroeppel TJ, Sharpe JP, Magnotti LJ, Weinberg JA, Croce MA, Fabian TC. How to increase the burden on trauma centers: implement the 80-hour work week. Am Surg 2014;80(7):659-663. [Medline: 24987896]
15. Schroeppel TJ, Sharpe JP, Magnotti LJ, Weinberg JA, Croce MA, Fabian TC. How to further decrease the efficiency of care at a level I trauma center: implement the amended resident work hours. Am Surg 2015;81(7):698-703. [Medline: 26140890]

16. Christmas E, Johnson I, Locker T. The impact of 24 h consultant shop floor presence on emergency department performance: a natural experiment. Emerg Med J 2013;30(5):360-362. [doi: 10.1136/emermed-2012-201369] [Medline: 22660466]
17. Han JH, Zhou C, France DJ, Zhong S, Jones I, Storrow AB, et al. The effect of emergency department expansion on emergency department overcrowding. Acad Emerg Med 2007;14(4):338-343 [FREE Full text] [doi: 10.1197/j.aem.2006.12.005] [Medline: 17400996]
18. Mumma BE, McCue JY, Li CS, Holmes JF. Effects of emergency department expansion on emergency department patient flow. Acad Emerg Med 2014;21(5):504-509 [FREE Full text] [doi: 10.1111/acem.12366] [Medline: 24842500]
19. van Veelen MJ, van den Brand CL, Reijnen R, van der Linden MC. Effects of a general practitioner cooperative co-located with an emergency department on patient throughput. World J Emerg Med 2016;7(4):270-273 [FREE Full text] [doi: 10.5847/wjem.j.1920-8642.2016.04.005] [Medline: 27942343]
20. Morin C, Choukroun J, Callahan JC. Safety and efficiency of a redirection procedure toward an out of hours general practice before admission to an emergency department, an observational study. BMC Emerg Med 2018;18(1):26 [FREE Full text] [doi: 10.1186/s12873-018-0173-6] [Medline: 30134934]
21. Hospital spending: focus on the emergency department. Canadian Institute for Health Information. Ottawa: CIHI; 2020. URL: https://www.cihi.ca/sites/default/files/document/hospital-spending-highlights-2020-en.pdf [accessed 2021-12-06]
22. Eastern health's annual performance report 2015-2016. Eastern Health. 2016. URL: https://issuu.com/easternhealth/docs/rpt_eastern_health_apr_2015-16_fina [accessed 2017-01-09]
23. A strategy to reduce emergency department wait times in Newfoundland and Labrador. Health and Community Services, Government of Newfoundland. 2012. URL: https://www.gov.nl.ca/hcs/files/wait-times-pdf-emergency-department-strategy.pdf [accessed 2021-12-06]
24. Anaraki NR, Jewer J, Hurley O, Mariathas H, Young C, Norman P, et al. Implementation of an ED surge management platform: a study protocol. Research Square (forthcoming) 2021:-. [doi: 10.21203/rs.3.rs-764312/v1]
25. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
26. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
27. Barker D, McElduff P, D'Este C, Campbell MJ. Stepped wedge cluster randomised trials: a review of the statistical methodology used and available. BMC Med Res Methodol 2016;16:69 [FREE Full text] [doi: 10.1186/s12874-016-0176-5] [Medline: 27267471]
28. Beard E, Lewis JJ, Copas A, Davey C, Osrin D, Baio G, et al. Stepped wedge randomised controlled trials: systematic review of studies published between 2010 and 2014. Trials 2015;16:353 [FREE Full text] [doi: 10.1186/s13063-015-0839-2] [Medline: 26278881]
29. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. BMC Med Res Methodol 2006;6:54 [FREE Full text] [doi: 10.1186/1471-2288-6-54] [Medline: 17092344]
30. Asghari S. SurgeCon: an emergency department surge management platform. Clinical Trials. 2021. URL: https://clinicaltrials.gov/ct2/show/NCT04789902 [accessed 2021-03-08]
31. Patey C, Asghari S, Norman P, Hurley O. Redesign of a rural emergency department to prepare for the COVID-19 pandemic. CMAJ 2020;192(19):E518-E520 [FREE Full text] [doi: 10.1503/cmaj.200509] [Medline: 32317277]
32. Patey C, Norman P, Araee M, Asghari S, Heeley T, Boyd S, et al. SurgeCon: priming a community emergency department for patient flow management. West J Emerg Med 2019;20(4):654-665 [FREE Full text] [doi: 10.5811/westjem.2019.5.42027] [Medline: 31316707]
33. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. Ann Fam Med 2004;2(3):204-208 [FREE Full text] [doi: 10.1370/afm.141] [Medline: 15209195]
34. Heaton HA, Castaneda-Guarderas A, Trotter ER, Erwin PJ, Bellolio MF. Effect of scribes on patient throughput, revenue, and patient and provider satisfaction: a systematic review and meta-analysis. Am J Emerg Med 2016;34(10):2018-2028. [doi: 10.1016/j.ajem.2016.07.056] [Medline: 27534432]
35. Bellamkonda VR, Kumar R, Scanlan-Hanson LN, Hess JJ, Hellmich TR, Bellamkonda E, et al. Pilot study of Kano "Attractive Quality" techniques to identify change in emergency department patient experience. Ann Emerg Med 2016;68(5):553-561. [doi: 10.1016/j.annemergmed.2016.02.005] [Medline: 27125817]
36. Griffen D, Callahan CD, Markwell S, de la Cruz J, Milbrandt JC, Harvey T. Application of statistical process control to physician-specific emergency department patient satisfaction scores: a novel use of the funnel plot. Acad Emerg Med 2012;19(3):348-355 [FREE Full text] [doi: 10.1111/j.1553-2712.2012.01304.x] [Medline: 22435869]
37. Boudreaux ED, Mandry CV, Wood K. Patient satisfaction data as a quality indicator: a tale of two emergency departments. Acad Emerg Med 2003;10(3):261-268 [FREE Full text] [doi: 10.1111/j.1553-2712.2003.tb02000.x] [Medline: 12615592]
38. PROMS: background document. Canadian Institute for Health Information. 2015. URL: https://www.cihi.ca/sites/default/files/document/proms_background_may21_en-web.pdf [accessed 2018-11-27]

39. Patient experience. Canadian Institute for Health Information. 2018. URL: https://www.cihi.ca/en/patient-experience [accessed 2018-11-30]
40. Language highlight tables, 2016 Census. Statistics Canada. 2016. URL: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/lang/Table.cfm?Lang=E&T=11&Geo=00 [accessed 2021-12-10]
41. McCullagh P, Nelder JA. Generalized linear models. 2nd edition. London, UK: Chapman and Hall; 1989.
42. Jiang J. Consistent estimators in generalized linear mixed models. J Am Stat Assoc 1998;93(442):720-729. [doi: 10.2307/2670122]
43. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials 2007;28(2):182-191. [doi: 10.1016/j.cct.2006.05.007] [Medline: 16829207]
44. Martin J, Taljaard M, Girling A, Hemming K. Systematic review finds major deficiencies in sample size methodology and reporting for stepped-wedge cluster randomised trials. BMJ Open 2016;6(2):e010166 [FREE Full text] [doi: 10.1136/bmjopen-2015-010166] [Medline: 26846897]
45. Girling AJ, Hemming K. Statistical efficiency and optimal design for stepped cluster studies under linear mixed effects models. Stat Med 2016;35(13):2149-2166 [FREE Full text] [doi: 10.1002/sim.6850] [Medline: 26748662]
46. Hooper R, Teerenstra S, de Hoop E, Eldridge S. Sample size calculation for stepped wedge and other longitudinal cluster randomised trials. Stat Med 2016;35(26):4718-4728. [doi: 10.1002/sim.7028] [Medline: 27350420]
47. Kasza J, Hemming K, Hooper R, Matthews J, Forbes AB. Impact of non-uniform correlation structure on sample size and power in multiple-period cluster randomised trials. Stat Methods Med Res 2019;28(3):703-716. [doi: 10.1177/0962280217734981] [Medline: 29027505]
48. Kumar A, Chakraborty BS. Interim analysis: a rational approach of decision making in clinical trial. J Adv Pharm Technol Res 2016;7(4):118-122 [FREE Full text] [doi: 10.4103/2231-4040.191414] [Medline: 27833889]
49. Shippee ND, Domecq Garces JP, Prutsky Lopez GJ, Wang Z, Elraiyah TA, Nabhan M, et al. Patient and service user engagement in research: a systematic review and synthesized framework. Health Expect 2015;18(5):1151-1166 [FREE Full text] [doi: 10.1111/hex.12090] [Medline: 23731468]
50. Kamine TH, Rembisz A, Barron RJ, Baldwin C, Kromer M. Decrease in trauma admissions with COVID-19 pandemic. West J Emerg Med 2020;21(4):819-822 [FREE Full text] [doi: 10.5811/westjem.2020.5.47780] [Medline: 32726250]
51. COVID-19's impact on emergency departments. Canadian Institute of Health Information. 2021. URL: https://www.cihi.ca/en/covid-19-resources/impact-of-covid-19-on-canadas-health-care-systems/how-covid-19-affected [accessed 2021-12-06]
52. Covid-19 update regarding the provision of in-person and virtual care. The College of Physicians and Surgeons of Ontario. 2021. URL: https://www.cpso.on.ca/en/News/COVID-19-Updates/Information-for-Physicians/COVID-19-update-regarding-the-provision-of-in-(1) [accessed 2021-12-14]
53. COVID-19 update re: in-person and virtual care. College of Physicians and Surgeons of British Columbia. 2021. URL: https://www.cpsbc.ca/files/pdf/2021-09-03-COVID-19-Update.pdf [accessed 2021-12-14]
54. Registered nurses' union calls on parties to address growing shortage of registered nurses. Registered Nurses' Union. 2021. URL: https://rnunl.ca/registered-nurses-union-calls-on-parties-to-address-growing-shortage-of-registered-nurses/ [accessed 2021-12-14]
55. Physicians in NL stand in solidarity with our registered nurse colleagues on their day of action. Newfoundland and Labrador Medical Association (NLMA) News. URL: https://www.nlma.nl.ca/News-And-Events/NLMA-News/Page/0/Article/385 [accessed 2021-12-14]
56. Report: physician resource forecast for family medicine: Newfoundland and Labrador. Health Intelligence Inc. 2019. URL: https://nlma.nl.ca/FileManager/Documents/docs/2019.12.09_Physician_Resource_Forecast_Family_Medicine_FINAL_reduced.pdf [accessed 2021-12-14]
57. Health care in Canada, 2012: a focus on wait times. Canadian Institute for Health Information. 2012. URL: https://publications.gc.ca/collections/collection_2013/icis-cihi/H115-15-2012-eng.pdf [accessed 2021-12-14]
58. El Sayed MJ, El-Eid GR, Saliba M, Jabbour R, Hitti EA. Improving emergency department door to doctor time and process reliability: a successful implementation of lean methodology. Medicine (Baltimore) 2015;94(42):e1679 [FREE Full text] [doi: 10.1097/MD.0000000000001679] [Medline: 26496278]
59. Pang PS, McCarthy D, Schmidt M, Weiner M, Gavran G, Patterson B, et al. 440 Factors associated with perfect Press Ganey satisfaction scores for discharged emergency department patients. Ann Emerg Med 2011;58(4):S327-S328. [doi: 10.1016/j.annemergmed.2011.06.474]
60. Womack JP, Jones DT, Roos D. The machine that changed the world: the story of lean production--Toyota's secret weapon in the global car wars that is now revolutionizing world industry. New York, NY: Simon and Schuster; 1990.
61. Holweg M. The genealogy of lean production. J Oper Manag 2007;25(2):420-437. [doi: 10.1016/j.jom.2006.04.001]
62. Poole T, Mazur LM. Assessing readiness for lean change in emergency department. In: Proceedings of the 2010 Industrial Engineering Research Conference. 2010 Presented at: IIE '10; June 5-9, 2010; Cancun, Mexico p. 1-6.
63. Vashi AA, Sheikhi FH, Nashton LA, Ellman J, Rajagopal P, Asch SM. Applying lean principles to reduce wait times in a VA emergency department. Mil Med 2019;184(1-2):e169-e178. [doi: 10.1093/milmed/usy165] [Medline: 30007347]
64. Kaizen. Lean production. 2018. URL: https://www.leanproduction.com/kaizen/ [accessed 2018-11-28]

Abbreviations

CIHI: Canadian Institute for Health Information
ED: emergency department
EH: Eastern Health
GLMM: generalized linear mixed model
ICC: intracluster correlation coefficient
KPI: key performance indicator
LOS: length of stay
LWBS: leaving the emergency department without being seen
PIA: physician's initial assessment
RE-AIM: reach, effectiveness, adoption, implementation, and maintenance
RHA: regional health authority
SW-CRT: stepped wedge cluster randomized trial

Edited by G Eysenbach; submitted 16.05.21; peer-reviewed by L Dudley, C Hao; comments to author 18.11.21; revised version received 16.12.21; accepted 18.12.21; published 24.03.22.

Please cite as:
Mariathas HH, Hurley O, Anaraki NR, Young C, Patey C, Norman P, Aubrey-Bassler K, Wang PP, Gadag V, Nguyen HV, Etchegary H, McCrate F, Knight JC, Asghari S
A Quality Improvement Emergency Department Surge Management Platform (SurgeCon): Protocol for a Stepped Wedge Cluster Randomized Trial
JMIR Res Protoc 2022;11(3):e30454
URL: https://www.researchprotocols.org/2022/3/e30454
doi:10.2196/30454
PMID:35323121

Protocol

Effectiveness of a Mindfulness-Based Mobile Application for the Treatment of Depression in Ambulatory Care: Protocol for a Randomized Controlled Trial

Jan Sarlon[1], MUDR; Jessica P K Doll[1], MSc; Aline Schmassmann[1]; Serge Brand[1,2,3], PD; Naomi Ferreira[1]; Markus Muehlhauser[1], Dr med; Stefanie Urech-Meyer[1], MSc; Nina Schweinfurth[1], Dr med; Undine Emmi Lang[1], Prof Dr; Annette Beatrix Bruehl[1], Prof Dr

[1]University Psychiatric Clinics, University of Basel, Basel, Switzerland

[2]Department of Sport, Exercise and Health, Faculty of Medicine, University of Basel, Basel, Switzerland

[3]Sleep Disorders Research Center, Kermanshah University of Medical Sciences, Kermanshah, Iran

Corresponding Author:
Jan Sarlon, MUDR
University Psychiatric Clinics
University of Basel
Wilhelm Klein-Strasse 27
Basel, 4002
Switzerland
Phone: 41 613255097
Fax: 41 613255513
Email: jan.sarlon@upk.ch

Abstract

Background: Patients with major depressive disorder (MDD) often experience relapses despite regular treatment with pharmacotherapy and psychotherapy. Further, long waiting lists and more demand than treatment capacity characterize ambulatory settings. Mindfulness-based interventions proved to be effective in relapse prevention in MDD. Next, mindfulness-based interventions in the form of free mobile applications can be an effective augmentation of the treatment as usual and can fill a gap in ambulatory care.

Objective: Given this background, the aim of this randomized controlled study is to assess the effectiveness of additional MBI via a mobile app on the symptom severity and stress levels, compared to treatment as usual.

Methods: A total of 140 individuals with MDD will be randomly allocated to the intervention or control condition. The intervention consists of the daily use of the mindfulness mobile application Headspace for thirty days (up to 10 minutes a day). The control condition will be treatment as usual. At baseline and four weeks later, the following key outcome dimensions will be assessed: self-rated (Beck Depression Inventory) and experts' rated symptoms of MDD (Hamilton Depression Rating Scale); secondary outcome variables will be blood pressure, heart rate, and respiratory rate and changes in tobacco and alcohol consumption and medication as a proxy of perceived stress.

Results: This study was funded in February 2021 and approved by the institutional review board on April 15, 2021, and it started in May 2021. As of December 2021, we enrolled 30 participants. The findings are expected to be published in spring 2023.

Conclusions: We hypothesize that compared to the control conditions, individuals with MDD of the mobile app-condition will have both lower self- and experts' rated symptoms of MDD and more favorable stress-related levels. While the risk for medical events is low, the immediate benefit for participants could be a decrease in symptom severity and reduction of the stress level.

Trial Registration: Clinical Trials.gov NCT05060393; https://clinicaltrials.gov/ct2/show/NCT05060393.

International Registered Report Identifier (IRRID): DERR1-10.2196/33423

(JMIR Res Protoc 2022;11(3):e33423) doi:10.2196/33423

KEYWORDS

depression; mindfulness; mhealth; ehealth; stress level

Introduction

Major depression disorder (MDD) is a global, serious, and life-shortening disease, affecting about 300 million people worldwide and the main cause of disability-adjusted life years caused by mental diseases [1-3]. The limited capacities in ambulatory care and the increasing mental health needs request innovative, cost-effective interventions [4,5]. During the last decade, research has focused on cost-effective internet-based interventions as add-ons to existing services for individuals with MDD. Such interventions can potentially reduce the gap between the need and provision of psychiatric treatments, though being effective [6-8]. Recently, the need for remote psychological, psychiatric, and internet-based interventions in the mental health sector further increased during the worldwide COVID-19 pandemic and the health-related lockdown measures [9-11]. Accessible strategies can be an effective option to enhance mental health response capacity [12,13].

Mindfulness-based interventions (MBIs), such as mindfulness-based stress reduction or mindfulness-based cognitive therapy, have received considerable attention in modern psychotherapy because of emerging evidence regarding their efficacy in different clinical populations [14-19]. MBCT can both decrease depressive symptoms of a current depressive episode and avoid relapses in individuals with MDD [20-23]. The effects are comparable to those of other cognitive behavioral therapies [24-27] and showed favorable effects if used as an adjunct therapy to treatment as usual (TAU) [28-31]. There is growing evidence that web- and computer-based stress interventions can be effective in reducing stress, depressive, and anxiety symptoms [32-34].

In a study comparing internet-based psychological interventions in primary care, MBIs were effective in both decreasing depressive symptoms and increasing well-being compared to healthy lifestyle psychoeducational program of positive affect promotion [4] and compared to the waitlist control group [35].

As regards mobile applications, these may contribute to closing the treatment gap for depression by reaching large populations at relatively low costs [36]. There is evidence of positive effects of mindfulness-based mobile applications on well-being, stress level, affect, work engagement, irritability, mind wandering as well as sleep quality in different study populations [37-43]. Eight randomized controlled trials using a mindfulness-based app as an intervention group are reported in more detail in Table 1. The intervention period differed from 14 days up to eight weeks (median 28 days).

Table 1. Summary of randomized control trials using mindfulness-based mobile applications.

Author/year	Intervention (intervention group)	App	Control group	Study population	Outcome/psychometry	Results
O`Donnell [37], 2020	10 min a day over 30 days	Insight Timer	Wait-list control group	General adult population	Generalized Anxiety Disorder 7 (GAD-7); well being	Not available yet
Huberty [38], 2019	10 min a day over 56 days	Calm	Wait-list control group	College students	Stress, well-being, mindfulness, sleep, alcohol consumption, physical activity	Significant effects in all measured parameters in the intervention group
Möltner [39], 2017	Daily over 14 days	7Mind	Waitlist control group	Employees	Mindfulness, work engagement, job satisfaction, emotional exhaustion, emotional intelligence, innovation and creativity, and self-efficacy.	Significant effects in all measured parameter in the intervention group
Economides [40], 2018	10 min a day over 14 days	Headspace	Audiobook about mindfulness	General adult population with no history of psychiatric disorder	Stress-level, irritability, affection/Stress Overload Scale, Scale of Positive and Negative Experience (SPANE)	Significant reduction of stress level, irritability, significant changes in SPANE score in the intervention group
Howells [41] 2015	10 min a day over 14 days	Headspace	Catch Notes	Facebook and LinkedIn users	The Satisfaction with Life Scale, Flourishing scale, Positive and Negative Affect Scale, Center for Epidemiologic Studies Depression Scale (CESD)	Significant increase in positive effects in the intervention group
Bennike [42] 2020	10-20 a day over 28 days	Headspace	Cognitive training	General adult population with no history of psychiatric disorder	Mind wandering, Sustained Attention to Response Task (SART)	Sign. reduction of mind wandering in the intervention group
Lim [43] 2015	21 days	Headspace	Cognitive training	Students	Compassion	(not significant) increase of compassion in the intervention group
Ly [44] 2014	3-30 min a day over 56 days	App developed for the clinical trial	Behavioral activation via an app developed for the clinical trial	Patients with MDD (Major depression d isorder)	Symptom severity (Beck Depression Inventory II [BDI-II], Patient Health Questionnaire-9 [PHQ-9])	No significant differences between the intervention and the control group

In 4 of the 8 studies, the mindfulness-based intervention was delivered via the app Headspace, probably due to previous findings and a high score in the mobile application rating system [45], making this mobile app promising for further research. The application was brief and easy to use, free to download, and accessible via smartphones globally [41]. Users who sign up for the free trial have access to one of three guided foundation courses, titled “Basics.” There are 10 free sessions in each “Basics” course.

However, and surprisingly, from the reviewed studies, only one study targeted 40 individuals with MDD [44]; the results indicate comparable effects of a mindfulness-based app and behavioral activation via a major depression app.

Given this background and given the scarcity of app-based MBIs in individuals with MDD, the aim of this study is to examine the effects of a mindfulness-based mobile application (Headspace) on the symptom severity and physiological effects in participants with MDD in the real daily clinical practice as an adjunct therapy to the TAU.

The following hypotheses are formulated: first, following Howells et al [41], we assume that compared to a control condition, individuals in the intervention group will achieve a decrease in depressive symptoms. Second, following Huberty et al [38] and Economides [40], we expect a reduction of the stress level in the intervention group compared to the control group. For the following reasons, we hold that this study is of both practical and clinical importance. First, apps allow an individually tailored and self-paced intervention to improve symptoms of MDD and stress; second, such interventional add-ons improve individuals’ self-management and self-responsibility; third, interventional add-ons are easy to use, follow, complete, and monitor, allowing the user to get motivational-enhancing feedback; fourth, such apps provide easy-to-manage data to further improve the dialogue between the user and health care experts; fifth, once a person is accustomed to using the app, the odds are much higher than the person continues to use the app and the appropriate exercising long term.

Methods

Aims

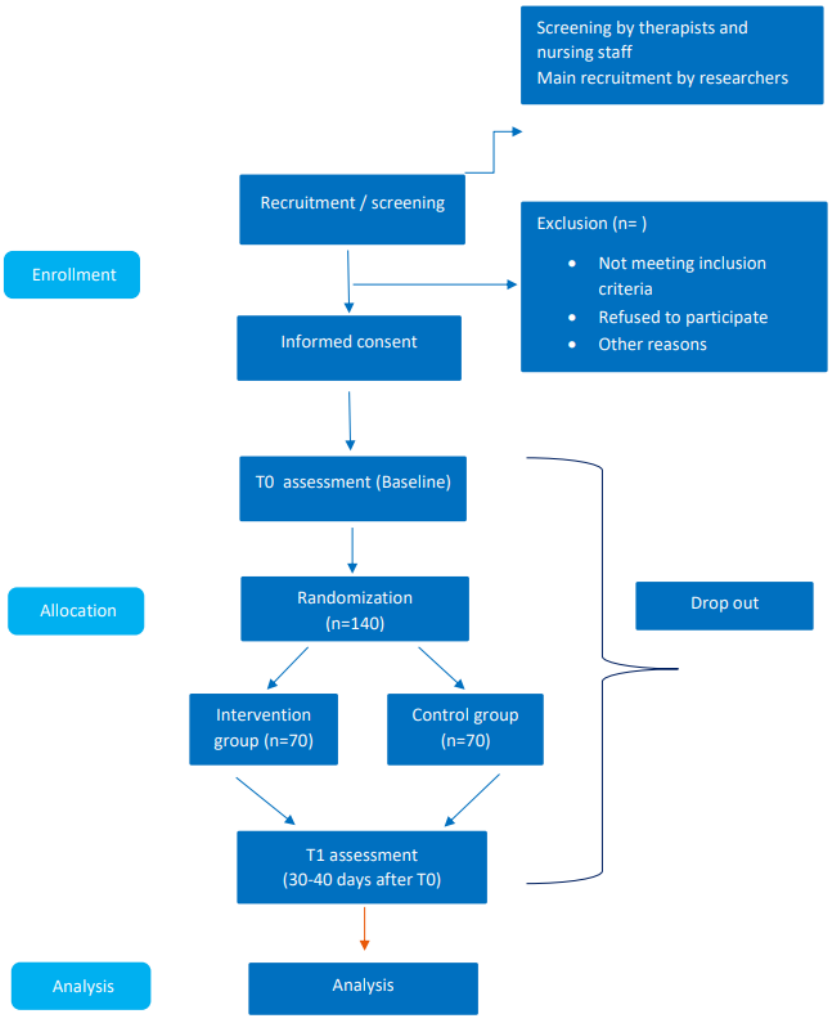
The aim of this study is to assess the efficacy of a mindfulness-based mobile application (Headspace) on the symptom severity and the stress level measured by physiological parameters in patients with MDD when compared to a control group. We assume a decrease of the symptom severity, measured by the Beck Depression Inventory II (BDI-II) and Hamilton Depression Rating Scales (HDRS), in the intervention group compared to the waitlist control group. Furthermore, we hypothesize that the stress level assessed by heart and respiratory rate as well as the blood pressure will be lower in the intervention group compared to the control group.

Study Design

This study will be a single-center, open-label and randomized wait-list-controlled trial, which will take place at the University Psychiatric Clinics (UPK) Basel (Switzerland).

Participants will be randomly allocated to two conditions: (1) MBI mobile application Headspace+ TAU (intervention group) and (2) waitlist + TAU (control group). They will be assessed at baseline and 30 days after baseline (study end). The study flowchart is presented in Figure 1 The study started in May 2021.

Figure 1. Flowchart study design.



The study protocol will be introduced to all participants, the name of the mobile application is not to be communicated in the first step. Previous experience with mobile applications for health will be assessed in all subjects prior to randomization. Randomization is performed by the project leader by using a computer-based randomization algorithm, performing the randomization allocation (ratio 1:1). Each new participant of the study will receive a participant number that corresponds to the following treatment number.

At baseline, participants will complete a series of questionnaires covering sociodemographic and treatment-related information, along with a self-rating (ie, BDI-II) of current symptoms of MDD. An expert will perform a thorough clinical-psychiatric interview and rate participants' symptoms of depression (ie, HDRS). To assess the current psychophysiological stress response at baseline conditions, a nurse will measure participants' heart and respiratory rate and blood pressure.

Next, participants will be informed about the study condition assignment. Participants of the control condition will have no additional task to their TAU instruction. Participants of the intervention group will be introduced to the Headspace app and

its handling. The necessary timeframe of the baseline assessment will be around 60 minutes.

The intervention group will be required to download the mobile application Headspace and start their daily usage in addition to their TAU (as explained in the "Intervention" section).

Regarding the control group, there will be no intervention planned. The control group will receive the same TAU as the intervention group during inpatient treatment. Participants of both groups will be supported in their outpatient treatment after discharge as usual. Compared to the treatment group, the control group will not be told about the application and its use but will be held on a waitlist. The control participants will be instructed to undergo their TAU (eg, private practice therapist), whereby they will be asked to come to the visit at measurement time T1 after 30 days (see the time points in Table 2).

At T1, both groups will undergo the same procedures as they did at baseline (maximum 60 min). Following the visit, the name of the mobile application will be communicated to the participants of the control group so they can use it after the end of the study as needed.

Table 2. Study timeline.

	Time points (days)		
		Baseline (T0)	T1
	–7 to 0	0	+30
Discharge (if inpatients)	√		
Inclusion/ exclusion criteria checked	√		
Signed written informed consent	√		
Medical history, medication	√		
Participant characteristics	√		
Randomization	√		
Visit T0 (baseline)		√	
Visit T1 (study end)			√

Participants and Recruitment

In total, 140 patients should be recruited (70 for the intervention and control group each).

Inclusion criteria: inpatients or partial inpatients (day clinic, at least 18 years old, before discharge to ambulatory care as well as outpatients, all diagnosed with MDD according to the International Classification of Diseases. All participants must have a smartphone and be generally willing to download and use a mindfulness-based app for at least 30 days.

Exclusion criteria: Acute suicidality, dementia, acute substance dependency, psychotic, schizophrenic, or schizoaffective disorders, serious health conditions like unstable cardiovascular, heart, lung, endocrine or neurological disorders, and momentary use of Headspace.

The study will be brought to the attention of potentially eligible patients of the UPK through their health care providers at UPK and through the study flyer. If they show interest in participating, the patients will receive a document with all the required information regarding the study in order to give their informed consent. After the patients sign the form of consent, it will take between 1 to 7 days until they will get scheduled for their first appointment. This will give them an adequate amount of time to consider their participation carefully. Participants will be informed about the scope and procedure of the study. Their written informed consent will be obtained.

The participants can be withdrawn from the project for any of the following reasons:

1. At their own request (withdrawal of consent)
2. In case of nonadherence to the study protocol (if participants use the app on less than 25 days or if they use other courses beyond the basic course)
3. If one or more exclusion criteria appear within the time of observation

If a participant wants to discontinue participation, this will be accepted without the need for reasons. However, the participant will be asked for potential reasons for discontinuation to document it for further analysis. The date and, if available, the reason for the withdrawal will be documented, and the participant's data will be destroyed.

Intervention

The intervention will consist of the download and the daily use of the mobile application Headspace in addition to the TAU. The participants will use the course basics of the application, which offers 10 guided meditation sessions. The duration of each session can be done in 3, 5, or 10 minutes. The study foresees that the course will be conducted in three series. The participants will make the first cycle of the course with 3 minutes per session, the second with 5 minutes, and the last one with 10 minutes per session. The participants will do one session a day, which adds up to 30 days of using the mobile application. The patients need no previous knowledge of mindfulness exercises or meditation. The course introduces and guides the patients in and through different meditation techniques. After 30 days of using the application, the patients will be examined at T1. In addition, at T1, all participants of the intervention

group will be asked for short feedback about the intervention ("comment/feedback to the app use").

To monitor the adherence to the protocol and keep participants engaged in using the app daily, participants of the intervention group will be contacted weekly via email or phone (as they prefer) and reminded of the daily use of the app.

Primary Outcomes

The BDI-II and HDRS total scores, both continuous variables, will be used as the primary outcome measure of depression severity; BDI-II and HDRS are commonly used psychometric instruments with broad applicability in research and clinical practice [46,47].

The BDI-II is a 21-item self-reporting questionnaire. It evaluates the severity of depression in normal and psychiatric populations [32]. The BDI-II total score can be obtained by adding up the answers, from a minimum of 0 to a maximum of 63 points.

The HDRS is the most widely used rating scale for depression severity. However, the 21-item version is most commonly used for the reason that the last four items (diurnal variation, depersonalization/derealization, paranoid symptoms, and obsessional and compulsive symptoms) should not contribute to the total score. This study will be using the 17-item rating scale as well. The total score can be acquired by adding up the answers. It can range from a minimum of 0 to a maximum of 52 points.

The possibility of relapse or rehospitalization between the two assessments will be assessed at T1.

Secondary Outcomes

Stress levels will be examined by autonomous nervous system parameters: resting heart rate, blood pressure as well as respiratory rate. The resting heart rate, as well as the blood pressure, will be measured by using a blood pressure monitor provided by the clinic. The respiratory rate will be assessed by counting the breaths taken per minute.

The same procedures will be used for all study participants to ensure that the data from all individuals will be comparable. All measurements will be taken in awake subjects in a sitting (slightly inclined) position and in the same room. All subjects will be asked to breathe normally and not move for five minutes to reduce the impact of pretest movements.

Furthermore, all current medications (inclusive anxiolytics and sedatives), general and momentary consumption of alcohol, as well as the consumption of nicotine, in particular pack-year and amount of daily smoked cigarettes, will be assessed.

Data Analysis

Basic hypothesis:

- H1: Decrease (or lower increase) of the BDI-II and HDRS sum scores at T1 compared to baseline in the intervention group compared to the control group.
- H0: The changes in the BDI-II and HDRS sum scores between baseline and T1 do not differ between the intervention and the control group.

For the statistical analysis, we will use tools provided by the software packages IBM SPSS statistics and R (R Core Team). The changes in the symptom severity, as well as physiological parameters between baseline and T1 measurements, will be assessed. When comparing the change scores in the intervention group with those of the control group, typical hypothesis-testing, including mixed linear models, will be used, as provided by the software packages IBM SPSS statistics and R with a significance level of .05.

Sample size estimation: The estimation is based on an expected effect size of Cohen d=0.50, which is based on prior published studies of mindfulness-based apps. When setting the significance criterion α=.05 and β=.2 (statistical power 80%), the required sample size for a two-sample t test is 64 for each group. The dropout rate is expected not to exceed 10%, resulting in 140 patients overall.

Ethics Considerations and Approval

This study will be conducted according to the international standards of the Declaration of Helsinki and subsequent amendments. This trial will be performed in compliance with the study protocol and with good clinical practice guidelines.

The study protocol has been approved by the Swiss Ethics Committee (North/Central Switzerland-EKZN; 2021-00452) and follows the guidelines of the Declaration of Helsinki and current Swiss legislation on privacy and data protection. The study participants will sign the written informed consent form before randomization. Participants will be provided with detailed information about the study and will be informed that they may leave the study at any time.

Results

This study was funded in February 2021 and approved by the institutional review board on April 15, 2021, and it started in May 2021. As of December 2021, we have enrolled 30 participants. The findings are expected to be published in spring 2023.

Discussion

This study aims to assess the efficacy of a mindfulness-based mobile application, Headspace, on depressive symptoms and physiological stress levels in outpatient MDD patients. We hypothesize fewer relapses of depressive symptoms and a decrease of symptom severity in the group of MDD patients using Headspace for 30 consecutive days, compared to the waitlist group. There is great potential in the use of application-based meditation since it is easily accessible, easy to use, cost-effective, and, most importantly, could help improve patients' well-being. However, there are some practical issues to mention in this study. One issue regards the duration of the effect; although the intervention of 30 days is within the range of other studies (median 28 days), we will not be able to measure and report middle- and long-term effects (eg, 6 months postintervention). If our hypothesis is confirmed, further studies are needed to investigate middle- and long-term effects, and implementation of MBI via an app in evidence-based care is needed. Furthermore, this study will use one single application

(Headspace) as the intervention method. Further studies will be needed to compare the effects of Headspace to other interventions. Another issue to consider is patient compliance. Due to the fact that our participants will be in an inpatient setting and the intervention consists of using the application by themselves, it will be difficult to monitor the actual use of the application and, therefore, patient compliance.

Fortunately, there are no risks of medical events as well as no known risks of the intervention planned. The only potential and very low risk could be due to unauthorized data access; however, this risk will be minimized due to the adherence to the study protocol. Further, electronic data will be saved on password-protected computers that are only accessible by authorized personnel. All paper data will be stored on-campus in locked cabinets. Only authorized personnel will have access to the study data.

The potential immediate benefit for the participants consists of a decrease in symptom severity and prevention of depressive relapse without side effects. The long-term benefits are that patient care, especially in the outpatient setting or after discharge from inpatient treatment, will be easily augmented and improved. Moreover, the patients will be supported by an easily accessible mindfulness-based intervention. Such benefits may accommodate the issue of limited capacities in ambulatory care. With the prevention of depressive relapse, people with MDD get the chance to not only improve in symptom severity but also to live in their homes, with their families, go to school, or to work, accompanied by a tool for stress reduction.

Acknowledgments

This research is financed by a grant from the Funding Fond (Forschungsförderungsfond) of the University Psychiatric Clinics (2021-00452). The funding body did not play any role in the design of the study, nor will it play any role in data collection, analysis, interpretation of data, or the writing of manuscripts.

Authors' Contributions

All the authors participated in the design and planning of the intervention evaluated here. JS designed the project and wrote the manuscript. SB planned the necessary statistical analysis. JPKD, AS, SB, and ABB contributed to the ethics approval and writing of the manuscript. ABB and UEL coordinated the process. JS, AS, and NF will collect the data. JPKD, MM, SUM, and NS will recruit or assist with recruitment. JS, AS, and JR will analyze the data.

Conflicts of Interest

None declared.

References

1. Herrman H, Kieling C, McGorry P, Horton R, Sargent J, Patel V. Reducing the global burden of depression: a Lancet–World Psychiatric Association Commission. The Lancet 2019 Jun;393(10189):e42-e43. [doi: 10.1016/s0140-6736(18)32408-5]
2. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet 2013 Nov 9;382(9904):1575-1586. [doi: 10.1016/S0140-6736(13)61611-6] [Medline: 23993280]
3. Saxena S, Thornicroft G, Knapp M, Whiteford H. Resources for mental health: scarcity, inequity, and inefficiency. The Lancet 2007 Sep;370(9590):878-889. [doi: 10.1016/s0140-6736(07)61239-2]
4. Gili M, Castro A, García-Palacios A, Garcia-Campayo J, Mayoral-Cleries F, Botella C, et al. Efficacy of Three Low-Intensity, Internet-Based Psychological Interventions for the Treatment of Depression in Primary Care: Randomized Controlled Trial. J Med Internet Res 2020 Jun 05;22(6):e15845 [FREE Full text] [doi: 10.2196/15845] [Medline: 32501276]
5. Breslow AS, Tran NM, Lu FQ, Alpert JE, Cook BL. Depression Treatment Expenditures for Adults in the USA: a Systematic Review. Curr Psychiatry Rep 2019 Sep 21;21(10):1-11. [doi: 10.1007/s11920-019-1083-3]
6. Karyotaki E, Ebert DD, Donkin L, Riper H, Twisk J, Burger S, et al. Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. Clin Psychol Rev 2018 Jul;63:80-92. [doi: 10.1016/j.cpr.2018.06.007] [Medline: 29940401]
7. Van't Hof E, Cuijpers P, Stein DJ. Self-help and Internet-guided interventions in depression and anxiety disorders: a systematic review of meta-analyses. CNS Spectr 2009 Feb 07;14(2 Suppl 3):34-40 [FREE Full text] [doi: 10.1017/s1092852900027279] [Medline: 19238128]
8. Hallgren M, Kraepelien M, Öjehagen A, Lindefors N, Zeebari Z, Kaldo V, et al. Physical exercise and internet-based cognitive-behavioural therapy in the treatment of depression: randomised controlled trial. Br J Psychiatry 2015 Sep 02;207(3):227-234. [doi: 10.1192/bjp.bp.114.160101] [Medline: 26089305]
9. Stefana A, Youngstrom EA, Hopwood CJ, Dakanalis A. The COVID-19 pandemic brings a second wave of social isolation and disrupted services. Eur Arch Psychiatry Clin Neurosci 2020 May 15;270(6):785-786. [doi: 10.1007/s00406-020-01137-8]
10. Humer E, Stippl P, Pieh C, Schimböck W, Probst T. Psychotherapy via the Internet: What Programs Do Psychotherapists Use, How Well-Informed Do They Feel, and What Are Their Wishes for Continuous Education? Int J Environ Res Public Health 2020 Nov 05;17(21):8182 [FREE Full text] [doi: 10.3390/ijerph17218182] [Medline: 33167478]

11. Shayganfard M, Mahdavi F, Haghighi M, Sadeghi Bahmani D, Brand S. Health Anxiety Predicts Postponing or Cancelling Routine Medical Health Care Appointments among Women in Perinatal Stage during the Covid-19 Lockdown. IJERPH 2020 Nov 09;17(21):8272. [doi: 10.3390/ijerph17218272]
12. Yue J, Yan W, Sun Y, Yuan K, Su S, Han Y, et al. Mental health services for infectious disease outbreaks including COVID-19: a rapid systematic review. Psychol. Med 2020 Nov 05;50(15):2498-2513. [doi: 10.1017/s0033291720003888]
13. Augusterfer EF, O'Neal CR, Martin SW, Sheikh TL, Mollica RF. The Role of Telemental Health, Tele-consultation, and Tele-supervision in Post-disaster and Low-resource Settings. Curr Psychiatry Rep 2020 Nov 28;22(12):85 [FREE Full text] [doi: 10.1007/s11920-020-01209-5] [Medline: 33247315]
14. Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. J Consult Clin Psychol 2010 Apr;78(2):169-183 [FREE Full text] [doi: 10.1037/a0018555] [Medline: 20350028]
15. Ninomiya A, Sado M, Park S, Fujisawa D, Kosugi T, Nakagawa A, et al. Effectiveness of mindfulness‐based cognitive therapy in patients with anxiety disorders in secondary‐care settings: A randomized controlled trial. Psychiatry Clin. Neurosci 2019 Dec 20;74(2):132-139. [doi: 10.1111/pcn.12960]
16. Aalsma MC, Jones LD, Staples JK, Garabrant JM, Gordon JS, Cyr LR, et al. Mind-Body Skills Groups for Adolescents With Depression in Primary Care: A Pilot Study. Journal of Pediatric Health Care 2020 Sep;34(5):462-469. [doi: 10.1016/j.pedhc.2020.05.003]
17. Chesin MS, Benjamin-Phillips CA, Keilp J, Fertuck EA, Brodsky BS, Stanley B. Improvements in Executive Attention, Rumination, Cognitive Reactivity, and Mindfulness Among High–Suicide Risk Patients Participating in Adjunct Mindfulness-Based Cognitive Therapy: Preliminary Findings. The Journal of Alternative and Complementary Medicine 2016 Aug;22(8):642-649. [doi: 10.1089/acm.2015.0351]
18. Key BL, Rowa K, Bieling P, McCabe R, Pawluk EJ. Mindfulness-based cognitive therapy as an augmentation treatment for obsessive-compulsive disorder. Clin Psychol Psychother 2017 Feb 13;24(5):1109-1120. [doi: 10.1002/cpp.2076]
19. Segal ZV, Walsh KM. Mindfulness-based cognitive therapy for residual depressive symptoms and relapse prophylaxis. Current Opinion in Psychiatry 2016;29(1):7-12. [doi: 10.1097/yco.0000000000000216]
20. Kuyken W, Warren FC, Taylor RS, Whalley B, Crane C, Bondolfi G, et al. Efficacy of Mindfulness-Based Cognitive Therapy in Prevention of Depressive Relapse: An Individual Patient Data Meta-analysis From Randomized Trials. JAMA Psychiatry 2016 Jun 01;73(6):565-574 [FREE Full text] [doi: 10.1001/jamapsychiatry.2016.0076] [Medline: 27119968]
21. Cladder-Micus MB, Speckens AE, Vrijsen JN, T Donders AR, Becker ES, Spijker J. Mindfulness-based cognitive therapy for patients with chronic, treatment-resistant depression: A pragmatic randomized controlled trial. Depress Anxiety 2018 Oct 08;35(10):914-924 [FREE Full text] [doi: 10.1002/da.22788] [Medline: 30088834]
22. Geurts DE, Compen FR, Van Beek MH, Speckens AE. The effectiveness of mindfulness-based cognitive therapy for major depressive disorder: evidence from routine outcome monitoring data. BJPsych open 2020 Nov 25;6(6):1-7. [doi: 10.1192/bjo.2020.118]
23. Tickell A, Ball S, Bernard P, Kuyken W, Marx R, Pack S, et al. The Effectiveness of Mindfulness-Based Cognitive Therapy (MBCT) in Real-World Healthcare Services. Mindfulness 2019 Jan 12;11(2):279-290. [doi: 10.1007/s12671-018-1087-9]
24. Strauss C, Arbon A, Barkham M, Byford S, Crane R, de Visser R, et al. Low-Intensity Guided Help Through Mindfulness (LIGHTMIND): study protocol for a randomised controlled trial comparing supported mindfulness-based cognitive therapy self-help to supported cognitive behavioural therapy self-help for adults experiencing depression. Trials 2020 May 04;21(1):1-10. [doi: 10.1186/s13063-020-04322-1]
25. Chan SHW, Chan WWK, Chao JYW, Chan PKL. A randomized controlled trial on the comparative effectiveness of mindfulness-based cognitive therapy and health qigong-based cognitive therapy among Chinese people with depression and anxiety disorders. BMC Psychiatry 2020 Dec 14;20(1):1-14. [doi: 10.1186/s12888-020-02994-2]
26. Farb N, Anderson A, Ravindran A, Hawley L, Irving J, Mancuso E, et al. Prevention of relapse/recurrence in major depressive disorder with either mindfulness-based cognitive therapy or cognitive therapy. Journal of Consulting and Clinical Psychology 2018 Feb;86(2):200-204. [doi: 10.1037/ccp0000266]
27. Abolghasemi A, Gholami H, Narimani M, Gamji M. The Effect of Beck's Cognitive Therapy and Mindfulness-Based Cognitive Therapy on Sociotropic and Autonomous Personality Styles in Patients With Depression. Iran J Psychiatry Behav Sci 2015 Dec 23;9(4):1-6. [doi: 10.17795/ijpbs-3665]
28. Barnhofer T, Duggan D, Crane C, Hepburn S, Fennell MJV, Williams JMG. Effects of meditation on frontal alpha-asymmetry in previously suicidal individuals. Neuroreport 2007 May 7;18(7):709-712. [doi: 10.1097/wnr.0b013e3280d943cd]
29. Schanche E, Vøllestad J, Visted E, Svendsen JL, Osnes B, Binder PE, et al. The effects of mindfulness-based cognitive therapy on risk and protective factors of depressive relapse - a randomized wait-list controlled trial. BMC Psychol 2020 Jun 05;8(1):57 [FREE Full text] [doi: 10.1186/s40359-020-00417-1] [Medline: 32503649]
30. Ter Avest MJ, Dusseldorp E, Huijbers MJ, van Aalderen JR, Cladder-Micus MB, Spinhoven P, et al. Added value of Mindfulness-Based Cognitive Therapy for Depression: A Tree-based Qualitative Interaction Analysis. Behav Res Ther 2019 Nov;122:103467. [doi: 10.1016/j.brat.2019.103467] [Medline: 31557693]

31. Godfrin K, van Heeringen C. The effects of mindfulness-based cognitive therapy on recurrence of depressive episodes, mental health and quality of life: A randomized controlled study. Behaviour Research and Therapy 2010 Aug;48(8):738-746. [doi: 10.1016/j.brat.2010.04.006]
32. Heber E, Ebert DD, Lehr D, Cuijpers P, Berking M, Nobis S, et al. The Benefit of Web- and Computer-Based Interventions for Stress: A Systematic Review and Meta-Analysis. J Med Internet Res 2017 Feb 17;19(2):e32. [doi: 10.2196/jmir.5774]
33. Lippmann M, Laudel H, Heinzle M, Narciss S. Relating Instructional Design Components to the Effectiveness of Internet-Based Mindfulness Interventions: A Critical Interpretive Synthesis. J Med Internet Res 2019 Nov 27;21(11):e12497 [FREE Full text] [doi: 10.2196/12497]
34. Lyzwinski LN, Caffery L, Bambling M, Edirippulige S. A Systematic Review of Electronic Mindfulness-Based Therapeutic Interventions for Weight, Weight-Related Behaviors, and Psychological Stress. Telemedicine and e-Health 2018 Mar;24(3):173-184. [doi: 10.1089/tmj.2017.0117]
35. Wahbeh H. Internet Mindfulness Meditation Intervention (IMMI) Improves Depression Symptoms in Older Adults. Medicines 2018 Nov 02;5(4):119. [doi: 10.3390/medicines5040119]
36. Kerst A, Zielasek J, Gaebel W. Smartphone applications for depression: a systematic literature review and a survey of health care professionals' attitudes towards their use in clinical practice. Eur Arch Psychiatry Clin Neurosci 2019 Jan 03:139-152. [doi: 10.1007/s00406-018-0974-3] [Medline: 30607530]
37. O'Donnell KT, Dunbar M, Speelman DL. Effectiveness of using a meditation app in reducing anxiety and improving well-being during the COVID-19 pandemic: A structured summary of a study protocol for a randomized controlled trial. Trials 2020 Dec 09;21(1):1006 [FREE Full text] [doi: 10.1186/s13063-020-04935-6] [Medline: 33298117]
38. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the Mindfulness Meditation Mobile App “Calm” to Reduce Stress Among College Students: Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273. [doi: 10.2196/14273]
39. Möltner H, Leve J, Esch T. Burnout-Prävention und mobile Achtsamkeit: Evaluation eines appbasierten Gesundheitstrainings bei Berufstätigen. Gesundheitswesen 2017 Oct 10;57(03):295-300. [doi: 10.1055/s-0043-114004]
40. Economides M, Martman J, Bell MJ, Sanderson B. Improvements in Stress, Affect, and Irritability Following Brief Use of a Mindfulness-based Smartphone App: A Randomized Controlled Trial. Mindfulness 2018 Mar 1;9(5):1584-1593. [doi: 10.1007/s12671-018-0905-4]
41. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the ‘app’ in Happiness: A Randomised Controlled Trial of a Smartphone-Based Mindfulness Intervention to Enhance Wellbeing. J Happiness Stud 2014 Oct 29:163-185. [doi: 10.1007/s10902-014-9589-1]
42. Bennike IH, Wieghorst A, Kirk U. Online-based Mindfulness Training Reduces Behavioral Markers of Mind Wandering. J Cogn Enhanc 2017 Apr 25;1(2):172-181. [doi: 10.1007/s41465-017-0020-9]
43. Lim D, Condon P, DeSteno D. Mindfulness and Compassion: An Examination of Mechanism and Scalability. PLoS ONE 2015 Feb 17;10(2):e0118221. [doi: 10.1371/journal.pone.0118221]
44. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014 Jan 09;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
45. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and Evaluation of Mindfulness-Based iPhone Apps. JMIR mHealth uHealth 2015 Aug 19;3(3):e82. [doi: 10.2196/mhealth.4328]
46. Wang Y, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. Braz J Psychiatry 2013;35(4):416-431 [FREE Full text] [doi: 10.1590/1516-4446-2012-1048] [Medline: 24402217]
47. Williams JBW. Standardizing the Hamilton Depression Rating Scale: past, present, and future. Eur Arch Psychiatry Clin Nuerosci 2001 Jun;251(S2):6-12. [doi: 10.1007/bf03035120]

Abbreviations

BDI-II: Beck depression inventory II
HDRS: Hamilton Depression Rating Scale
MBI: mindfulness-based intervention
MDD: major depression disorder
TAU: treatment as usual
UPK: University Psychiatric Clinics

Edited by G Eysenbach; submitted 12.10.21; peer-reviewed by AR Mohseni, R Eckhoff; comments to author 30.12.21; revised version received 13.01.22; accepted 23.01.22; published 31.03.22.

Please cite as:
Sarlon J, Doll JPK, Schmassmann A, Brand S, Ferreira N, Muehlhauser M, Urech-Meyer S, Schweinfurth N, Lang UE, Bruehl AB Effectiveness of a Mindfulness-Based Mobile Application for the Treatment of Depression in Ambulatory Care: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2022;11(3):e33423
URL: https://www.researchprotocols.org/2022/3/e33423
doi:10.2196/33423
PMID:35357325

Protocol

Treatment of Persistent Postconcussion Syndrome With Repetitive Transcranial Magnetic Stimulation Using Functional Near-Infrared Spectroscopy as a Biomarker of Response: Protocol for a Randomized Controlled Clinical Trial

Sané du Plessis[1,2], BSc; Ibukunoluwa K Oni[1,2,3], MEng; Andrew P Lapointe[2,3], PhD; Christina Campbell[1], MSc; Jeff F Dunn[1,2,3], PhD; Chantel T Debert[1,2], MD, MSc

[1]Department of Clinical Neurosciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[2]Hotchkiss Brain Institute, Calgary, AB, Canada

[3]Department of Radiology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Corresponding Author:
Sané du Plessis, BSc
Department of Clinical Neurosciences
Cumming School of Medicine
University of Calgary
3330 Hospital Dr NW
Calgary, AB, T2N4N1
Canada
Phone: 1 9448643
Email: sane.duplessis@ucalgary.ca

Abstract

Background: Approximately one-third of all concussions lead to persistent postconcussion syndrome (PPCS). Repetitive transcranial magnetic stimulation (rTMS) is a form of noninvasive brain stimulation that has been extensively used to treat refractory major depressive disorder and has a strong potential to be used as a treatment for patients with PPCS. Functional near-infrared spectroscopy (fNIRS) has already been used as a tool to assess patients with PPCS and may provide insight into the pathophysiology of rTMS treatment in patients with PPCS.

Objective: The primary objective of this research is to determine whether rTMS treatment improves symptom burden in patients with PPCS compared to sham treatment using the Rivermead postconcussion symptom questionnaire. The secondary objective is to explore the neuropathophysiological changes that occur following rTMS in participants with PPCS using fNIRS. Exploratory objectives include determining whether rTMS treatment in participants with PPCS will also improve quality of life, anxiety, depressive symptoms, cognition, posttraumatic stress, and function secondary to headaches.

Methods: A total of 44 adults (18-65 years old) with PPCS (>3 months to 5 years) will participate in a double-blind, sham-controlled, concealed allocation, randomized clinical trial. The participants will engage in either a 4-week rTMS treatment protocol or sham rTMS protocol (20 treatments). The left dorsolateral prefrontal cortex will be located through Montreal Neurologic Institute coordinates. The intensity of the rTMS treatment over the left dorsolateral prefrontal cortex will be 120% of resting motor threshold, with a frequency of 10 Hz, 10 trains of 60 pulses per train (total of 600 pulses), and intertrain interval of 45 seconds. Prior to starting the rTMS treatment, participant and injury characteristics, questionnaires (symptom burden, quality of life, depression, anxiety, cognition, and headache), and fNIRS assessment will be collected. Repeat questionnaires and fNIRS will occur immediately after rTMS treatment and at 1 month and 3 months post rTMS. Outcome parameters will be analyzed by a 2-way (treatment × time) mixed analysis of variance.

Results: As of May 6, 2021, 5 participants have been recruited for the study, and 3 have completed the rTMS protocol. The estimated completion date of the trial is May 2022.

Conclusions: This trial will expand our knowledge of how rTMS can be used as a treatment option of PPCS and will explore the neuropathophysiological response of rTMS through fNIRS analysis.

Trial Registration: ClinicalTrials.gov NCT04568369; https://clinicaltrials.gov/ct2/show/NCT04568369

International Registered Report Identifier (IRRID): DERR1-10.2196/31308

(JMIR Res Protoc 2022;11(3):e31308) doi:10.2196/31308

KEYWORDS

concussion; mild traumatic brain injury; persistent postconcussion syndrome; repetitive transcranial magnetic stimulation; functional near-infrared spectroscopy; traumatic brain injury; TBI; brain injury; brain; symptom burden; mental health; quality of life; neuroscience; neurology

Introduction

Concussion incidence has been rising among Canadians with the Ontario Physician Billing codes reporting an annual incidence of 1.2% of the population [1]. According to the Canadian Community Health Survey, over the last 2 decades, the number of concussions or other brain injury reports among Canadians is increasing, with 1 in 200 Canadians reporting a concussion or other brain injury to be their most disabling injury [2]. While many recover quickly from concussions, approximately one-third of patients with concussion have prolonged symptoms [3]. The symptoms that these individuals experience are collectively known as persistent postconcussion syndrome (PPCS), and can include cognitive, physical, and emotional impairments [4].

Transcranial magnetic stimulation (TMS) is a tool used to provide noninvasive electromagnetic induction through magnetic fields to the brain in conscious humans [5]. In general, single pulse TMS is used to explore brain function while repetitive TMS (rTMS) is used to change brain activity that lasts well beyond stimulation. rTMS has been shown to be an effective treatment for depressive disorders [6] as well as several other neurological conditions [7,8]. Furthermore, a recent systematic review showed that rTMS may be an effective treatment for patients experiencing PPCS, but studies with larger sample sizes are required to further support this evidence [9].

There is potential in using rTMS to improve symptoms for patients with PPCS, but the mechanism behind rTMS treatment for PPCS remains unknown. Recent studies have suggested that rTMS treatment modifies recovery pace and improves symptomatology [10,11]. Currently, very few studies have looked at biomarkers for rTMS treatment response in patients with PPCS. Our group previously conducted a 2-patient case study demonstrating how functional near-infrared spectroscopy (fNIRS) may be useful as a sensitive tool to predict treatment response to rTMS in patients with PPCS [12]. fNIRS is a noninvasive imaging technique that measures the differences in the absorption of local oxygenated and deoxygenated hemoglobin in the brain. The changes in cerebral tissue oxygenation are then used to monitor brain activity [13,14]. fNIRS has demonstrated comparability to functional magnetic resonance imaging for reliably detecting changes in cerebral vascular reactivity [15], and several studies have already examined the use of fNIRS for assessing concussions [15-17] and PPCS in adults [18]. Previous work assessing pediatric and adult patients with PPCS have shown reduced connectivity in both frontal and motor regions when compared to healthy controls during rest as well as during a finger tapping exercise and working memory task [18,19]. These findings are comparable to literature on functional magnetic resonance imaging, which has shown a reduction in connectivity in patients with mild traumatic brain injury during both resting state and task-based measures [20-22]. fNIRS has many advantages to being used in the clinical setting because of its portability and ability to replace other neuroimaging techniques that may be cost prohibitive.

The primary objective for this study is to determine if the application of rTMS treatment to the left dorsolateral prefrontal cortex (DLPFC) in patients with PPCS will significantly improve postconcussion symptoms using the Rivermead postconcussion symptom questionnaire (RPQ) compared to sham treatment. The second objective of the study is to explore the pathophysiology of rTMS response in participants with PPCS using fNIRS immediately following rTMS treatment and at 1-month and 3-month posttreatment follow-ups. An exploratory aim of the study is to determine whether rTMS treatment to the left DLPFC will improve quality of life, anxiety, depressive symptoms, fatigue, posttraumatic stress, and function secondary to headaches compared to sham treatment. We hypothesize that participants who show a clinically significant improvement in PPCS symptoms will also show an improved hemodynamic response in fNIRS recordings.

Methods

Study Design

This is a double-blind, sham-controlled, concealed allocation, randomized clinical trial. Recruitment will occur through 3 locations in Calgary, Alberta, Canada including the Calgary Brain Injury Program, the University of Calgary Sports Medicine Centre concussion clinic, and the Calgary Pain Program. The estimated study duration is from March 2021 to May 2022. Participant enrollment, fNIRS, and rTMS treatment will be completed by authors SD and CC.

Ethics Approval and Trial Registration

This study was approved by the University of Calgary Conjoint Health Research Ethics Board (REB15-1786). All participants will sign informed consent prior to study participation. This study has been registered with ClinicalTrials (NCT04568369).

Participants

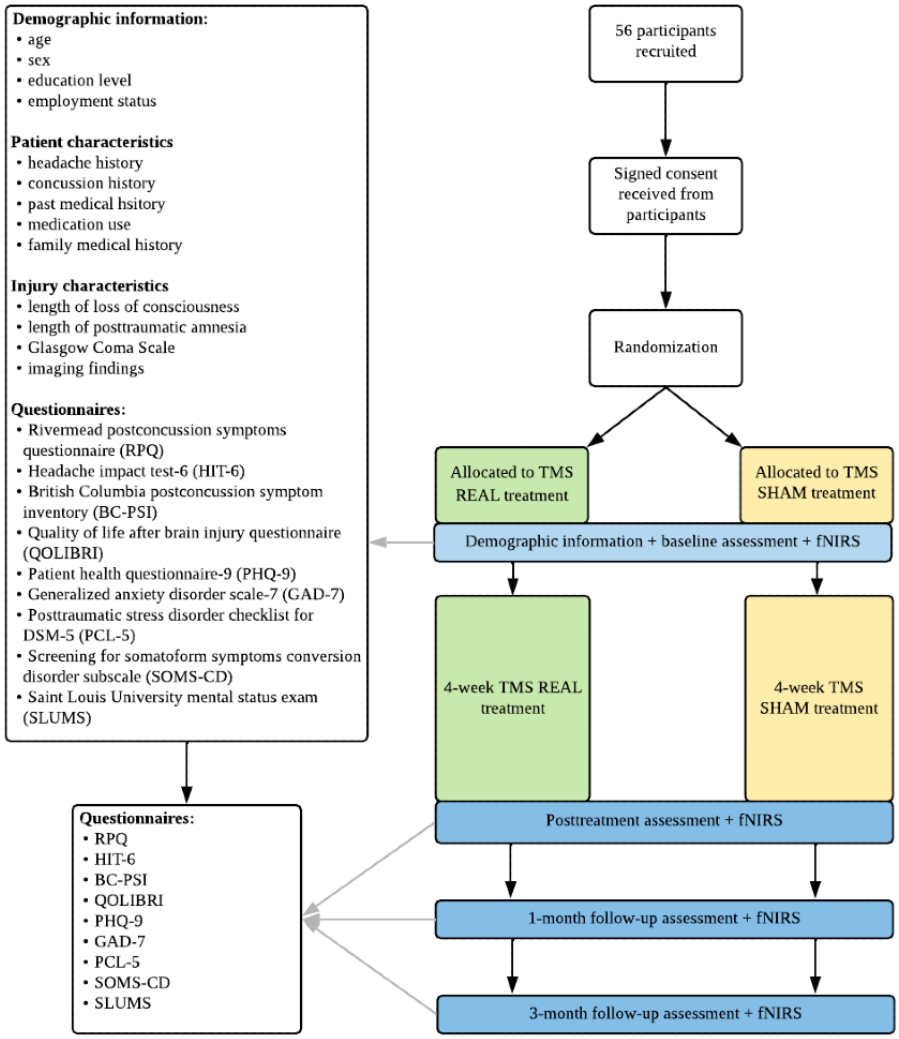
All participants will be between the ages of 18 and 65 years. The participants will meet the inclusion and exclusion criteria explained in this section. Inclusion criteria include a diagnosis of PPCS based on the International Classification of Diseases (ICD)-10 criteria for at least 3 months to a maximum of 5 years [23]. The ICD-10 states that PPCS must be followed by head trauma with loss of consciousness as well as any 3 of the

following symptoms: feelings of unwellness (ie, headaches, dizziness, general malaise, and excessive fatigue or noise intolerance), emotional changes (ie, irritability, emotional lability, depression, and anxiety), difficulty concentrating or performing mental tasks or memory complaints, insomnia, reduced tolerance to alcohol, preoccupation with the above symptoms, and fear of permanent brain damage [23]. Current pharmacologic management will be maintained without change during the treatment study (ie, consistent use of triptans, opioids, tricyclic antidepressants, and antiseizure medications). Participants undergoing Botox treatment will undergo rTMS treatment 6-8 weeks following their injection, which is around the time of peak Botox efficacy [24]. Exclusion criteria comprise a history of prior TMS therapy, TMS-related contraindications (pacemaker, metallic implants, and large intracranial lesion), change in medication, and other neurological or mental health medical conditions such as structural brain disease, previous seizure, psychiatric disorders excluding depression and anxiety (eg, schizophrenia and bipolar disorder), liver or kidney disease, malignancy, uncontrolled hypertension or diabetes, and pregnancy.

Study Procedure

The study procedure is outlined in Figure 1. Prior to starting the study, participant and injury characteristics and questionnaires will be completed. Demographic information will be collected including age, sex, education level, and employment status. The participant characteristics collected will include headache history, concussion history, past medical history, medication use, and family medical history. Finally, injury characteristics will include length of loss of consciousness, length of posttraumatic amnesia, initial Glasgow Coma Scale, and imaging findings. The participants will be randomized to either active or sham rTMS. Once randomized, the participants will complete questionnaires and fNIRS. They will then complete 4 weeks (20 sessions) of treatment. Questionnaires and fNIRS will be repeated immediately following treatment and then at 1 month and 3 months post treatment.

Figure 1. Study design protocol. DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; fNIRS: functional near-infrared spectroscopy; TMS: transcranial magnetic stimulation.



Blinding and Randomization

The participants will be randomized to either active or sham rTMS treatment by a research assistant external to the study using sequentially numbered opaque concealed envelopes. All individuals involved in the study except the research assistant administering the rTMS will be blinded to the treatment protocol, and allocation will be concealed. Following study completion, the participants and study personnel will be unblinded. Once recruitment is completed, if rTMS treatment proves beneficial after a full statistical analysis, the participants in the sham groups will be offered active rTMS treatment.

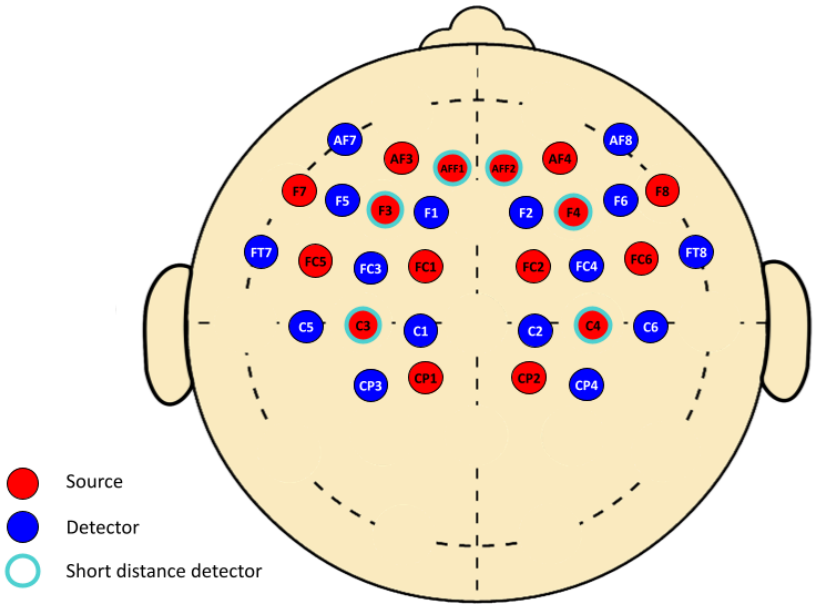
TMS Protocol

The participants will engage in a 4-week treatment protocol (20 treatments). This was chosen as it reflects prior depression, PPCS, and migraine protocol durations [25-27]. A standardized atlas brain with Montreal Neurologic Institute coordinates will be used for navigation. rTMS treatment will be to the left DLPFC and will be located through Montreal Neurologic Institute coordinates (-50, 30, 36) [28]. The resting motor threshold will be determined using electromyography electrodes attached to the right abductor digiti minimi muscle and a TMS stimulation coil placed over the left motor cortex, as previously described by Stilling et al [11]. The intensity of the rTMS will be 100%-120% of resting motor threshold amplitude, with a frequency of 10 Hz, 10 trains of 60 pulses per train (total of 600 pulses) and intertrain interval of 45 seconds. In the sham condition, a sham coil will be applied to the scalp after the resting motor threshold is determined. Recent studies exploring rTMS implementing a 100%-120% resting motor threshold amplitude as a treatment for PPCS have been successful for patient retainment and treatment success [25,27,29,30]. Our group has previously used an intensity of 70% resting motor threshold in a similar patient population [11]. This protocol will use 120% resting motor threshold to attempt to achieve a greater neurophysiological effect [31]. Furthermore, previous studies treating posttraumatic headache and migraine show benefit using a frequency of 10 Hz, 600 pulses, and intertrain interval of 45 seconds, suggesting this protocol would also be beneficial in PPCS [32-35]. The sham coil will produce sounds and vibration similar to the real rTMS coil but will not produce any effective stimulation. This blinding method has been demonstrated to be effective in previous sham TMS studies [36].

Functional Near-Infrared Spectroscopy

fNIRS will be completed before treatment, within 1 week following the 4-week treatment, and then at 1 month and 3 months post treatment. The TechEn (TechEn, Inc) system and a headcap with transmit and receive fibers positioned over the parietal and frontal lobes will be used (Figure 2). Data will be collected for 5 minutes at rest and then during a finger tapping exercise and working memory tasks. The resting task will consist of the participant sitting and resting while fixating on a small white cross for 5 minutes [37,38]. Finger tapping will be used as a simple, repeatable motor activation task, which has prominently shown a motor response [18,19]. The finger tapping exercise will consist of the participant tapping their right forefinger to their thumb at a frequency of 1 Hz, while being prompted on screen by “Tap” or a white cross indicating rest. The participant will alternate between tapping for 10 seconds and resting for 15 seconds, which will be repeated 10 times. Finally, the working memory task will consist of a 2-back variation of the n-back exercise [39], where a series of letters will appear on screen in 2-second intervals, and the participant will be asked to match each new letter to the letter presented 2 trials back. If the letters match, the participant is asked to press the right arrow key indicating correct, or, if the letters do not match, they are asked to press the left arrow key indicating incorrect. The participant will alternate between 20 presentations of letters followed by 20 seconds of rest, which will be repeated 3 times. Data on caffeine and other substance consumption from the last 8 hours before fNIRS assessment will also be collected. This protocol was adapted from the protocol described by Hocke et al [18].

Figure 2. Optode configuration of functional near-infrared spectroscopy headcap based on 10-10 and 10-5 electroencephalogram coordinate system. AF: anterior frontal; AFF: anterior prefrontal; C: central; CP: central parietal; F: frontal; FC: fronto-central.



Rivermead Postconcussion Symptom Questionnaire

The RPQ is the primary outcome measure for this study. The RPQ assesses the severity of 16 commonly experienced PPCS symptoms. The participants are instructed to rate the extent to which they have suffered from each of the listed symptoms in the past 24 hours, as compared to preinjury levels, using a scale of 0 (“not experienced at all”) to 4 (“a severe problem”). It is advised to use this assessment as 2 separate scales (RPQ-13 and RPQ-3). Using these subscales, the instrument has good test-retest reliability and external construct validity [40]. This questionnaire probes the separate cognitive, emotional, and somatic components of PPCS.

Exploratory Outcome Measures

Multiple other questionnaires will be administered as exploratory outcomes. The questionnaires administered will be described in the following section.

Quality of Life After Brain Injury Questionnaire

Quality of life will be measured by the quality of life after brain injury (QOLIBRI) questionnaire assessment. The QOLIBRI questionnaire is a health-related quality of life instrument developed specifically for participants who have experienced a traumatic brain injury. It has 6 subscales with a total of 37 items and is scored from 0 to 100 (with 100 being the best possible quality of life). This tool has been validated in the traumatic brain injury literature and is shown to have adequate (>0.5) to excellent (>0.7) test-retest reliability [41].

Headache Impact Test

The headache impact test-6 (HIT-6) is a 6-item tool used as a global measure of headache impact. The HIT-6 addresses the 6 categories of headache impact including social functioning, role functioning, cognitive functioning, vitality, psychological distress, and severity of headache pain. Each question is scored on a 5-point scale. The total score can range from 36 to 78, with higher total score indicating greater impact [42]. The HIT-6 has been validated in participants with episodic and chronic migraine [43]. Based on a chronic tension type headache study, a minimal clinically important change has been reported as a decrease of at least 8 points on the HIT-6 [44].

Patient Health Questionnaire-9

The patient health questionnaire-9 (PHQ-9) is a 9-item tool used to assess the presence and severity of depressive symptoms. In the PHQ-9, a score below 4 is classified as minimal, 5 to 9 is mild, 10 to 14 is moderate, 15 to 19 is moderate-to-severe, and anything higher than 20 is severe [45]. The PHQ-9 has been validated in participants with mild to severe traumatic brain injury with a screening cut-off score of greater than 12 for major depressive disorder [46]. The test-retest reliability is excellent within 7 days of the initial assessment. The minimal clinically important difference for the PHQ-9 is 5 points [47].

Generalized Anxiety Disorder Scale

The generalized anxiety disorder (GAD-7) scale is a 7-item tool where, like the PHQ-9, each item is rated on frequency over a 2-week period based on a 0 to 3 scale [48]. An analysis confirmed screens for depression (PHQ-9) and anxiety (GAD-7) to be distinct, and it is suggested that the 2 highly comorbid disorders be assessed independently [49].

Saint Louis University Mental Status Exam

The Saint Louis University mental status exam will be used to screen the participants for mild cognitive impairment. The Saint Louis University mental status exam is an 11-item scale with questions that correspond to attention, recall (immediate and delayed with interference), orientation, number calculation, memory, registration, visual spatial information, and executive function [50].

Screening for Somatoform Symptoms Conversion Disorder Subscale

Somatoform symptoms will be monitored through the screening for somatoform symptoms conversion disorder subscale. This screening lists 14 symptoms associated with somatization disorder. Each symptom is rated on degree of impairment experienced in the past 7 days on a 0 to 4 scale, with 0 indicating “not at all” and 4 indicating “very severe” [51].

Posttraumatic Stress Disorder Checklist for DSM-5

The posttraumatic stress disorder (PTSD) checklist for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; PCL-5) is a screening tool used to assess PTSD. The PLC-5 is a 20-item self-report measure with each question scaled from 0 to 4, with the total score ranging from 0 to 80. A score greater than 53 may indicate more severe PTSD symptoms [52].

Analysis

Preprocessing fNIRS Data

All preprocessing data will be done using MATLAB (The MathWorks Inc). Raw intensity data will first be converted to optical density using Homer3 (hmrR_Intensity2OD). Optical density will then be assessed for signal quality over several steps. First, channels that display low signal (below 90 dB) or have an oversaturated signal (above 140 dB) will be discarded. Channels that have a signal to noise ratio less than 1 will also be discarded, as previously described by Hocke et al [18]. Next, a spline motion correction step will be applied to correct for motion artifacts and baseline shifts using Homer3 (hmrR_MotionCorrectSpline). Short distance detectors located approximately 8 mm from select sources will be used to remove the influence of extracerebral signal obtained from superficial layers [53,54]. A bandpass filter will be applied to the data with a low pass filter of 0.01 Hz and a high pass filter of 2 Hz (hmrR_BandpassFilt). Optical density data will then be converted to delta concentration of oxyhemoglobin and deoxyhemoglobin using the modified Beer-Lambert law with age-dependent differential path length factors. Finally, the concentration data will be down sampled to 5 Hz.

Postprocessing fNIRS Data

Preprocessed data between 4 major brain regions (ie, left DLPFC, left motor cortex, right DLPFC, and right motor cortex) will be assessed using a coherence analysis at 0.04 to 0.1 Hz. Coherence will be calculated for all channels over the 4 regions. Coherence is defined as an estimate of a linear-time variant

relationship between 2 signals at a particular frequency domain [19]. The coherence value used will be the maximum coherence between each combination of regions. The interregional coherence values will be quantified between the right and left DLPFC, the left and right motor cortex, the left DLPC and left motor cortex, and finally between the right DLPFC and right motor cortex [18].

Statistical Analysis

The effect size of the RPQ score extrapolated from a similar study (d=0.77) [10] provided the basis for a power calculation performed with an alpha of .05, and 80% power suggested that 44 participants should be included in this study (22 per group). Anticipating a 20% attrition [10,22,55], 56 participants will be recruited for this study.

Descriptive statistics will be used to analyze baseline characteristics. A 2-way (treatment × time) mixed ANOVA will be used to assess whether any changes in clinical measurements and outcome parameters were the result of the interaction of type of treatment (real vs sham) and time. fNIRS data will also be analyzed using a 2-way mixed ANOVA where interregional coherence will act as the dependent variable, sham vs real TMS treatment will act as the between-subject variable, and brain region and assessment (ie, pre-TMS, post-TMS, 1-month follow-up, and 3-month follow-up) will act as the within subject variables. A P value of ≤.05 will be considered significant. A post hoc analysis will be used for exploratory measures. Simple effects testing using a Bonferroni correction will be performed when a significant group by time interaction is detected.

fNIRS data will be assessed to determine the location and magnitude of task-evoked changes in oxyhemoglobin, deoxyhemoglobin, and total hemoglobin at the left DLPC for each participant. Coherence of hemodynamic frequencies between major brain regions will also be quantified to observe brain connectivity during tasks as previous works have done [18,19]. Our group has previously attempted to decrease bias by restricting change in medication and other treatments throughout the trial [11]; however, it is unknown how some medications may influence rTMS and fNIRS data. Full patient characteristics, injury characteristics, and medication profile will be collected to better understand this bias.

Results

Participant recruitment for the study started on March 25, 2021. To date, 5 participants are enrolled, and 3 have completed the rTMS protocol as of May 6, 2021. The study protocol will be completed in May 2022. The results are expected in June 2022 with submitted manuscript in August 2022.

Discussion

The proposed project will substantially contribute to the growing body of literature exploring rTMS as a safe and noninvasive treatment for participants with PPCS. The mechanism by which rTMS influences brain function and subsequently alters participants' symptoms and function is unknown. fNIRS analysis will provide insight into the neuropathophysiological changes that can occur following rTMS treatment in participants with PPCS. Several studies have explored fNIRS and TMS in the clinical setting outside of PPCS [56] including depression [57], panic disorder [58-60], tinnitus [61], phobias [62], bulimia nervosa [63], and stroke [64-66]. Our group has already shown through a 2-patient case study that fNIRS might be useful as a sensitive tool to predict response to rTMS in participants with postconcussion syndrome. This study will help grow the understanding of rTMS as a treatment for PPCS and provide additional support for the use of fNIRS as a tool to understand the effects of neuromodulation on participants with PPCS.

Acknowledgments

We would like to thank the Hotchkiss Brain Institute for their continual support through the Noninvasive Neuromodulation Network and the leadership of Adam Kirton and Alexander McGirr. In addition, we would like to thank Cari Jahraus for helping with the study protocol and ethics. Finally, we would also like to thank the members of the Canadian Traumatic Brain Consortium and the Integrated Concussion Research Program for their valuable feedback on this protocol.

Conflicts of Interest

None declared.

References

1. Langer L, Levy C, Bayley M. Increasing Incidence of Concussion: True Epidemic or Better Recognition? J Head Trauma Rehabil 2020;35(1):E60-E66. [doi: 10.1097/HTR.0000000000000503] [Medline: 31246881]
2. Gordon KE, Kuhle S. 'Reported concussion' time trends within two national health surveys over two decades. Brain Inj 2018;32(7):843-849. [doi: 10.1080/02699052.2018.1463105] [Medline: 29652525]
3. McMahon P, Hricik A, Yue JK, Puccio AM, Inoue T, Lingsma HF, TRACK-TBI Investigators. Symptomatology and functional outcome in mild traumatic brain injury: results from the prospective TRACK-TBI study. J Neurotrauma 2014 Jan 01;31(1):26-33 [FREE Full text] [doi: 10.1089/neu.2013.2984] [Medline: 23952719]
4. Ryan LM, Warden DL. Post concussion syndrome. Int Rev Psychiatry 2003 Nov;15(4):310-316. [doi: 10.1080/09540260310001606692] [Medline: 15276952]

5. Klomjai W, Katz R, Lackmy-Vallée A. Basic principles of transcranial magnetic stimulation (TMS) and repetitive TMS (rTMS). Ann Phys Rehabil Med 2015 Sep;58(4):208-213 [FREE Full text] [doi: 10.1016/j.rehab.2015.05.005] [Medline: 26319963]
6. Brunoni AR, Chaimani A, Moffa AH, Razza LB, Gattaz WF, Daskalakis ZJ, et al. Repetitive Transcranial Magnetic Stimulation for the Acute Treatment of Major Depressive Episodes: A Systematic Review With Network Meta-analysis. JAMA Psychiatry 2017 Feb 01;74(2):143-152. [doi: 10.1001/jamapsychiatry.2016.3644] [Medline: 28030740]
7. Yang W, Liu T, Song X, Zhang Y, Li Z, Cui Z, et al. Comparison of different stimulation parameters of repetitive transcranial magnetic stimulation for unilateral spatial neglect in stroke patients. J Neurol Sci 2015 Dec 15;359(1-2):219-225. [doi: 10.1016/j.jns.2015.08.1541] [Medline: 26671118]
8. Vonloh M, Chen R, Kluger B. Safety of transcranial magnetic stimulation in Parkinson's disease: a review of the literature. Parkinsonism Relat Disord 2013 Jun;19(6):573-585 [FREE Full text] [doi: 10.1016/j.parkreldis.2013.01.007] [Medline: 23473718]
9. Mollica A, Safavifar F, Fralick M, Giacobbe P, Lipsman N, Burke MJ. Transcranial Magnetic Stimulation for the Treatment of Concussion: A Systematic Review. Neuromodulation 2021 Jul;24(5):803-812. [doi: 10.1111/ner.13319] [Medline: 33184973]
10. Moussavi Z, Suleiman A, Rutherford G, Ranjbar Pouya O, Dastgheib Z, Zhang W, et al. A Pilot Randomised Double-Blind Study of the Tolerability and efficacy of repetitive Transcranial Magnetic Stimulation on Persistent Post-Concussion Syndrome. Sci Rep 2019 Apr 02;9(1):5498 [FREE Full text] [doi: 10.1038/s41598-019-41923-6] [Medline: 30940870]
11. Stilling J, Paxman E, Mercier L, Gan LS, Wang M, Amoozegar F, et al. Treatment of Persistent Post-Traumatic Headache and Post-Concussion Symptoms Using Repetitive Transcranial Magnetic Stimulation: A Pilot, Double-Blind, Randomized Controlled Trial. J Neurotrauma 2020 Jan 15;37(2):312-323. [doi: 10.1089/neu.2019.6692] [Medline: 31530227]
12. Stilling JM, Duszynski CC, Oni I, Paxman E, Dunn JF, Debert CT. Using Functional Near-Infrared Spectroscopy to Study the Effect of Repetitive Transcranial Magnetic Stimulation in Concussion: A Two-Patient Case Study. Front Neurol 2019;10:476 [FREE Full text] [doi: 10.3389/fneur.2019.00476] [Medline: 31139136]
13. Almajidy RK, Mankodiya K, Abtahi M, Hofmann UG. A Newcomer's Guide to Functional Near Infrared Spectroscopy Experiments. IEEE Rev Biomed Eng 2020;13:292-308. [doi: 10.1109/RBME.2019.2944351] [Medline: 31634142]
14. Scholkmann F, Kleiser S, Metz AJ, Zimmermann R, Mata Pavia J, Wolf U, et al. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. Neuroimage 2014 Jan 15;85 Pt 1:6-27. [doi: 10.1016/j.neuroimage.2013.05.004] [Medline: 23684868]
15. Amyot F, Kenney K, Spessert E, Moore C, Haber M, Silverman E, et al. Assessment of cerebrovascular dysfunction after traumatic brain injury with fMRI and fNIRS. Neuroimage Clin 2020;25:102086 [FREE Full text] [doi: 10.1016/j.nicl.2019.102086] [Medline: 31790877]
16. Forcione M, Colonnese C, Belli A. Cerebral Hemodynamic Influences in Task-Related Functional Magnetic Resonance Imaging and Near-Infrared Spectroscopy in Acute Sport-Related Concussion: A Review. J. Imaging 2018 Apr 17;4(4):59. [doi: 10.3390/jimaging4040059]
17. Bishop SA, Neary JP. Assessing prefrontal cortex oxygenation after sport concussion with near-infrared spectroscopy. Clin Physiol Funct Imaging 2018 Jul;38(4):573-585. [doi: 10.1111/cpf.12447] [Medline: 28626873]
18. Hocke LM, Duszynski CC, Debert CT, Dleikan D, Dunn JF. Reduced Functional Connectivity in Adults with Persistent Post-Concussion Symptoms: A Functional Near-Infrared Spectroscopy Study. J Neurotrauma 2018 Jun 01;35(11):1224-1232 [FREE Full text] [doi: 10.1089/neu.2017.5365] [Medline: 29373947]
19. Urban KJ, Barlow KM, Jimenez JJ, Goodyear BG, Dunn JF. Functional near-infrared spectroscopy reveals reduced interhemispheric cortical communication after pediatric concussion. J Neurotrauma 2015 Jun 01;32(11):833-840 [FREE Full text] [doi: 10.1089/neu.2014.3577] [Medline: 25387354]
20. Mayer AR, Mannell MV, Ling J, Gasparovic C, Yeo RA. Functional connectivity in mild traumatic brain injury. Hum Brain Mapp 2011 Nov;32(11):1825-1835 [FREE Full text] [doi: 10.1002/hbm.21151] [Medline: 21259381]
21. Johnson B, Zhang K, Gay M, Horovitz S, Hallett M, Sebastianelli W, et al. Alteration of brain default network in subacute phase of injury in concussed individuals: resting-state fMRI study. Neuroimage 2012 Jan 02;59(1):511-518 [FREE Full text] [doi: 10.1016/j.neuroimage.2011.07.081] [Medline: 21846504]
22. Zhou Y, Lui YW, Zuo X, Milham MP, Reaume J, Grossman RI, et al. Characterization of thalamo-cortical association using amplitude and connectivity of functional MRI in mild traumatic brain injury. J Magn Reson Imaging 2014 Jun;39(6):1558-1568 [FREE Full text] [doi: 10.1002/jmri.24310] [Medline: 24014176]
23. The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research. Geneva, Switzerland: World Health Organization; 1993.
24. Dong Y, Wu T, Hu X, Wang T. Efficacy and safety of botulinum toxin type A for upper limb spasticity after stroke or traumatic brain injury: a systematic review with meta-analysis and trial sequential analysis. Eur J Phys Rehabil Med 2017 May;53(2):256-267 [FREE Full text] [doi: 10.23736/S1973-9087.16.04329-X] [Medline: 27834471]
25. Rao V, Bechtold K, McCann U, Roy D, Peters M, Vaishnavi S, et al. Low-Frequency Right Repetitive Transcranial Magnetic Stimulation for the Treatment of Depression After Traumatic Brain Injury: A Randomized Sham-Controlled Pilot Study. J Neuropsychiatry Clin Neurosci 2019;31(4):306-318. [doi: 10.1176/appi.neuropsych.17110338] [Medline: 31018810]

26. Koski L, Kolivakis T, Yu C, Chen J, Delaney S, Ptito A. Noninvasive brain stimulation for persistent postconcussion symptoms in mild traumatic brain injury. J Neurotrauma 2015 Jan 01;32(1):38-44. [doi: 10.1089/neu.2014.3449] [Medline: 24955920]
27. Hoy KE, McQueen S, Elliot D, Herring SE, Maller JJ, Fitzgerald PB. A Pilot Investigation of Repetitive Transcranial Magnetic Stimulation for Post-Traumatic Brain Injury Depression: Safety, Tolerability, and Efficacy. J Neurotrauma 2019 Jul 01;36(13):2092-2098. [doi: 10.1089/neu.2018.6097] [Medline: 30712461]
28. Rusjan PM, Barr MS, Farzan F, Arenovich T, Maller JJ, Fitzgerald PB, et al. Optimal transcranial magnetic stimulation coil placement for targeting the dorsolateral prefrontal cortex using novel magnetic resonance image-guided neuronavigation. Hum Brain Mapp 2010 Nov;31(11):1643-1652 [FREE Full text] [doi: 10.1002/hbm.20964] [Medline: 20162598]
29. Siddiqi SH, Trapp NT, Hacker CD, Laumann TO, Kandala S, Hong X, et al. Repetitive Transcranial Magnetic Stimulation with Resting-State Network Targeting for Treatment-Resistant Depression in Traumatic Brain Injury: A Randomized, Controlled, Double-Blinded Pilot Study. J Neurotrauma 2019 Apr 15;36(8):1361-1374 [FREE Full text] [doi: 10.1089/neu.2018.5889] [Medline: 30381997]
30. Lee SA, Kim M. Effect of Low Frequency Repetitive Transcranial Magnetic Stimulation on Depression and Cognition of Patients with Traumatic Brain Injury: A Randomized Controlled Trial. Med Sci Monit 2018 Dec 04;24:8789-8794 [FREE Full text] [doi: 10.12659/MSM.911385] [Medline: 30513530]
31. Padberg F, Zwanzger P, Keck ME, Kathmann N, Mikhaiel P, Ella R, et al. Repetitive transcranial magnetic stimulation (rTMS) in major depression: relation between efficacy and stimulation intensity. Neuropsychopharmacology 2002 Oct;27(4):638-645. [doi: 10.1016/S0893-133X(02)00338-X] [Medline: 12377400]
32. Misra UK, Kalita J, Bhoi SK. High frequency repetitive transcranial magnetic stimulation (rTMS) is effective in migraine prophylaxis: an open labeled study. Neurol Res 2012 Jul;34(6):547-551. [doi: 10.1179/1743132812Y.0000000045] [Medline: 22732675]
33. Misra UK, Kalita J, Bhoi SK. High-rate repetitive transcranial magnetic stimulation in migraine prophylaxis: a randomized, placebo-controlled study. J Neurol 2013 Nov;260(11):2793-2801. [doi: 10.1007/s00415-013-7072-2] [Medline: 23963471]
34. Kalita J, Laskar S, Bhoi SK, Misra UK. Efficacy of single versus three sessions of high rate repetitive transcranial magnetic stimulation in chronic migraine and tension-type headache. J Neurol 2016 Nov;263(11):2238-2246. [doi: 10.1007/s00415-016-8257-2] [Medline: 27541044]
35. Kalita J, Bhoi SK, Misra UK. Effect of high rate rTMS on somatosensory evoked potential in migraine. Cephalalgia 2017 Nov;37(13):1222-1230. [doi: 10.1177/0333102416675619] [Medline: 27780897]
36. Broadbent HJ, van den Eynde F, Guillaume S, Hanif EL, Stahl D, David AS, et al. Blinding success of rTMS applied to the dorsolateral prefrontal cortex in randomised sham-controlled trials: a systematic review. World J Biol Psychiatry 2011 Jun;12(4):240-248. [doi: 10.3109/15622975.2010.541281] [Medline: 21426265]
37. Wang J, Dong Q, Niu H. The minimum resting-state fNIRS imaging duration for accurate and stable mapping of brain connectivity network in children. Sci Rep 2017 Jul 25;7(1):6461 [FREE Full text] [doi: 10.1038/s41598-017-06340-7] [Medline: 28743886]
38. Geng S, Liu X, Biswal BB, Niu H. Effect of Resting-State fNIRS Scanning Duration on Functional Brain Connectivity and Graph Theory Metrics of Brain Network. Front Neurosci 2017;11:392 [FREE Full text] [doi: 10.3389/fnins.2017.00392] [Medline: 28775676]
39. Owen AM, McMillan KM, Laird AR, Bullmore E. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum Brain Mapp 2005 May;25(1):46-59 [FREE Full text] [doi: 10.1002/hbm.20131] [Medline: 15846822]
40. Eyres S, Carey A, Gilworth G, Neumann V, Tennant A. Construct validity and reliability of the Rivermead Post-Concussion Symptoms Questionnaire. Clin Rehabil 2005 Dec;19(8):878-887. [doi: 10.1191/0269215505cr905oa] [Medline: 16323387]
41. von Steinbüchel N, Wilson L, Gibbons H, Hawthorne G, Höfer S, Schmidt S, QOLIBRI Task Force. Quality of Life after Brain Injury (QOLIBRI): scale development and metric properties. J Neurotrauma 2010 Jul;27(7):1167-1185. [doi: 10.1089/neu.2009.1076] [Medline: 20486801]
42. Kosinski M, Bayliss MS, Bjorner JB, Ware JE, Garber WH, Batenhorst A, et al. A six-item short-form survey for measuring headache impact: the HIT-6. Qual Life Res 2003 Dec;12(8):963-974. [doi: 10.1023/a:1026119331193] [Medline: 14651415]
43. Yang M, Rendas-Baum R, Varon SF, Kosinski M. Validation of the Headache Impact Test (HIT-6™) across episodic and chronic migraine. Cephalalgia 2011 Feb;31(3):357-367 [FREE Full text] [doi: 10.1177/0333102410379890] [Medline: 20819842]
44. Castien RF, Blankenstein AH, van der Windt DAWM, Dekker J. Minimal clinically important change on the Headache Impact Test-6 questionnaire in patients with chronic tension-type headache. Cephalalgia 2012 Jul;32(9):710-714. [doi: 10.1177/0333102412449933] [Medline: 22707519]
45. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]

46. Fann JR, Bombardier CH, Dikmen S, Esselman P, Warms CA, Pelzer E, et al. Validity of the Patient Health Questionnaire-9 in assessing depression following traumatic brain injury. J Head Trauma Rehabil 2005;20(6):501-511. [doi: 10.1097/00001199-200511000-00003] [Medline: 16304487]
47. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. Med Care 2004 Dec;42(12):1194-1201. [doi: 10.1097/00005650-200412000-00006] [Medline: 15550799]
48. Spitzer RL, Kroenke K, Williams JBW, 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]
49. Teymoori A, Gorbunova A, Haghish FE, Real R, Zeldovich M, Wu Y, Center-Tbi Investigators And Participants, et al. Factorial Structure and Validity of Depression (PHQ-9) and Anxiety (GAD-7) Scales after Traumatic Brain Injury. J Clin Med 2020 Mar 23;9(3):873 [FREE Full text] [doi: 10.3390/jcm9030873] [Medline: 32210017]
50. Tariq SH, Tumosa N, Chibnall JT, Perry MH, Morley JE. Comparison of the Saint Louis University mental status examination and the mini-mental state examination for detecting dementia and mild neurocognitive disorder--a pilot study. Am J Geriatr Psychiatry 2006 Nov;14(11):900-910. [doi: 10.1097/01.JGP.0000221510.33817.86] [Medline: 17068312]
51. Rief W, Hiller W. A new approach to the assessment of the treatment effects of somatoform disorders. Psychosomatics 2003;44(6):492-498. [doi: 10.1176/appi.psy.44.6.492] [Medline: 14597684]
52. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and Initial Psychometric Evaluation. J Trauma Stress 2015 Dec;28(6):489-498. [doi: 10.1002/jts.22059] [Medline: 26606250]
53. Wyser D, Mattille M, Wolf M, Lambercy O, Scholkmann F, Gassert R. Short-channel regression in functional near-infrared spectroscopy is more effective when considering heterogeneous scalp hemodynamics. Neurophotonics 2020 Jul;7(3):035011 [FREE Full text] [doi: 10.1117/1.NPh.7.3.035011] [Medline: 33029548]
54. Gagnon L, Cooper RJ, Yücel MA, Perdue KL, Greve DN, Boas DA. Short separation channel location impacts the performance of short channel regression in NIRS. Neuroimage 2012 Mar 01;59(3):2518-2528 [FREE Full text] [doi: 10.1016/j.neuroimage.2011.08.095] [Medline: 21945793]
55. Leung A, Shukla S, Fallah A, Song D, Lin L, Golshan S, et al. Repetitive Transcranial Magnetic Stimulation in Managing Mild Traumatic Brain Injury-Related Headaches. Neuromodulation 2016 Feb;19(2):133-141. [doi: 10.1111/ner.12364] [Medline: 26555886]
56. Curtin A, Tong S, Sun J, Wang J, Onaral B, Ayaz H. A Systematic Review of Integrated Functional Near-Infrared Spectroscopy (fNIRS) and Transcranial Magnetic Stimulation (TMS) Studies. Front Neurosci 2019;13:84 [FREE Full text] [doi: 10.3389/fnins.2019.00084] [Medline: 30872985]
57. Eschweiler GW, Wegerer C, Schlotter W, Spandl C, Stevens A, Bartels M, et al. Left prefrontal activation predicts therapeutic effects of repetitive transcranial magnetic stimulation (rTMS) in major depression. Psychiatry Res 2000 Oct 30;99(3):161-172. [doi: 10.1016/s0925-4927(00)00062-7] [Medline: 11068197]
58. Dresler T, Ehlis A, Plichta MM, Richter MM, Jabs B, Lesch K, et al. Panic disorder and a possible treatment approach by means of high-frequency rTMS: a case report. World J Biol Psychiatry 2009;10(4 Pt 3):991-997. [doi: 10.1080/15622970902898147] [Medline: 19396702]
59. Deppermann S, Vennewald N, Diemer J, Sickinger S, Haeussinger FB, Notzon S, et al. Does rTMS alter neurocognitive functioning in patients with panic disorder/agoraphobia? An fNIRS-based investigation of prefrontal activation during a cognitive task and its modulation via sham-controlled rTMS. Biomed Res Int 2014;2014:542526 [FREE Full text] [doi: 10.1155/2014/542526] [Medline: 24757668]
60. Deppermann S, Vennewald N, Diemer J, Sickinger S, Haeussinger FB, Dresler T, et al. Neurobiological and clinical effects of fNIRS-controlled rTMS in patients with panic disorder/agoraphobia during cognitive-behavioural therapy. Neuroimage Clin 2017;16:668-677 [FREE Full text] [doi: 10.1016/j.nicl.2017.09.013] [Medline: 29085773]
61. Schecklmann M, Giani A, Tupak S, Langguth B, Raab V, Polak T, et al. Functional near-infrared spectroscopy to probe state- and trait-like conditions in chronic tinnitus: a proof-of-principle study. Neural Plast 2014;2014:894203 [FREE Full text] [doi: 10.1155/2014/894203] [Medline: 25478237]
62. Deppermann S, Notzon S, Kroczek A, Rosenbaum D, Haeussinger FB, Diemer J, et al. Functional co-activation within the prefrontal cortex supports the maintenance of behavioural performance in fear-relevant situations before an iTBS modulated virtual reality challenge in participants with spider phobia. Behav Brain Res 2016 Jul 01;307:208-217. [doi: 10.1016/j.bbr.2016.03.028] [Medline: 26996315]
63. Sutoh C, Koga Y, Kimura H, Kanahara N, Numata N, Hirano Y, et al. Repetitive Transcranial Magnetic Stimulation Changes Cerebral Oxygenation on the Left Dorsolateral Prefrontal Cortex in Bulimia Nervosa: A Near-Infrared Spectroscopy Pilot Study. Eur Eat Disord Rev 2016 Jan;24(1):83-88. [doi: 10.1002/erv.2413] [Medline: 26481583]
64. Hara T, Abo M, Kakita K, Mori Y, Yoshida M, Sasaki N. The Effect of Selective Transcranial Magnetic Stimulation with Functional Near-Infrared Spectroscopy and Intensive Speech Therapy on Individuals with Post-Stroke Aphasia. Eur Neurol 2017;77(3-4):186-194. [doi: 10.1159/000457901] [Medline: 28161706]
65. Urushidani N, Kinoshita S, Okamoto T, Tamashiro H, Abo M. Low-Frequency rTMS and Intensive Occupational Therapy Improve Upper Limb Motor Function and Cortical Reorganization Assessed by Functional Near-Infrared Spectroscopy in

a Subacute Stroke Patient. Case Rep Neurol 2018;10(2):223-231 [FREE Full text] [doi: 10.1159/000492381] [Medline: 30283320]

66. Tamashiro H, Kinoshita S, Okamoto T, Urushidani N, Abo M. Effect of baseline brain activity on response to low-frequency rTMS/intensive occupational therapy in poststroke patients with upper limb hemiparesis: a near-infrared spectroscopy study. Int J Neurosci 2019 Apr;129(4):337-343. [doi: 10.1080/00207454.2018.1536053] [Medline: 30311827]

Abbreviations

DLPFC: dorsolateral prefrontal cortex
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
fNIRS: functional near-infrared spectroscopy
GAD-7: generalized anxiety disorder scale
HIT: headache impact test
ICD: International Classification of Diseases
PCL-5: posttraumatic stress disorder checklist for DSM-5
PHQ-9: patient health questionnaire
PPCS: persistent postconcussion symptom
PTSD: posttraumatic stress disorder
QOLIBRI: quality of life after brain injury questionnaire
RPQ: Rivermead postconcussion symptom questionnaire
rTMS: repetitive transcranial magnetic stimulation
TMS: transcranial magnetic stimulation

Edited by T Leung; submitted 18.06.21; peer-reviewed by Z Moussavi, R Dewey; comments to author 04.08.21; revised version received 29.09.21; accepted 25.01.22; published 22.03.22.

Please cite as:
du Plessis S, Oni IK, Lapointe AP, Campbell C, Dunn JF, Debert CT
Treatment of Persistent Postconcussion Syndrome With Repetitive Transcranial Magnetic Stimulation Using Functional Near-Infrared Spectroscopy as a Biomarker of Response: Protocol for a Randomized Controlled Clinical Trial
JMIR Res Protoc 2022;11(3):e31308
URL: https://www.researchprotocols.org/2022/3/e31308
doi:10.2196/31308
PMID:35315783

Protocol

mHealth Intervention to Improve Treatment Outcomes Among People With HIV Who Use Cocaine: Protocol for a Pilot Randomized Controlled Trial

Yerina S Ranjit[1], PhD; Archana Krishnan[2], PhD; Debarchana Ghosh[3], PhD; Claire Cravero[4], MPH; Xin Zhou[5], PhD; Frederick L Altice[5], MD

[1]Department of Communication, University of Missouri, Columbia, MO, United States
[2]Department of Communication, University at Albany, State University of New York, New York, NY, United States
[3]Department of Geography, University of Connecticut, Storrs, CT, United States
[4]Harvard T H Chan School of Public Health, Harvard University, Boston, MA, United States
[5]Department of Internal Medicine, Section of Infectious Diseases, Yale University, New Haven, CT, United States

Corresponding Author:
Yerina S Ranjit, PhD
Department of Communication
University of Missouri
108 Switzler Hall
Columbia, MO, 65211
United States
Phone: 1 573 822 0881
Email: ranjity@missouri.edu

Related Article:

This is a corrected version. See correction statement: https://www.researchprotocols.org/2022/4/e37925

Abstract

Background: Antiretroviral therapy is effective in reducing HIV-related morbidity, mortality, and transmission among people with HIV. However, adherence and persistence to antiretroviral therapy are crucial for successful HIV treatment outcomes. People with HIV who use cocaine have poor access to HIV services and lower retention in care.

Objective: The primary goal of this paper is to provide a detailed description of a mobile health intervention. This study is designed to improve medication adherence among people with HIV who use cocaine. A secondary goal is to list the important challenges and adaptations incorporated in the study design.

Methods: This study, titled Project SMART, used a wireless technology–based intervention, including cellular-enabled electronic pillboxes called TowerView Health and smartphones, to provide reminders and feedback on adherence behavior. The intervention design was based on the theoretical frameworks provided by the self-determination theory and the Motivation Technology Model. The 12-week pilot randomized controlled trial with four arms provided three types of feedback: automated feedback, automated+clinician feedback, and automated feedback+social network feedback.

Results: The study was funded by the National Institute of Drug Abuse (R21DA039842) on August 1, 2016. The institutional review board for the study was approved by Yale University on March 21, 2017. Data collection lasted from June 2017 to January 2020. The final enrollment was 71 participants, of whom 57 (80%) completed the study. The data are currently undergoing analysis, and the manuscript is being developed for publication in early 2022.

Conclusions: Implementing complex mobile health interventions for high-risk and marginalized populations with multicomponent interventions poses certain challenges, such as finding companies with adequate technology for clients and financial stability and minimizing the research-related burden for the study population. Conducting feasibility studies is important to recognize these challenges and the opportunity to address these challenges with solutions while keeping the design of a randomized controlled trial as true as possible.

Trial Registration: Clinicaltrials.gov NCT04418076; https://clinicaltrials.gov/ct2/show/NCT04418076

International Registered Report Identifier (IRRID): DERR1-10.2196/28332

(JMIR Res Protoc 2022;11(3):e28332) doi:10.2196/28332

KEYWORDS

smart pillbox; smartphone; mHealth intervention; people with HIV; cocaine use; antiretroviral therapy; description of feasibility and acceptability; mobile phone

Introduction

Background

Antiretroviral therapy (ART), a daily medication regimen, is extremely effective in reducing HIV-related morbidity and mortality and reducing transmission among people with HIV [1-3]. ART suppresses HIV viral load (VL) and slows the progression of the virus in the body by keeping a high CD4 cell count [4]. The Centers for Disease Control and Prevention defines viral suppression as having <200 copies of HIV per milliliter of blood [4]. However, successful HIV treatment outcomes depend on patients' adherence to and persistence in ART [5-9]. When adherence is optimal (≥90%) [10], improvements are seen not only for HIV outcomes but also for several comorbidities, including substance use disorders, tuberculosis, viral hepatitis, and depression [11-15]. Among the approximately 1.2 million people with HIV in the United States, nonadherence to ART results in as few as 19% achieving viral suppression [16]. For people with HIV with co-occurring substance use disorders, there are additional negative health outcomes. Active drug use is associated with clinicians not prescribing or delaying prescribing ART [17], which can overshadow a patient's health care priorities, including medication adherence and retention in HIV care [18]. As a result, people with HIV who use drugs experience worse clinical outcomes and greater mortality than people who do not use drugs [19]. People who use cocaine especially have poor measures of success in accessing and retaining HIV care [20,21]. They are less likely to receive ART compared with people who use other drugs [21,22], who have poor adherence to ART when they do receive it [23-25], and who have an extraordinarily low likelihood of achieving viral suppression [26]. Overall, AIDS-related death rates among people who use cocaine are 3 times higher than among those who do not use cocaine, controlling for age, race, duration of illness, and self-reported high ART adherence [26]. Effective strategies for improving ART adherence are urgently needed to address this health crisis among people with HIV who use cocaine.

For people with HIV who use cocaine, currently, the only effective adherence intervention is directly administered ART, which is expensive and labor intensive [27]. International guidelines suggest that adherence interventions need to be scaled back, both in terms of cost and personnel [28]. Mobile health (mHealth) technologies satisfy these recommendations and can provide innovative, efficacious, and cost-effective strategies for improving ART adherence and optimizing HIV treatment outcomes [29,30]. In fact, the World Health Organization guidelines now strongly recommend sending SMS text messages to promote adherence and retention in general [31]. mHealth has already shown promise in resource-limited settings [32] and improvement in treatment efficacy among patients with diabetes [33,34], tuberculosis [35], malaria [36], asthma [37], and HIV [38-40]. mHealth interventions have also been efficacious in reducing substance use [41]. In addition, bidirectional SMS text messaging interventions have significantly improved adherence to ART among people with HIV who use substances [42]. However, there has *never* been any systematic development or testing of mHealth interventions for people with HIV who use cocaine. Moreover, existing mHealth interventions have not optimized the integration of feedback to determine what type of feedback (eg, predetermined by researchers or clinicians, tailored to patients' preferences, dynamically tailored based on patient's adherence level, or feedback from friends and family members) improves outcomes. Prior studies have shown the effects of dynamically tailored feedback via SMS text messages on viral suppression compared with standard HIV care [43-45]. However, in these studies, the feedback was designed for people with HIV in a regulated clinical setting, which may not be effective for people with HIV using drugs in an unregulated community setting.

Study Objective

The primary objective of this paper is to provide a detailed description of a pilot mHealth-based randomized controlled trial (RCT) for improving ART adherence among people with HIV who use cocaine. A secondary goal is to list the critical challenges faced during the duration of the study and adaptations incorporated into the study design.

Pilot RCTs are small-scale studies designed to evaluate the feasibility and acceptability of a proposed intervention. Findings related to feasibility prepare investigators to conduct large-scale interventions, whereas acceptability shows how the population for whom the intervention is designed and the researchers who implement the study react to the intervention [46]. As part of a National Institute on Drug Abuse–funded study, we developed and conducted a feasibility study titled Project SMART, which developed a wireless technology–based intervention with multiple forms of feedback to improve ART adherence. In this paper, we aim to describe the research design of this pilot study, including the timeline, assessments, technologies, and challenges faced by researchers in implementing a multicomponent intervention. We also outline some practical adjustments made to the protocol during the study. By sharing our experiences, we aim to inform and educate other researchers interested in designing mHealth interventions for populations with similar characteristics. The following sections describe the methodologies and components of the pilot study as these are preanalysis protocol descriptions. The primary and secondary outcomes will be analyzed in forthcoming papers.

Methods

Study Setting

This study was conducted between May 2017 and January 2020. Participants were recruited from community centers, HIV care clinics, and a mobile clinic that has strategic liaisons with patients and providers throughout the Yale University community. The study took place within an HIV and opioid use disorder treatment facility, which is in an urban location in the northeastern part of the United States. The study was funded by the National Institute of Drug Abuse (R21DA039842).

Recruitment, Screening for Eligibility, and Consent

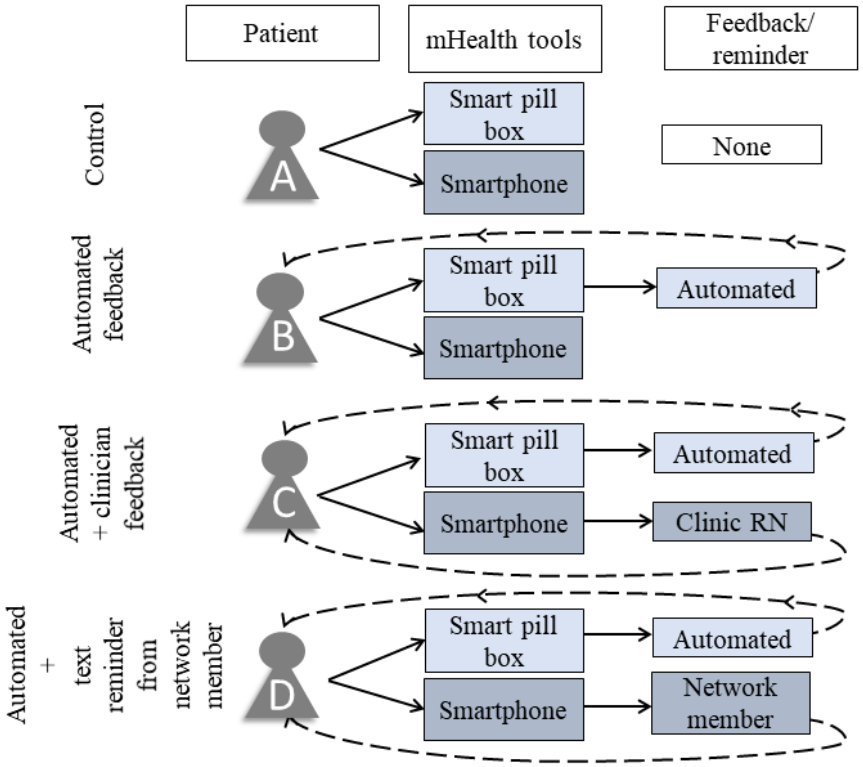
Recruitment started in May 2017 and continued until January 2020. Institutional review board (IRB)-approved flyers containing a toll-free number were posted in numerous venues such as community centers, HIV care clinics, and a mobile clinic that had strategic liaisons with patients and providers throughout the community. The initial screening eligibility criteria included patients (1) with a self-reported HIV diagnosis, (2) with self-reported use of cocaine ever, (3) having ART prescription and taking ART, and (4) willing and able to use a smartphone and electronic pillbox for a 16-week study period. After screening, the research staff obtained informed consent from the participants at their first visit and explained the purpose, protocol, expectations, and possible risks and benefits of the study.

Study Design

The design of this study is based on the fundamental components of transactional communication, which are the sender, receiver, and message [47]. In the transactional model of face-to-face communication, the sender and receiver engage in communication facilitated by feedback (verbal or nonverbal). With technologies such as the electronic pillbox, this exchange can take place only when the technology is responsive in some form to its user [48]. Interactive technologies enhance the engagement of its users [48]. Self-determination theory posits that the motivation to perform an action is influenced if an individual's need for competence, autonomy, and relatedness is fulfilled [49]. Using this assumption, the Motivation Technology Model (MTM) predicts that individuals can be motivated to change their behaviors if and when they achieve a sense of relatedness while using the technology [50]. MTM argues that this interactive nature of technology, which fulfills the need to be socially connected, can motivate individuals to perform or maintain certain health behaviors [50]. This perception of interactivity can be established when the technology provides timely responses and feedback to its users [50]. On the basis of these theoretical concepts of the self-determination theory and MTM, this study uses feedback as the central motivational factor to study adherence behavior.

The study design of Project SMART is shown in Figure 1. This was a 12-week pilot RCT that examined the use of two mHealth tools (cellular-enabled electronic pillboxes and smartphones) and four types of feedback—(1) no feedback, (2) automated feedback, (3) automated+clinician feedback, and (4) automated feedback+social network feedback (Figure 1)—on primary (ART adherence and persistence) and secondary outcomes (HIV viral suppression, cocaine use, retention in HIV care, and retention in RCT). The study period was 16 weeks, with a 12-week intervention period and a 4-week postintervention period. The following sections describe each of the components of the study design in detail.

Figure 1. Study design of Project SMART. RN: registered nurse.



mHealth Tools

Overview

Two mHealth tools—an electronic pillbox and smartphones—were used for intervention delivery. All study participants received these two mHealth tools but differed in the type of communication feedback they received: no feedback, automated feedback, automated and clinician feedback, and automated and social network feedback. The types of feedback determined the four arms of the RCT: 1 control arm and 3 intervention arms.

Electronic Pillbox

Overview

As shown in Figure 2, the TowerView Health pillbox is 10×7.5×1.5 inches, which is comparable with a handheld device such as an iPad, and can be linked to a smartphone. It has the capacity to hold pills or tablets for a weekly medication regimen of up to 4 dosages a day. The pills are nestled in 28 individual wells, each with its own sensor, to indicate when a dosage has been removed. The box has a built-in alarm that emits both a low sound and a flash of light. The pillboxes have 4G cellular capability via an integrated chip and hence have the ability to communicate with other devices such as smartphones. This capability is useful for transmitting real-time adherence data to the secure TowerView Health server and communicating with linked smartphones via SMS text messages.

Figure 2. Electronic pillbox TowerView Health used for Project SMART.

Medication and Refills

At the start of study enrollment, the clinician associated with Project SMART requested permission from participants' physicians to dispense ART medications from the selected pharmacy for the study period. The pharmacy collaborated with TowerView Health and delivered 4 sets (one for each week) of blister packs, with 1 months' worth of ART medication specifically designed to fit the dimensions of the pillbox. Although the pillboxes are designed to hold medication pills independently, blister packs were used by TowerView Health for ease of delivery and organization as they did not require manual insertion of individual pills. Our clinical staff placed these blister packs in the box in the participants' presence for the first time and trained the participants to replace the blister packs at the end of each week. Participants were instructed to bring their pillbox to the research site each month, at which time the research staff resupplied participants with a month's supply of medications. A 1-month supply of medication is the community standard of care for the dispensation of medications for patients with HIV [51].

Automated Feedback Messages

All participants' phones were connected to their assigned pillbox. The pillbox was programmed to send SMS text messages in the event of a missed dosage or messages of encouragement in the event of a successful daily dosage regimen. For example, when the pillbox was opened, and a segment of the blister pack was broken for medication intake, the sensor within the individual wells detected pressure, prompting a signal to be sent to the TowerView Health server. This then activated the automatic delivery of an SMS text

message to the participant's smartphone. The content of the SMS text messages or feedback delivered to a participant's smartphone was determined by the research team. Refer to section *Randomization Trial and Intervention* for examples of messages sent to participants.

Smartphones

The study used Samsung phones with an Android operating system, which were purchased through Yale University's agreement with Verizon. The participants received either a Samsung 7 or Samsung 8 model for the duration of the study, and the phone was configured for our study. The messaging app was activated so that participants could start receiving SMS text messages from the electronic pillbox. A total of 2 apps, *CommCare* and *SMS Back-Up and Restore*, were installed before handing out a phone to a participant (the role of these apps is described in the *Study Instruments* section). The research staff trained the participants on how to charge and use the electronic pillbox and smartphone (including using the voicemail and SMS text messaging features). In addition, the research staff programmed pertinent contact information (eg, clinic care and research staff contact) on each participant's phone.

Randomization Trial and Intervention

After completing the baseline interview, participants were assigned to 1 of the 4 arms systematically at a 1:1:1:1 ratio. Arm A was the control group, and participants in this group did not receive feedback as an intervention via the mHealth tools. The in-built and automatic feedback features of the TowerView Health smart pillboxes were also deactivated.

Arm B participants received automated feedback messages as an intervention. The box sent text dosage reminders to the participants' phones at the dosage time, and the rim around the box flashed blue, purple, and magenta light-emitting diode lights with a gentle ringing alarm. An example of the SMS text message is as follows: "Hi [participant name], don't forget to take your [dose time, eg, 3 PM] medications." If the pillbox was not opened at a dosage time, the box sent a reminder at 30 minutes and another at 90 minutes past the missed dosage time. The reminder said: "Hello [participant name], here's a quick reminder to take your missed pill of [dose time]." If at the end of the day, the participant took all the dosages, they received a note of encouragement on their phones; for example, "Congratulations, [participant name]! You took all your pills yesterday." These kinds of positive feedback were accompanied by emoticons such as a thumbs-up or clap. The positive messages varied daily and were selected from a set of messages designed by the team before the start of the study.

Arm C participants, in addition to the automated feedback from the pillbox, received feedback about their ART adherence from a clinician via phone calls on a weekly basis. The clinician's phone conversation was guided by the real-time adherence data provided by the TowerView Health dashboard.

For arm D, in addition to the automated feedback from the pillbox, participants received SMS text message reminders to take their medication from a person in their social network or a preselected social designee. The designee was a family member, friend, or acquaintance (excluding clinicians) who was willing to be part of the intervention and consented to their participation. The research team sent a weekly notification to the designee's mobile phone, prompting them to deliver the intervention via SMS text messages to the participant. The predefined automated notification was:

> *This is from Project SMART. In the next hour, send a text to [participant'name] @ [participant'number] to remind them to take their medication on time. Thank you for your help.!*

The designee then sent SMS text messages to remind the participant's phone about their ART dosages. No predefined SMS text messages for dyads (participant–social designee) were developed for this intervention. Instead, the actual contents of the SMS text message reminders between dyads were recorded as intervention data.

Participants in arms B, C, and D received their interventions for a duration of 12 weeks. After 12 weeks, the built-in feedback features of the electronic pillbox were deactivated such that during the 4-week postintervention period, the pillboxes only measured participants' pill-taking behavior, similar to arm A.

Study Instruments

Overview

Project SMART collected data using a range of instruments and standardized assessments administered at various time points of a participant's visits during preintervention (screening and baseline), intervention (4, 8, and 12 weeks), and postintervention (16 weeks) phases, spanning a period of 16 weeks. Table 1 contains a list of survey instruments.

Table 1. Study activity and instruments.

Study activity and measures	Preintervention			Intervention				Postintervention, week 16
	Before 4 weeks to before 1 week	Baseline (enrollment)	Preparation time	Week 0	Week 4	Week 8	Week 12	
Study activity								
Screening for eligibility	✓							
Release of information		✓						
Randomization			✓					
Device feature setting and preparation			✓					
Acquiring permission to change pharmacy and ordering medication			✓					
Device and medication distribution				✓				
Assigning social network designee		✓						
Phlebotomy		✓					✓	
Weekly smartphone survey (conducted weekly)					✓	✓	✓	✓
Demographic and social information								
Demographics		✓						
Social circumstances		✓			✓	✓	✓	✓
Health care status								
Health literacy [52]		✓						✓
Social support [52]		✓						✓
Trust in physician [52]								
Mental health and risk factors								
BINI[a]		✓						
Depression (CES-D[b]) [53]		✓						✓
AUDIT-C[c] [54]		✓						
Cocaine use disorder [55]		✓						
Substance use disorder		✓						
ASI-Lite[d] to assess cocaine use, polysubstance use, and severity [56,57]		✓			✓	✓	✓	✓
Timeline follow back [58]								
LifeWindows (LW-IMB-AAQ[e]) questionnaire to measure the information, motivation, and behavioral constructs of adherence [59,60]		✓						
Adherence								
Visual analog scale [61,62]		✓			✓	✓	✓	✓
Wilson adherence scale [63]		✓			✓	✓	✓	✓
Barriers to taking medication		✓					✓	
Technology perceptions								
Technology acceptance [64]		✓			✓		✓	

Study activity and measures	Preintervention			Intervention				Postintervention, week 16
	Before 4 weeks to before 1 week	Baseline (enrollment)	Preparation time	Week 0	Week 4	Week 8	Week 12	
Credibility of SMS text message source [65]					✓		✓	
Attitudes toward text and study							✓	
Geolocation and social network								
Activity space					✓		✓	
Social network					✓		✓	
Social network designee		✓						
Smartphone survey					✓	✓	✓	✓
Phone call history					✓	✓	✓	✓
Payments (US $)								
Research interviews	N/A[f]	30	N/A	N/A	30	20	30	20
Return of the devices (smartphone and pillbox)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	90
Weekly smartphone surveys	N/A	N/A	N/A	N/A	N/A	N/A	N/A	80
Incentive to social network designee	N/A	N/A	N/A	N/A	N/A	N/A	N/A	20

[a]BINI: Brief Inventory of Neurocognitive Impairment.

[b]CES-D: Center for Epidemiologic Studies–Depression.

[c]AUDIT-C: Alcohol Use Disorders Identification Test–Concise.

[d]ASI-Lite: Addiction Severity Index–Lite.

[e]LW-IMB-AAQ: Life Windows Information Motivation Behavioral skills ART Adherence Questionnaire.

[f]N/A: not applicable.

Screening

The screening instrument included questions regarding HIV status, health insurance, current prescription of ART, social designees, and ability and willingness to use a smart pillbox and smartphone for the duration of the study. This instrument determined the eligibility of a participant.

Baseline

The research staff completed the baseline survey after participants provided consent and a signed medical release of information for self-reported HIV status and HIV treatment regimen. Neurocognitive impairment (NCI) was assessed at baseline before assigning participants to the study arms. In addition, questions regarding their social networks (requested for arm D) were also administered at baseline. Phlebotomy measures using Quest Diagnostics were also conducted at baseline to conduct tests that included HIV-1 RNA testing (Amplicor 1.5, range 50-750,000 copies/mL), CD4 lymphocyte count using flow cytometry, and HIV genotypic mutations (if VL>500 copies/mL).

Weekly Survey

During the intervention period (12 weeks), participants completed a short survey per week using the *CommCare* app on their smartphones. *CommCare* is a survey platform designed specifically for mobile devices. The survey had a multiple-choice format with the questions regarding cocaine use in the past 7 days, triggers of cocaine use, protected and unprotected sexual intercourse following cocaine use, ART adherence in the past 7 days, and reasons for nonadherence. The research team sent an automated SMS text message reminder to complete the survey.

Monthly Research Interviews

A total of 4 monthly interviews took place at weeks 4, 8, 12, and 16, the latter of which was the postintervention interview. At the week 8 follow-up, participants also completed a set of questions on the acceptability and feasibility of the study components using the Technology Acceptance Model. The questions were about the ease of use and usefulness of smart pillboxes and smartphones, attitude toward adherence-related SMS text messages, and perceived credibility of those messages. For participants in arm D of the RCT, there was an additional instrument measuring the acceptability and receptivity of the social network intervention.

SMS Text Messages

SMS text message data were collected from arm D participants at weeks 4, 8, and 12. A total of 4 SMS text messages (1 SMS text message per week) sent by the social designee to the participant were downloaded from participants' phones by a

research staff member. These messages were then saved in a secure and password-protected drive at X University's Health Insurance Portability and Accountability Act–compliant server.

Study Completion and Device Recycling

At the week 16 interview, study participants were asked to return their pillbox and their smartphone. Incentives were built in for device returns (Table 1). The study team disinfected the phones and boxes and applied a factory reset to remove all the saved information from the prior user.

Payments

All participants were given payments in the form of gift cards for their time and participation, following the guidelines of our IRB. Participants were paid for monthly interviews, answering the weekly smartphone survey, and returning the smart pillboxes and smartphones at the end of the study. Refer to Table 1 for a description of the payments.

Analytical Plan

Primary Outcomes

The primary outcomes of this study were ART adherence and persistence, measured as a composite score ranging from 0% to 100%, and as the number of continuous days of taking ART without a 7-day (or 3-day) gap [66,67], respectively.

Secondary Outcomes

Outcome analyses will be conducted for (1) drug outcomes (cocaine use), (2) HIV treatment outcomes (VL<20 copies/mL; change in CD4+ T lymphocytes), and (3) ART-related information, motivation, and behavioral skills constructs. The general framework for discerning differences between study conditions on the main outcome variables will be to perform a 4 (intervention condition)×3 (repeated follow-up assessments) mixed design analysis by entering preintervention scores as covariates and additional theoretically and empirically relevant variables (such as baseline mental status). For variables that are approximately normally distributed, we will use multivariate analyses of covariance, with significant multivariate tests with Bonferroni correction followed by subsequent analyses of covariance. Variables that violate distributional assumptions of normality will be analyzed using generalized linear model procedures, which enable the use of nonparametric functions such as non–Gaussian error models. On the basis of a review of the literature, the generalized linear model is the most appropriate approach for variables with strong positive skew, such as adherence [63,68].

Ethical Oversight

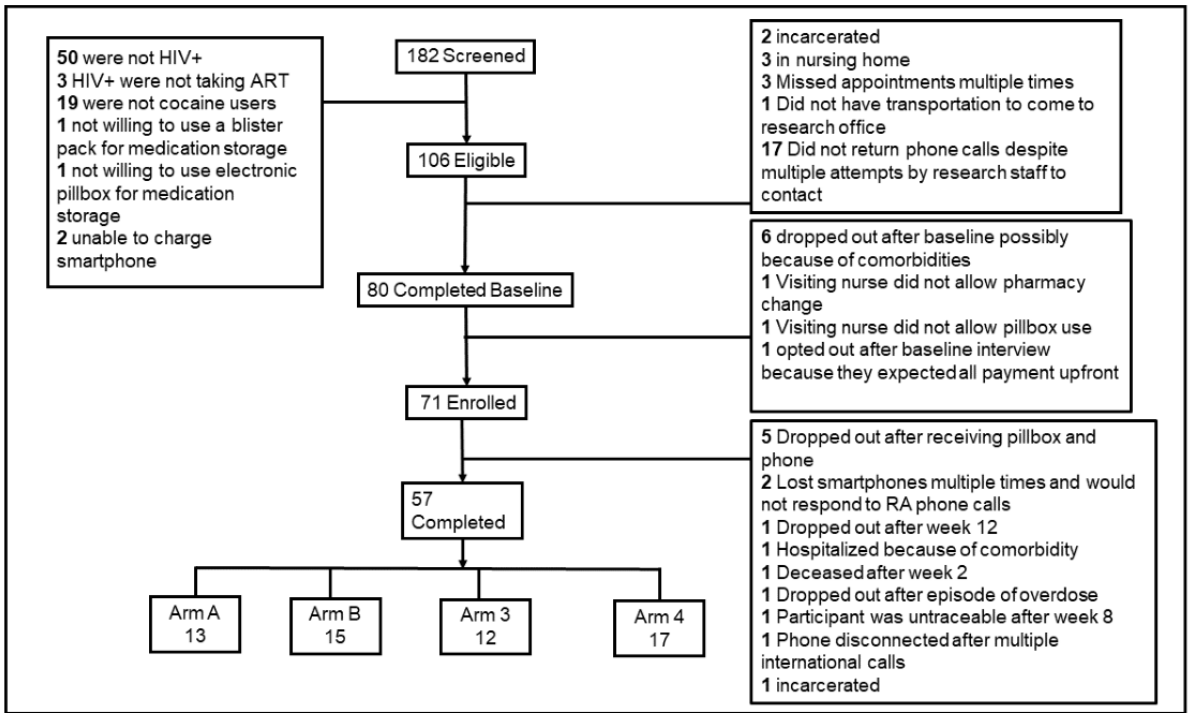
The IRB at Yale University reviewed and approved all study procedures (IRB:1508016342). Additional protections were provided by the National Institute of Health, which issued a certificate of confidentiality.

Results

Sample Size

On the basis of a review of pilot studies that have ranged in sample size from 12 per group [69-71] to 9% to 50% [72,73] of a large RCT's sample size (typically 20-30 per arm), we aimed to enroll 20 participants for each of the arms for a total of 80 participants. The final enrollment was 71 participants, of whom 57 (80%) completed the study. Figure 3 shows a complete CONSORT (Consolidated Standards of Reporting Trials) diagram. This was deemed appropriate for obtaining preliminary findings and optimizing a large-scale RCT for future studies [74].

Figure 3. Study flow of participants. ART: antiretroviral therapy; RA: research assistant.



Randomization

Participants were assigned to their arms after completing the baseline interview and before receiving ART medication, a pillbox, and a smartphone. Owing to the time required in acquiring permission to change pharmacy and receive medication for 4 weeks from the pharmacy, there was a time-lapse between baseline interview and actual start time of the study, termed *preparation time* (Table 1). During the preparation time, participants were randomized to their respective arms, the pillbox communicative features were set to the requirements of the study arms, and the smartphones were set up with the survey (*CommCare*) and the SMS text message (*SMS Back-Up and Restore*) apps.

We followed a systematic sequence of assigning participants to the study arms—the first participant to arm A, second to arm B, third to arm C, and fourth to arm D; this sequence was repeated until the desired sample size was reached. In addition, during the baseline interview, questions related to the NCI were administered using the Brief Inventory of Neurocognitive Impairment questionnaire [75,76]. As it is well-documented that individuals with underlying mental illness experience worse outcomes, we stratified participants based on the presence or absence of NCI using 155 out of 228 as the cutoff point [75,76]. Consecutive participants with high NCI were placed in different arms. The data are currently undergoing analysis, and the manuscript is being developed for publication in early 2022.

Discussion

Study Challenges and Adaptations

Smart Pillbox

The research team chose the Health Insurance Portability and Accountability Act–compliant TowerView Health electronic pillbox because of its ability to record adherence data in real time and send automated reminders to a participant's smartphone via SMS text messages. The *smart* features made this particular device more useful than frequently used pillboxes or bottles such as Glow Caps, MEMS Caps, or Wisepill in adherence research. As a start-up company based in Philadelphia, TowerView Health seemed to be the right choice for the study of the aforementioned reasons and for their active adherence research [77]. However, because of the unstable nature of start-up companies, this study experienced challenges worth noting.

First, the pillbox required a customized medication blister pack that would fit easily in the box. The local pharmacies at our study sites were not equipped with a printer to deliver these blister packs, creating the first barrier to adopting this pillbox. However, a pharmacy that had the capability to produce the customized blister packs required for the pillboxes was identified, located approximately 30 miles from the research site. The research team, along with TowerView Health, established a memorandum of understanding with the pharmacy that it would deliver participants' HIV medication after receiving the prescription from the health care provider. This required participants to change their pharmacy for HIV medication for the duration of the study, which sometimes contributed to complications for the timely delivery of medication.

Second, some devices had technical issues similar to any other cellular device with wireless capabilities, such as loss of connection or cellular towers not responding. Our team worked diligently with the company's engineering department and our participants to resolve such technical issues. One such incident, where pillboxes were not connected to the server because of negligence from the company's side, led to data loss. Our research team was able to address this data loss by replacing faulty pillboxes with new devices and extending the study period for the affected participants to compensate for the data loss. We reached this solution by fully disclosing the errors and discussing the serious consequences on the study outcomes with our participants and the IRB. With IRB approval and full consent of the affected participants, we were able to resume regular intervention delivery and data collection.

Third, toward the end of Project SMART, when we were 8 participants away from reaching our target enrollment, TowerView Health informed us that their company had ceased to exist because of their inability to find investors. This was an unexpected challenge for our team; nevertheless, we addressed this issue by introducing a new method for the collection of adherence data for the last 8 participants. We extended the traditional and well-accepted pill count method of adherence to a picture format, whereby participants uploaded pictures of their ART pills in a blister pack on a weekly basis [78,79]. The collected pictures will allow us to count the pills taken in a given week and calculate the necessary adherence metrics based on prescribed dosages and time elapsed from the last refill [78-80]. Although we lost the real-time adherence data collected by the electronic pillboxes, the picture pill count provides an appropriate alternative to other more invasive and labor-intensive methods of measuring adherence.

Smartphones

Most participants were able to return their smartphones at the end of the study, which was then factory reset and reused for other study participants. However, some (9/57, 16%) participants reported losing or damaging their phones, such that phones had to be replaced. mHealth studies with incentives have reported a similar proportion of returned devices with return incentives. This is encouraging as the return rate among people with multiple challenges because of cocaine use is comparable with that of other populations.

The use of prepared Samsung phones for the participants provided numerous lessons for future digital health interventions. Each participant phone needed 2 separate apps (*CommCare* and *SMS Back-Up and Restore*) downloaded from the Google Play Store installed before participant enrollment. To ensure the consistency of the mobile phone hardware and software provided to each participant, the research team needed to register a new Google account on each Samsung device for each new participant. To prepare batches of phones with separate Google accounts, the research team needed to validate each new account by receiving an SMS text message on their personal phones. At the beginning of the study, Google account permissions allowed a single phone number to validate

numerous new accounts. However, midway through the study, Google enacted more stringent account validation protocols, and the research team could not set up numerous accounts using 1 phone number. Thus, the creation of Google accounts to set up participant devices became a significant bottleneck for the project. The research team overcame this problem by using multiple phones to set up the accounts.

Time Length for Cocaine Use

One of the initial inclusion criteria was cocaine use in the past 30 days. This criterion was limiting, as most of the eligible participants reported using it infrequently or trying to quit. Hence, to address low recruitment, the study was adapted to first open up to people with HIV who had used cocaine in the past 60 days and then to people who had *ever used* to reach the target sample size, which was approved by the IRB.

Demands of the Study

Project SMART was conducted among people with HIV who were active cocaine users. Although most participants were able to adhere to the study protocol of using the electronic pillbox, using their smartphones to complete weekly surveys, and completing in-person interviews at the study site, almost one-fifth were unable to complete the study (14/71, 20%). The reasons for dropping out of the study were (1) comorbidities such as cancer and heart conditions, which required an extended period of hospitalization, affecting the ability to use the pillbox and participate in the study; (2) incarceration because of violent behavior while using illicit substances; (3) inability to answer phone calls and reminders for in-person interviews, leading to regularly missed in-person interviews; and (4) not being able to receive permission from nurses at the community homes to use medication adherence devices.

The RCT was a complex study that placed multiple demands on a population that already experienced constant instabilities, including insecurities related to health, social relations, or living conditions, because of drug use and other marginalized identities. The components of the study added to the burdens of living, including traveling to the research site, participating 4 times in an hour-long interview, learning how to use new technology, and completing multiple tasks throughout the period of the study. These demands created additional challenges for our participants, which resulted in a high dropout rate. Hence, in the future, similar mHealth interventions should aim to minimize these demands to reduce the burden on marginalized participants and maintain retention in the study.

This study lends itself to the emerging theoretical work on complex health intervention frameworks [81]. Complex health interventions are characterized by the complexity of the intervention at the following levels: the design of the intervention, the people involved (those who implement as well as those to receive it), at the organization and workforce, and at the outcome level. The 2 important constructs of this framework are the *core function* and *forms* of the intervention. *Core function* refers to the main purpose or goal that the intervention seeks to achieve, and *form* includes the strategies or activities of the intervention to achieve those goals [81]. The complex health intervention framework, which was based on research on patient-centered medical home interventions, argues that the core function of health interventions can be achieved through forms, which may vary greatly based on the cultural and contextual needs of the individuals that they serve. Hence, complex health interventions will change based on the context-specific needs of the patients or populations they serve as well as those who implement the intervention [81]. In this study, we acknowledge that the electronic pillbox was designed for individuals who have greater stability and fewer comorbidities than individuals with HIV who use cocaine. Therefore, future interventions should account for the complex and changing needs of this population, their health care providers (eg, nurses at community homes), and other characteristics in the early stages of the implementation and adapt the services to the changing demands of the intervention. Although this study was able to adjust the intervention in the latter part of the study by replacing the electronic pillbox with a manual blister package, earlier assessment and adaptation would have better facilitated the core function. In addition, the framework suggests that researchers should develop a *function and form matrix* in anticipation of challenges they may face in implementing complex interventions [81]. Hence, in the future, researchers should create a framework with the core functions of the interventions and possible adaptations of the forms to be prepared for the complexities that could arise at the various levels of the intervention.

Conclusions

This paper provides detailed information about an RCT involving an mHealth intervention to improve adherence to ART for people with HIV who use cocaine. The 4-armed RCT with a control group sought to assess the feasibility and acceptability of technology interventions such as smart pillboxes and smartphones among at-risk, vulnerable populations, for whom currently the only effective adherence intervention is directly administered ART, which is expensive and labor intensive. The pilot study described here presents the different components of a technology-based intervention that can provide options for scaled-back strategies, both in terms of cost and personnel, as well as highlighting some challenges and adaptations required to consider when planning for a large-scale study.

Acknowledgments

The authors would like to express their gratitude to the research and clinical staff of Yale Center for Community and Clinical Research Office, who navigated through the complexity of this study with positivity and sincerity. The authors would also like to acknowledge the administrative prowess of Dr Ruthanne Marcus, who ensured the functioning of Project SMART. Finally, the entire team would like to thank the research participants without whose cooperation this study would not have been possible.

Funding support for this work was provided by R21 (AK) and K01 (DG).

Conflicts of Interest

FLA declares the following conflicts of interest, which are unrelated to this research: research grant support to Yale University where he is listed as principal investigator or co-principal investigator (National Institute of Health, National Institute on Drug Abuse, National Institute on Alcohol Abuse and Alcoholism, Health Resources and Services Administration, Substance Abuse and Mental Health Services Administration, Gilead Foundation, Merck Pharmaceuticals); Speakers Bureau (Gilead Sciences, Practice Point Communications Simply Speaking HIV, Clinical Care Options); and Advisory Board Membership (Abbvie, Gilead Sciences, Alliance for Patient Access).

Multimedia Appendix 1
Peer review report from the Behavioral and Social Consequences of HIV/AIDS Study Section - Center for Scientific Review (National Institutes of Health).
[PDF File (Adobe PDF File), 125 KB - resprot_v11i3e28332_app1.pdf]

References

1. Hammond R, Harry TC. Efficacy of antiretroviral therapy in Africa: effect on immunological and virological outcome measures - a meta-analysis. Int J STD AIDS 2008 May 01;19(5):291-296. [doi: 10.1258/ijsa.2007.007248] [Medline: 18482957]
2. Bartlett J, Fath MJ, Demasi R, Hermes A, Quinn J, Mondou E, et al. An updated systematic overview of triple combination therapy in antiretroviral-naive HIV-infected adults. AIDS 2006 Oct 24;20(16):2051-2064. [doi: 10.1097/01.aids.0000247578.08449.ff] [Medline: 17053351]
3. HIV treatment overview. HIV - U.S. Department of Health & Human Services. URL: https://www.hiv.gov/hiv-basics/staying-in-hiv-care/hiv-treatment/hiv-treatment-overview [accessed 2022-01-28]
4. HIV treatment. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/hiv/basics/livingwithhiv/treatment.html [accessed 2022-01-28]
5. Velasco-Hernandez J, Gershengorn H, Blower S. Could widespread use of combination antiretroviral therapy eradicate HIV epidemics? Lancet Infect Dis 2002 Aug;2(8):487-493. [doi: 10.1016/s1473-3099(02)00346-8]
6. Montaner JS, Hogg R, Wood E, Kerr T, Tyndall M, Levy AR, et al. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. Lancet 2006 Aug;368(9534):531-536. [doi: 10.1016/s0140-6736(06)69162-9]
7. Lima V, Johnston K, Hogg R, Levy A, Harrigan P, Anema A, et al. Expanded access to highly active antiretroviral therapy: a potentially powerful strategy to curb the growth of the HIV epidemic. J Infect Dis 2008 Jul 01;198(1):59-67. [doi: 10.1086/588673] [Medline: 18498241]
8. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, HPTN 052 Study Team. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011 Aug 11;365(6):493-505 [FREE Full text] [doi: 10.1056/NEJMoa1105243] [Medline: 21767103]
9. Altice F, Evuarherhe O, Shina S, Carter G, Beaubrun AC. Adherence to HIV treatment regimens: systematic literature review and meta-analysis. Patient Prefer Adher 2019 Apr;Volume 13:475-490. [doi: 10.2147/ppa.s192735]
10. Ortego C, Huedo-Medina TB, Llorca J, Sevilla L, Santos P, Rodríguez E, et al. Adherence to highly active antiretroviral therapy (HAART): a meta-analysis. AIDS Behav 2011 Oct 6;15(7):1381-1396. [doi: 10.1007/s10461-011-9942-x] [Medline: 21468660]
11. Altice FL, Kamarulzaman A, Soriano VV, Schechter M, Friedland GH. Treatment of medical, psychiatric, and substance-use comorbidities in people infected with HIV who use drugs. Lancet 2010 Jul;376(9738):367-387. [doi: 10.1016/s0140-6736(10)60829-x]
12. Uthman OA, Okwundu C, Gbenga K, Volmink J, Dowdy D, Zumla A, et al. Optimal timing of antiretroviral therapy initiation for HIV-infected adults with newly diagnosed pulmonary tuberculosis. Ann Intern Med 2015 Jul 07;163(1):32-39. [doi: 10.7326/m14-2979]
13. Zhang S, Rust G, Cardarelli K, Felizzola J, Fransua M, Stringer HG. Adherence to highly active antiretroviral therapy impact on clinical and economic outcomes for Medicaid enrollees with human immunodeficiency virus and hepatitis C coinfection. AIDS Care 2015 Mar 27;27(7):829-835 [FREE Full text] [doi: 10.1080/09540121.2015.1021745] [Medline: 25814041]
14. Paschen-Wolff MM, Campbell AN, Tross S, Choo T, Pavlicova M, Jarlais DD. DSM-5 substance use disorder symptom clusters and HIV antiretroviral therapy (ART) adherence. AIDS Care 2020 May 04;32(5):645-650 [FREE Full text] [doi: 10.1080/09540121.2019.1686600] [Medline: 31682153]
15. Surie D, Borgdorff MW, Cain KP, Click ES, DeCock KM, Yuen CM. Assessing the impact of antiretroviral therapy on tuberculosis notification rates among people with HIV: a descriptive analysis of 23 countries in sub-Saharan Africa, 2010-2015. BMC Infect Dis 2018 Sep 26;18(1):481 [FREE Full text] [doi: 10.1186/s12879-018-3387-z] [Medline: 30257667]

16. Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. Clin Infect Dis 2011 Mar 15;52(6):793-800 [FREE Full text] [doi: 10.1093/cid/ciq243] [Medline: 21367734]
17. Westergaard RP, Ambrose BK, Mehta SH, Kirk GD. Provider and clinic-level correlates of deferring antiretroviral therapy for people who inject drugs: a survey of North American HIV providers. J Int AIDS Soc 2012 Feb 23;15(1):10 [FREE Full text] [doi: 10.1186/1758-2652-15-10] [Medline: 22360788]
18. Lucas GM, Cheever LW, Chaisson RE, Moore RD. Detrimental effects of continued illicit drug use on the treatment of HIV-1 infection. J Acquir Immune Defic Syndr 2001;27(3):251-259. [doi: 10.1097/00042560-200107010-00006]
19. Lewden C, Chene G, Morlat P, Raffi F, Dupon M, Dellamonica P, Agence Nationale de Recherches sur le Sida et les Hepatites Virales (ANRS) CO8 APROCO-COPILOTE Study Group, Agence Nationale de Recherches sur le Sida et les Hepatites Virales (ANRS) CO3 AQUITAINE Study Group. HIV-infected adults with a CD4 cell count greater than 500 cells/mm3 on long-term combination antiretroviral therapy reach same mortality rates as the general population. J Acquir Immune Defic Syndr 2007 Sep 01;46(1):72-77. [doi: 10.1097/QAI.0b013e318134257a] [Medline: 17621240]
20. Cunningham CO, Sohler NL, Berg KM, Shapiro S, Heller D. Type of substance use and access to HIV-related health care. AIDS Patient Care STDS 2006 Jun;20(6):399-407. [doi: 10.1089/apc.2006.20.399] [Medline: 16789853]
21. Metsch LR, Pereyra M, Brewer TH. Use of HIV health care in HIV-seropositive crack cocaine smokers and other active drug users. J Subst Abuse 2001 Sep;13(1-2):155-167. [doi: 10.1016/s0899-3289(01)00064-5]
22. Sohler NL, Wong MD, Cunningham WE, Cabral H, Drainoni M, Cunningham CO. Type and pattern of illicit drug use and access to health care services for HIV-infected people. AIDS Patient Care STDS 2007 Jun;21 Suppl 1(s1):68-76. [doi: 10.1089/apc.2007.9985] [Medline: 17563292]
23. Arnsten JH, Demas PA, Grant RW, Gourevitch MN, Farzadegan H, Howard AA, et al. Impact of active drug use on antiretroviral therapy adherence and viral suppression in HIV-infected drug users. J Gen Intern Med 2002 May;17(5):377-381. [doi: 10.1007/s11606-002-0044-3]
24. Ingersoll K. The impact of psychiatric symptoms, drug use, and medication regimen on non-adherence to HIV treatment. AIDS Care 2004 Feb 12;16(2):199-211. [doi: 10.1080/09540120410001641048] [Medline: 14676026]
25. Sharpe TT, Lee LM, Nakashima AK, Elam-Evans LD, Fleming PL. Crack cocaine use and adherence to antiretroviral treatment among HIV-infected black women. J Community Health 2004 Apr;29(2):117-127. [doi: 10.1023/b:johe.0000016716.99847.9b]
26. Cook J, Burke-Miller JK, Cohen MH, Cook RL, Vlahov D, Wilson TE, et al. Crack cocaine, disease progression, and mortality in a multicenter cohort of HIV-1 positive women. AIDS 2008 Jul 11;22(11):1355-1363 [FREE Full text] [doi: 10.1097/QAD.0b013e32830507f2] [Medline: 18580615]
27. Maru DS, Bruce RD, Walton M, Mezger JA, Springer SA, Shield D, et al. Initiation, adherence, and retention in a randomized controlled trial of directly administered antiretroviral therapy. AIDS Behav 2008 Mar 18;12(2):284-293 [FREE Full text] [doi: 10.1007/s10461-007-9336-2] [Medline: 18085432]
28. Thompson MA, Mugavero MJ, Amico KR, Cargill VA, Chang LW, Gross R, et al. Guidelines for improving entry into and retention in care and antiretroviral adherence for persons with HIV: evidence-based recommendations from an International Association of Physicians in AIDS Care panel. Ann Intern Med 2012 Jun 05;156(11):817-833 [FREE Full text] [doi: 10.7326/0003-4819-156-11-201206050-00419] [Medline: 22393036]
29. Boyer EW, Smelson D, Fletcher R, Ziedonis D, Picard RW. Wireless technologies, ubiquitous computing and mobile health: application to drug abuse treatment and compliance with HIV therapies. J Med Toxicol 2010 Jun 10;6(2):212-216 [FREE Full text] [doi: 10.1007/s13181-010-0080-z] [Medline: 20623215]
30. Lester R, Gelmon L, Plummer FA. Cell phones: tightening the communication gap in resource-limited antiretroviral programmes? AIDS 2006 Nov 14;20(17):2242-2244. [doi: 10.1097/QAD.0b013e3280108508] [Medline: 17086071]
31. World Health Organization. Wireless Technologies, Ubiquitous Computing and Mobile Health: Application to Drug Abuse Treatment and Compliance With HIV Therapies. Geneva: World Health Organization; 2013. [doi: 10.1007/s13181-010-0080-z]
32. Leach-Lemens C. Using mobile phones in HIV care and prevention. HIV & AIDS Treat Pract 2009;137:7 [FREE Full text]
33. Hanauer DA, Wentzell K, Laffel N, Laffel LM. Computerized Automated Reminder Diabetes System (CARDS): e-mail and SMS cell phone text messaging reminders to support diabetes management. Diabetes Technol Ther 2009 Feb;11(2):99-106 [FREE Full text] [doi: 10.1089/dia.2008.0022] [Medline: 19848576]
34. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
35. Hoffman JA, Cunningham JR, Suleh AJ, Sundsmo A, Dekker D, Vago F, et al. Mobile direct observation treatment for tuberculosis patients: a technical feasibility pilot using mobile phones in Nairobi, Kenya. Am J Prev Med 2010 Jul;39(1):78-80. [doi: 10.1016/j.amepre.2010.02.018] [Medline: 20537846]
36. Asiimwe C, Gelvin D, Lee E, Ben Amor Y, Quinto E, Katureebe C, et al. Use of an innovative, affordable, and open-source short message service-based tool to monitor malaria in remote areas of Uganda. Am J Trop Med Hyg 2011 Jul;85(1):26-33 [FREE Full text] [doi: 10.4269/ajtmh.2011.10-0528] [Medline: 21734120]

37. Ryan D, Pinnock H, Lee AJ, Tarassenko L, Pagliari C, Sheikh A, et al. The CYMPLA trial. Mobile phone-based structured intervention to achieve asthma control in patients with uncontrolled persistent asthma: a pragmatic randomised controlled trial. Prim Care Respir J 2009 Dec 25;18(4):343-345 [FREE Full text] [doi: 10.4104/pcrj.2009.00064] [Medline: 19940961]
38. Lester R, Kariri A. Mobilizing cell phones to improve antiretroviral adherence and follow-up in Kenya: a randomized controlled trial in progress. WHO Essential Medicines Monitor. 2009. URL: https://www.who.int/medicines/publications/monitor/EMMArt1Issue22009.pdf [accessed 2022-01-28]
39. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. Lancet 2010 Nov;376(9755):1838-1845. [doi: 10.1016/s0140-6736(10)61997-6]
40. Harris LT, Lehavot K, Huh D, Yard S, Andrasik MP, Dunbar PJ, et al. Two-way text messaging for health behavior change among human immunodeficiency virus-positive individuals. Telemed J E Health 2010 Dec;16(10):1024-1029 [FREE Full text] [doi: 10.1089/tmj.2010.0050] [Medline: 21087122]
41. Kazemi DM, Borsari B, Levine MJ, Li S, Lamberson KA, Matta LA. A systematic review of the mHealth interventions to prevent alcohol and substance abuse. J Health Commun 2017 May 10;22(5):413-432 [FREE Full text] [doi: 10.1080/10810730.2017.1303556] [Medline: 28394729]
42. Ingersoll KS, Dillingham RA, Hettema JE, Conaway M, Freeman J, Reynolds G, et al. Pilot RCT of bidirectional text messaging for ART adherence among nonurban substance users with HIV. Health Psychol 2015 Dec;34S(Suppl):1305-1315 [FREE Full text] [doi: 10.1037/hea0000295] [Medline: 26651472]
43. Lewis MA, Uhrig JD, Bann CM, Harris JL, Furberg RD, Coomes C, et al. Tailored text messaging intervention for HIV adherence: a proof-of-concept study. Health Psychol 2013 Mar;32(3):248-253. [doi: 10.1037/a0028109] [Medline: 22545972]
44. Uhrig JD, Lewis MA, Bann CM, Harris JL, Furberg RD, Coomes CM, et al. Addressing HIV knowledge, risk reduction, social support, and patient involvement using SMS: results of a proof-of-concept study. J Health Commun 2012 May 02;17 Suppl 1(sup1):128-145. [doi: 10.1080/10810730.2011.649156] [Medline: 22548606]
45. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database Syst Rev 2012 Mar 14(3):CD009756 [FREE Full text] [doi: 10.1002/14651858.CD009756] [Medline: 22419345]
46. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
47. Shannon CE, Weaver W. The Mathematical Theory of Communication. Champaign, Illinois: University of Illinois Press; 1949.
48. Sundar SS, Bellur S, Oh J, Jia H, Kim H. Theoretical importance of contingency in human-computer interaction. Commun Res 2014 May 22;43(5):595-625. [doi: 10.1177/0093650214534962]
49. Ryan RM, Deci EL. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness. New York: Guilford Press; 2017.
50. Sundar S, Bellur S, Jia H. Motivational technologies: a theoretical framework for designing preventive health applications. In: Proceedings of the International Conference on Persuasive Technology. 2012 Presented at: International Conference on Persuasive Technology; June 6-8, 2012; Linköping, Sweden p. 112-122. [doi: 10.1007/978-3-642-31037-9_10]
51. Dybul M, Fauci AS, Bartlett JG, Kaplan JE, Pau AK, Panel on Clinical Practices for the Treatment of HIV. Guidelines for using antiretroviral agents among HIV-infected adults and adolescents. Recommendations of the Panel on Clinical Practices for Treatment of HIV. MMWR Recomm Rep 2002 May 17;51(RR-7):1-55 [FREE Full text] [Medline: 12027060]
52. 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]
53. Radloff LS. The CES-D Scale: a self report depression scale for research in the general population. Appl Psychol Measur 2016 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]
54. 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]
55. American Psychiatric Association. Diagnostic And Statistical Manual Of Mental Disorders. 5 Ed. Washington, DC: American Psychiatric Association Publishing; 2013.
56. McLellan A, Kushner H, Metzger D, Peters R, Smith I, Grissom G, et al. The fifth edition of the addiction severity index. J Subst Abuse Treat 1992 Jun;9(3):199-213. [doi: 10.1016/0740-5472(92)90062-s]
57. Rosen C, Henson BR, Finney JW, Moos RH. Consistency of self-administered and interview-based Addiction Severity Index composite scores. Addiction 2000 Mar;95(3):419-425. [doi: 10.1046/j.1360-0443.2000.95341912.x] [Medline: 10795362]
58. Robinson SM, Sobell LC, Sobell MB, Leo GI. Reliability of the Timeline Followback for cocaine, cannabis, and cigarette use. Psychol Addict Behav 2014 Mar;28(1):154-162. [doi: 10.1037/a0030992] [Medline: 23276315]

59. Amico KR, Toro-Alfonso J, Fisher JD. An empirical test of the information, motivation and behavioral skills model of antiretroviral therapy adherence. AIDS Care 2005 Aug 20;17(6):661-673. [doi: 10.1080/09540120500038058] [Medline: 16036253]
60. Fisher JD, Amico KR, Fisher WA, Cornman DH, Shuper PA, Trayling C, LifeWindows Team. Computer-based intervention in HIV clinical care setting improves antiretroviral adherence: the LifeWindows Project. AIDS Behav 2011 Nov 31;15(8):1635-1646. [doi: 10.1007/s10461-011-9926-x] [Medline: 21452051]
61. Amico K, Fisher WA, Cornman DH, Shuper PA, Redding CG, Konkle-Parker DJ, et al. Visual analog scale of ART adherence: association with 3-day self-report and adherence barriers. J Acquir Immune Defic Syndr 2006 Aug 01;42(4):455-459. [doi: 10.1097/01.qai.0000225020.73760.c2] [Medline: 16810111]
62. Wilson IB, Lee Y, Michaud J, Fowler FJ, Rogers WH. Validation of a new three-item self-report measure for medication adherence. AIDS Behav 2016 Nov 20;20(11):2700-2708 [FREE Full text] [doi: 10.1007/s10461-016-1406-x] [Medline: 27098408]
63. Neal DJ, Simons JS. Inference in regression models of heavily skewed alcohol use data: a comparison of ordinary least squares, generalized linear models, and bootstrap resampling. Psychol Addict Behav 2007 Dec;21(4):441-452. [doi: 10.1037/0893-164x.21.4.441]
64. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
65. Hass RG. Effects of source characteristics on cognitive responses on persuasion. In: Cognitive Responses in Persuasion. Hillsdale, N.J: Lawrence Erlbaum Associates; 1981.
66. Ing EC, Bae JW, Maru DS, Altice FL. Medication persistence of HIV-infected drug users on directly administered antiretroviral therapy. AIDS Behav 2013 Jan 22;17(1):113-121. [doi: 10.1007/s10461-011-0082-0] [Medline: 22105340]
67. Bae J, Guyer W, Grimm K, Altice FL. Medication persistence in the treatment of HIV infection: a review of the literature and implications for future clinical care and research. AIDS 2011 Jan 28;25(3):279-290. [doi: 10.1097/QAD.0b013e328340feb0] [Medline: 21239892]
68. Ng VK, Cribbie RA. Using the Gamma Generalized Linear Model for modeling continuous, skewed and heteroscedastic outcomes in psychology. Curr Psychol 2016 Jan 9;36(2):225-235. [doi: 10.1007/s12144-015-9404-0]
69. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharmaceut Statist 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
70. Birkett MA, Day SJ. Internal pilot studies for estimating sample size. Stat Med 1994 Dec 15;13(23-24):2455-2463. [doi: 10.1002/sim.4780132309] [Medline: 7701146]
71. Browne RH. On the use of a pilot sample for sample size determination. Stat Med 1995 Sep 15;14(17):1933-1940. [doi: 10.1002/sim.4780141709] [Medline: 8532986]
72. Cocks K, Torgerson DJ. Sample size calculations for pilot randomized trials: a confidence interval approach. J Clin Epidemiol 2013 Feb;66(2):197-201. [doi: 10.1016/j.jclinepi.2012.09.002] [Medline: 23195919]
73. Wittes J, Brittain E. The role of internal pilot studies in increasing the efficiency of clinical trials. Stat Med 1990 Jan;9(1-2):65-71. [doi: 10.1002/sim.4780090113] [Medline: 2345839]
74. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun 19;25(3):1057-1073 [FREE Full text] [doi: 10.1177/0962280215588241] [Medline: 26092476]
75. Stern Y, McDermott MP, Albert S, Palumbo D, Selnes OA, McArthur J, Dana Consortium on the Therapy of HIV-DementiaRelated Cognitive Disorders. Factors associated with incident human immunodeficiency virus-dementia. Arch Neurol 2001 Mar 01;58(3):473-479. [doi: 10.1001/archneur.58.3.473] [Medline: 11255452]
76. Anand P, Springer SA, Copenhaver MM, Altice FL. Neurocognitive impairment and HIV risk factors: a reciprocal relationship. AIDS Behav 2010 Dec 16;14(6):1213-1226 [FREE Full text] [doi: 10.1007/s10461-010-9684-1] [Medline: 20232242]
77. George J. DreamIt Health company makes pill taking easier with help from big name partners. Philadelphia Business Journal. 2015. URL: https://www.bizjournals.com/philadelphia/blog/health-care/2015/05/dreamit-health-company-makes-pill-taking-easier.html [accessed 2022-01-28]
78. Wu P, Johnson B, Nachega J, Wu B, Ordonez C, Hare A, et al. The combination of pill count and self-reported adherence is a strong predictor of first-line ART failure for adults in South Africa. Curr HIV Res 2014 Nov 21;12(5):366-375. [doi: 10.2174/1570162x1205141121102501] [Medline: 25426940]
79. Basu S, Garg S. Methods for estimating medication adherence to antiretroviral therapy: response to Mehta et al. (2016). J Family Med Prim Care 2017;6(1):171-172 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_9_17] [Medline: 29026778]
80. Holstad MM, Higgins M, Bauman M, Farber EW, Waldrop-Valverde D, Okwandu O, et al. Picture pill count: an innovative, reliable, valid and feasible method to measure adherence to ART. AIDS Behav 2019 Aug 30;23(8):2210-2217. [doi: 10.1007/s10461-019-02513-9] [Medline: 31041625]
81. Jolles MP, Lengnick-Hall R, Mittman BS. Core functions and forms of complex health interventions: a patient-centered medical home illustration. J Gen Intern Med 2019 Jun 8;34(6):1032-1038 [FREE Full text] [doi: 10.1007/s11606-018-4818-7] [Medline: 30623387]

Abbreviations

ART: antiretroviral therapy
CONSORT: Consolidated Standards of Reporting Trials
IRB: institutional review boardmHealth: mobile healthMTM: Motivation Technology Model
NCI: neurocognitive impairment
RCT: randomized controlled trial
VL: viral load

Edited by G Eysenbach; submitted 01.03.21; this is a non–peer-reviewed article;accepted 22.11.21; published 07.03.22.

Please cite as:
Ranjit YS, Krishnan A, Ghosh D, Cravero C, Zhou X, Altice FL
mHealth Intervention to Improve Treatment Outcomes Among People With HIV Who Use Cocaine: Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2022;11(3):e28332
URL: https://www.researchprotocols.org/2022/3/e28332
doi:10.2196/28332
PMID:35254270

Protocol

Developing a Smartphone-Based Adjunct Intervention to Reduce Cannabis Use Among Juvenile Justice-Involved Adolescents: Protocol for a Multiphase Study

Sarah A Helseth[1], PhD; John Guigayoma[1], MPH; Dayna Price[1], BA; Anthony Spirito[2], PhD; Melissa A Clark[3,4], PhD; Nancy P Barnett[1,2], PhD; Sara J Becker[1,2], PhD

[1]Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, United States

[2]Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, Providence, RI, United States

[3]Department of Health Services Policy & Practice, Brown University School of Public Health, Providence, RI, United States

[4]Department of Obstetrics & Gynecology, Alpert Medical School of Brown University, Providence, RI, United States

Corresponding Author:
Sarah A Helseth, PhD
Department of Behavioral and Social Sciences
Brown University School of Public Health
121 South Main Street, 5th Floor
Providence, RI,
United States
Phone: 1 401 863 6881
Email: sarah_helseth@brown.edu

Abstract

Background: Adolescents involved in the juvenile justice system who use cannabis are at an increased risk of future substance use disorders and rearrest. Many court-involved, nonincarcerated (CINI) youth are referred for services in the community and often encounter multiple barriers to care, highlighting the need for minimally burdensome services that can be delivered in justice settings. Digital health interventions are accessible, easy to implement, and can provide ongoing support but have not been developed to address the unique needs of CINI youth who use cannabis.

Objective: This multiphase study will aim to develop, implement, and pilot test a novel smartphone app, Teen Empowerment through Computerized Health (TECH), to reduce cannabis and other substance use among CINI youth. TECH is conceptualized as a digital adjunct to a brief computerized intervention delivered by our family court partner.

Methods: Following the principles of user-centered design, phase I interviews with CINI youth aged 14-18 years (n=14-18), their caregivers (n=6-8), and behavioral health app developers (n=6-8) will guide the TECH design decisions. Next, in phase II, CINI youth (n=10) will beta test the TECH app prototype for 1 month; their feedback regarding feasibility and acceptability will directly inform the app refinement process. Finally, in phase III, CINI youth (n=60) will participate in a pilot randomized controlled trial for 6 months, comparing the preliminary effectiveness of the adjunctive TECH app on cannabis use outcomes.

Results: Phase I data collection began in September 2020 and was completed in December 2021; 14 CINI youth, 8 caregivers, and 11 behavioral health app developers participated in the study. Phases II and III will occur in 2022 and 2023 and 2023 and 2025, respectively.

Conclusions: This body of work will provide insight into the feasibility and acceptability of a smartphone-based adjunctive intervention designed for CINI youth. Phase III results will offer a preliminary indication of the effectiveness of the TECH app in reducing cannabis use among CINI youth.

International Registered Report Identifier (IRRID): DERR1-10.2196/35402

(JMIR Res Protoc 2022;11(3):e35402) doi:10.2196/35402

KEYWORDS

mobile intervention; juvenile justice; justice; court; adolescent; teenager; substance use; cannabis; youth; adolescence; protocol; mHealth; mobile health; substance use; user design; behavioral app; health app; development; pilot; prototype; feasibility; acceptability; smartphone app; marijuana; mobile phone

Introduction

Background

Cannabis use is common among court-involved, nonincarcerated (CINI) youth, who comprise 74% of justice-involved youth [1]—of those who test positive for a substance at the time of their arrest, 92% test positive for cannabis [2]. Early-onset cannabis use has been linked to many negative long-term outcomes, including violent behavior [3], criminal justice involvement [4,5], and substance use (SU) disorders [6]. In addition, cannabis use is a risk factor for subsequent rearrest and detention among CINI youth, perpetuating their involvement in the juvenile justice system [7]. The synergistic, maladaptive relationship between early cannabis use and juvenile justice involvement [8] highlights the need for early interventions targeting cannabis use in CINI youth [9]. Unlike detained youth who can be funneled directly into treatment on-site [10], CINI youth are typically referred to community providers, shifting treatment-seeking responsibilities to CINI youth and their families. Families involved with the legal system often face multiple competing demands that can make it difficult to engage in treatment [11], even when court-mandated, placing CINI youth at risk for escalating legal consequences and SU problems. Recent research indicates that only 28% of detained youth [12] and <30% of CINI youth [13] in need of SU services ultimately receive any treatment. Given that many court agencies were not designed to deliver SU treatments, on-site services need to minimize resource demands, maximize accessibility, and above all, be effective.

Although research on effective interventions to reduce cannabis use among CINI youth is in its infancy, theory and evidence from the broader adolescent SU literature may offer valuable insight. To date, theories of adolescent SU [14] have centered on several key intrapersonal and interpersonal mechanisms of SU behavior. Intrapersonal mechanisms, including one's attitudes, self-efficacy, beliefs, and SU expectancies, are thought to be informed via didactic (eg, SU education [15]) and social (witnessing parent [16-18] or peer [19] SU) learning and later refined by direct experiences. Theorized interpersonal mechanisms of SU behavior encompass both direct influence from others, such as social reinforcement or modeling [20,21], and indirect influences, including perceived norms of peers' SU [14,22].

Recent systematic and meta-analytic reviews of SU treatments for youth [23-25] indicate that cognitive-behavioral, family-based, and motivation-enhancing approaches are most effective at decreasing SU behaviors and problems in adolescents. Established intervention approaches primarily target intrapersonal mechanisms of SU. Cognitive-behavioral and family-based approaches can take 12 to 16 sessions [26], whereas motivation-enhancing models require fewer sessions and organizational resources, suggesting they could potentially be more cost-effective [27-29] and disseminable [30] in justice settings. Although brief motivation-enhancing approaches have shown promise among justice-involved youth [31,32], such interventions could be strengthened by incorporating components that also target interpersonal mechanisms of SU. For example, a recent randomized controlled trial (RCT) of a motivation-enhancing intervention incorporating accurate peer cannabis use norms found that participants reported more accurate peer norms after treatment [33]. In turn, changes in peer norms and one's own approval of cannabis use were linked to posttreatment improvements in cannabis use and related outcomes. Finding new ways to foster peer engagement and build supportive communities in adolescent treatment contexts akin to SU recovery (eg, 12-step programs, peer recovery specialists, and online support groups) could prove to be a powerful means of promoting SU behavior change among youth.

Digital health interventions that combine motivation-enhancement and accurate peer norms may be especially well-suited to treat CINI youth in justice settings because they are accessible, customizable, portable, and show evidence of reducing cannabis use [34-39]. Among adolescents who receive brief computerized SU treatments [40], effects appear to fade over time [41], suggesting that adjunct treatments may be needed to help sustain gains. Smartphone apps are an especially appealing means of providing youth ongoing, on-demand care as an adjunct to concurrent or recently completed services (eg, continuing care and recovery support). More than 91% of US youth have their own smartphones [42], which they use for everything from schoolwork to seeking support from peers during challenging times [43]. Overwhelming evidence shows that teens find mobile apps to be a confidential and acceptable intervention platform, one with which they engage much more frequently than adults [44,45].

Most publicly available smartphone apps for SU do not incorporate evidence-based practices or resources related to SU treatment [46,47], but the availability of empirically-supported apps is increasing. A 2021 systematic review of smartphone apps for SU [48] identified 2 adjunctive smartphone apps specifically for youth with cannabis use. Although studies of both apps established the feasibility and potential efficacy of adjunctive smartphone apps for cannabis use [44,49-52], these apps were built for older, heavy cannabis users who were also receiving in-person therapy from skilled clinicians, and neither was designed for youth in low-resource settings like the juvenile justice system. Critically, no existing apps built to help youth reduce cannabis or other SU have any peer networking features. In other words, no currently available app designed for youth who use substances leverages interpersonal mechanisms of change. This represents a significant limitation to existing apps as well as an opportunity to advance the field. Therefore, new research is needed to establish the feasibility, acceptability, and effectiveness of an adjunct smartphone app targeting intrapersonal and interpersonal mechanisms to reduce cannabis use among high-risk teens in low-resource settings.

Study Aims and Hypotheses

This protocol aims to develop, feasibility test, and evaluate the preliminary effectiveness of Teen Empowerment through Computerized Health (TECH) app as an adjunct to standard family court services (ie, treatment-as-usual [TAU]) to reduce cannabis and other SU among CINI youth. We conceptualize TECH as a peer-facilitated adjunct to brief, motivation-enhancing treatments for cannabis use—a virtual

community where CINI youth can give and receive ongoing support around SU-related behavior change. Study activities will address three scientific aims and two hypotheses:

- Aim 1: develop a user-driven smartphone app, TECH, to reduce cannabis use among CINI youth
- Aim 2: examine the feasibility and acceptability of TECH
  - *Hypothesis 1*: TECH will be feasible for and acceptable to CINI youth
- Aim 3: test the preliminary effectiveness of TECH as an adjunct to TAU (ie, TAU-only vs TAU+TECH) on cannabis and other SU (primary outcome)
  - *Hypothesis 2*: CINI youth who receive TAU+TECH will demonstrate greater reductions in cannabis and other SU-related outcomes relative to TAU-only CINI youth. We will also explore possible intervention effects on a secondary outcome (ie, delinquent behavior) and multiple putative mediators (ie, intrapersonal and interpersonal mechanisms of change)
  - *Exploratory Hypothesis 1*: youth in the TAU+TECH group will show greater change on the secondary outcome and putative mediators

Methods

Study Overview

The study aims will be achieved across 3 distinct study phases. Phase I will consist of qualitative interviews with CINI youth (n=14-18), their caregivers (n=6-8), and behavioral health app developers (n=6-8) to determine how, why, and when CINI youth would most prefer to engage with a smartphone app to reduce cannabis use. Qualitative results will guide decision-making on the key features and overall design of the TECH app, in partnership with our software developer. Phase II will beta test the TECH app with CINI youth (n=10) to guide decision-making to refine the app. Phase III will encompass a pilot RCT with CINI youth (n=60) to compare the treatment effects of TAU-only and TAU+TECH.

Theoretical Frameworks

The study will be guided by 2 frameworks. First, the research approach will follow principles of user-centered design [53] for accelerating digital health research [54]. In this context, the user-centered design approach will engage key stakeholders in all aspects of TECH development and refinement, to ensure the resulting app is tailored to the specific needs and preferences of CINI youth who use cannabis. Second, the Behavior Intervention Technology (BIT) model will guide the research team's efforts to build the TECH app prototype [55]. The BIT model helps researchers clarify why a digital health product is needed (eg, clinical and usage aims for the TECH app), how it is expected to change users' behavior (eg, interpersonal and intrapersonal mechanisms), and what technical components will achieve the desired effects (eg, elements, characteristics, and workflow of the TECH app features).

Study Setting

This research will be done in partnership with a state-wide family court in the Northeastern United States. The partner family court has jurisdiction over all delinquent, wayward, dependent, psychiatrically disordered, and diverted youth aged <18 years.

Standard of Care Services

All CINI youth who participate in any phase of this study will receive standard services from the family court, including juvenile intake and any indicated referrals. In the juvenile intake, court staff workers screen youth for mental health and SU problems. Intake workers then make referrals; options for continued system involvement include formal delinquency charges, diversion programs (eg, mandated community service, mental health services, or SU services), ongoing supervision, or probation.

In 2017, the family court began recommending the eCHECKUP-TO-GO (ECTG) program for cannabis [56] as their standard of care for CINI youth who present with a history of cannabis use. ECTG is a computerized single-session, motivation-enhancing intervention that generates personalized behavior change goals related to cannabis and SU across personal, academic, and social domains and has shown positive outcomes among cannabis-using [37,57] and abstinent [58] college students. All CINI youth who participate in any phase of this study will receive the computerized ECTG program for cannabis use.

Inclusion Criteria

Three types of participants will be recruited: CINI youth, their caregivers, and behavioral app developers. Notably, caregivers of CINI youth and CINI youth will participate independently in the study (ie, we are not enrolling caregiver+CINI youth dyads). CINI youth will be recruited for all study phases and must be (1) aged 14-18 years, (2) able to speak and read English, (3) have access to a smartphone, (4) report past-year cannabis use and screen positive on either the CRAFFT (car, relax, alone, forget, friends, trouble) [59] or the Massachusetts Youth Screening Instrument-Version 2 (MAYSI-2 [60])—the family court's preferred evidence-based SU screening measures, and (5) have a caregiver willing and able to provide consent for their participation. In phases II and III, eligible CINI youth must also be (6) willing to adhere to the TECH app's user safety agreement during the consent process. Importantly, caregivers providing consent for CINI youth to participate must be able to complete the consent process in English or Spanish. Caregivers and behavioral health app developers will only participate in phase I interviews. Caregiver participants must (1) speak in English and (2) be a legal guardian and primary caregiver of an eligible CINI youth. Behavioral health app developers must (1) speak and read in English and (2) be lead developers of a research-based smartphone app targeting behavioral health outcomes. Exclusion criteria are limited to enhance generalizability and only include conditions that preclude CINI youth from actively participating in an interview or using the proposed app (eg, psychosis, cognitive impairment, no smartphone access, or non-English speaking).

Recruitment

CINI youth (phase I, n=14-18; phase II, n=10; and phase III, n=60; in total, n=84-88) will be recruited through the family

court's juvenile intake department. Court intake staff will receive a brief training on the study objectives, eligibility criteria, and referral procedures. Upon screening a study-eligible adolescent, intake staff will describe whichever phase of the study is actively recruiting participants; CINI youth aged >18 years and caregivers of interested minors will sign a consent-to-contact form allowing contact with the investigators. The phase I consent-to-contact form allows families to indicate whether the CINI youth, the caregiver, or both would be interested in interviewing. When recruiting CINI youth who are minors, the research team will initiate contact with their caregivers.

Caregivers of CINI youth (phase I only; n=6-8) will be recruited through the family court's juvenile intake department via the same methods used to recruit CINI youth. Following receipt of the consent-to-contact form, the principal investigator will contact eligible caregivers to describe the study interviews.

Behavioral health app developers (phase I only; n=6-8) will be identified via searches of the literature and repositories of federally funded grants to ensure they were a lead developer of a research-based smartphone app targeting behavioral health outcomes. Given TECH's focus, preference will be given to developers of youth- or SU-focused apps. The principal investigator will contact eligible individuals to describe the study and invite them to reply if interested.

Compensation

All participants will receive financial compensation for their participation in the study. Phase I participants will receive a US $50 gift card upon completion of the study interview. Phase II and III CINI youth will receive up to US $90 and US $200 in gift cards, respectively.

Retention

To promote retention of phase II and III CINI youth at follow-up, we will collect multiple sources of contact information at baseline, including social media handles [61]. Participants will also receive gift cards escalating in value for each completed assessment (ie, US $40 at baseline, US $50 at 1 and 3 months, and US $60 at 6 months).

Informed Consent

Overview

All participants will be fully informed of the purposes and procedures of the study, both verbally and through a written description in assent and consent forms. Individuals aged >18 years (ie, caregivers, behavioral health app developers, and CINI youth who are legally adults) will provide informed consent before study enrollment; for minor CINI youth, we will obtain consent from their legal guardian and assent from the youth. The informed consent and assent processes will be conducted in English by a research team member; a bilingual, trained research assistant will complete the process with any Spanish-speaking caregivers providing consent for minor CINI youth. This process will occur either in-person, by phone, or via videoconference, depending on the potential participant's preference.

Potential participants and caregivers providing consent for CINI youth will be reminded that their participation is strictly voluntary, they can refuse to answer any questions, and they may withdraw from the study at any time. Potential adolescent and caregiver participants will also be assured that they may decide to participate, not participate, or withdraw from the study without fear of penalty from the family court. To ensure adequate comprehension, a research team member will ask potential participants to describe the purpose of the study and the basic study procedures. If the potential participant does not appear to understand the study, the research team member will read aloud key portions of the consent or assent form and provide brief verbal summaries of each section of the consent or assent form. Once adequate understanding has been demonstrated, potential participants will be asked to electronically sign their consent or assent form via a Qualtrics survey. Participants will be sent a copy of their completed form, for their records.

User Safety Agreement

As part of the consent/assent process, phase II and III CINI youth will also be required to agree to adhere to a user safety agreement describing appropriate use of the TECH app. Digital and social media products often require users to agree to adhere to specific terms regarding their use; researchers have begun to include such agreements in their technology-based studies [62]. The TECH app user safety agreement will specify guidelines for appropriate peer interactions in the digital research context (eg, maintaining privacy and anonymity, respect for others, relevant content, the role of the moderator, and protecting the safety of all users) and outline possible consequences for violating the terms of use (eg, moderators will delete posts and remove the participants' access to interactive app features).

Sample Size Feasibility

We aim to recruit 84 to 88 CINI youths throughout the 5-year, multiphase protocol. Approximately 1500 youths are referred to the family court annually. In a recent study of a sample of youth screened by the family court, 173 (50%) of 348 youths screened positive for problematic cannabis use [63], indicating that the recruitment pool is sufficiently robust to support the proposed recruitment targets.

Phase I: Formative Research and TECH Prototype Development

Overview

Semistructured interviews will be conducted with CINI youth (n=14-18), their caregivers (n=6-8), and behavioral health app developers (n=6-8). The primary goal is to gather data from CINI youth, to guide decision-making and prioritization of the TECH app features. Broadly, the interviewers will uncover potential barriers to app development or implementation.

Interview Guides

Separate interview guides will be used for each type of participant, but all will explore the following BIT [55] model dimensions: (1) clinical aims, to learn what types of intervention goals would be most relevant and appealing to CINI youth (eg, avoid arrest and decrease cannabis use); (2) usage aims, to

understand how CINI youth use behavioral health apps; (3) behavior change strategies, to identify which theory-driven approaches to behavior change appeal most to CINI youth (eg, goal-setting, monitoring, education, and peer support or advice); (4) elements, to identify those app features most relevant and appealing to CINI youth to address their cannabis use (eg, notifications, text messages, and newsfeed); (5) characteristics, to establish CINI youths' preferences around personalization, complexity, types of media employed, and general esthetics (eg, settings, design, and individualization); and (6) workflow, to understand CINI youths' preferred conditions for engaging with an app (eg, time- vs goal-based) to define a user-driven workflow.

Interviews with behavioral health app developers will focus on their experiences developing, implementing, and refining smartphone apps for behavioral health. Caregiver interviews will capture thoughts about their teen's potential use of a smartphone app for cannabis use, potential barriers to their teen's ongoing use of the app, and the perceived feasibility and acceptability of digital health interventions for high-risk teens.

CINI youth participants will begin their interviews upon completing the computerized ECTG program for cannabis [56] (ie, TAU). Youth will describe their access to and use of smartphones, how they engage with peers digitally, and share their perceptions of and prior experience with behavior change apps, to capture the feasibility and acceptability of digital behavioral health interventions for CINI youth. Interviews will examine youths' concerns about using smartphone-based interventions in juvenile justice contexts, including security and privacy, court staff access to the app, and perceived risks of connecting digitally with other CINI youth. The bulk of the interview will be dedicated to BIT [55] domains to inform how the proposed TECH app could best support CINI youth enacting ECTG-recommended changes in cannabis use and other behaviors. Youth will rate each app feature on a 10-point Likert scale of importance and name their top 3 preferred features, to ensure the TECH app features align with CINI youths' preferences.

Interview Procedures

All interviews will be conducted and audio recorded by the principal investigator (SH). Behavioral health app developer and caregiver interviews will take 30 to 60 minutes, whereas CINI youth interviews will take 60 to 90 minutes. Participants will complete any quantitative measures before their interviews.

Interview Quantitative Measures

All participants will be asked to provide basic demographic information (eg, age, biological sex, and education level).

CINI youth will be asked to describe their recent history of behavioral health services via the Child and Adolescent Services Assessment [64,65]. The Marijuana Use Questionnaire [66,67] will capture youths' preferred type (eg, plant or concentrate) and mode (eg, vape, smoke, and edible) of cannabis use. Youth will also report on their perceptions of normative cannabis use among their peers. Following standard practices [33,68], three types of norms will be captured for two reference groups (ie, your close friends and other teens your age): (1) descriptive norms ask participants to estimate, within each reference group, the percentage of youth who have ever used cannabis and who use cannabis regularly, along with their past-month and past-year frequency of use, on a 9-point Likert scale (0=*never* to 8=*every day*); (2) injunctive norms, which assess each reference group's approval of youth who have never used cannabis, used once or twice, use occasionally, and use regularly, on a 7-point Likert scale (1=*strongly disapprove* to 7=*strongly approve*); and (3) subjective norms, which measure each reference group's perceived approval of the participant's past 30-day cannabis use on a 7-point Likert scale (1=*strongly disapprove* to 7=*strongly approve*).

Interview Analysis

All phase I quantitative measures will be analyzed using percentages or means and SDs. These data will offer descriptive insights into each group of interviewees.

Interviews will be transcribed and checked by 2 research team members to identify any discrepancies before the analysis. Data from each type of participant will be analyzed separately using the following directed content analysis [69] approach. The analysis will begin with an a priori coding framework derived from BIT model constructs but will allow for the emergence of unanticipated codes. First, 2 team members will independently code the same transcript; the research team will convene to discuss identified codes to refine BIT codes and reach consensus regarding emergent codes. This process will be completed for each type of participant, resulting in a series of 3 codebooks. All transcripts will then be independently coded by 2 team members. The team will reconvene to discuss divergent codes with the goal of achieving 100% consensus; codes that remain discrepant following this discussion will be decided by a third, independent coder. Coding will occur on a rolling basis, to allow for periodic assessment of data saturation [70,71].

Building the TECH App

Anticipated Components

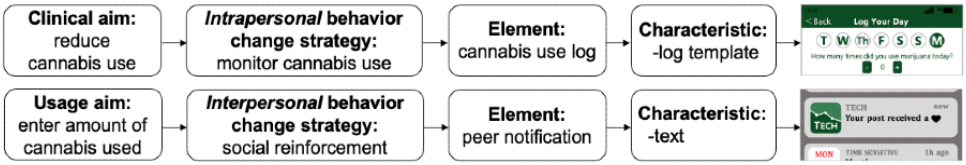
Although our final decision-making will be driven by phase I formative work, we anticipate including several theory-driven components in the TECH app. Specifically, in-app goal-setting and behavior tracking (eg, promoting self-efficacy) as well as delivery of accurate SU information (eg, helping to change attitudes and beliefs about SU) will leverage intrapersonal mechanisms of change. Interpersonal mechanisms of change will be targeted through a newsfeed for youth to post their progress (eg, modeling and social norms) and receive peer support (eg, social reinforcement). To protect against deviant peer influence [72], the app will be programmed to prevent negative (eg, *thumbs-down*) or covert (eg, direct messaging) features, and the moderator will monitor in-app activity daily.

App Development

The TECH app prototype will be built in partnership with Mooseworks Software LLC, a development company that builds apps from a software library of preprogrammed features, including, but not limited to, networking forums, community newsfeed, push notifications, tracking logs, and sharable posts. Guided by phase I formative work, the research team will engage in a series of meetings to select core elements and features of

TECH in consultation with the software developer. Following the BIT model [55] (Figure 1), this process will match theory-driven intrapersonal (ie, self-efficacy, attitudes, and beliefs about SU) and interpersonal (ie, modeling, social norms, and social reinforcement) mechanisms of behavior change with CINI youth's preferred app features. The research team will test the prototype's functionality before launching phase II. In addition, the research team will populate the app with content solicited during phase I interviews with CINI youth to ensure the app is engaging to initial participants.

Figure 1. Applying the behavior information technology model [55] to the proposed Teen Empowerment through Computerized Health app.



Phase II: Beta Testing the TECH Prototype

Overview

In the next phase of the study, CINI youth (n=10) will beta test the TECH prototype. Beta testing enables users to identify and refine unclear components and troubleshoot software bugs. Methodologically, beta testing helps researchers reduce dissatisfaction, assess acceptability, and test and refine key study procedures (eg, app introduction, user safety agreement adherence, and metadata collection) before an RCT [73-76].

Beta-Testing Procedures

CINI youth recruitment, caregiver consent, and assent procedures will be identical to phase I, with one exception. Phase II will use a phased approach to recruitment, wherein the research team will wait to schedule baseline assessments until receiving consent-to-contact forms for 4 CINI youths. This will help ensure sufficient in-app peer presence for initial participants.

Enrolled CINI youth will attend a study session to complete baseline measures, the computerized ECTG program for cannabis use [56] and receive an introduction to the TECH prototype. A research team member will help participants download the app, create an anonymous account and log in, review the TECH app user safety agreement, and learn how to use each app feature. The research team member will help teenagers enter at least two cannabis-related behavior change goals derived from the ECTG program [56]. The TECH app introductions will be audio recorded to assess researcher adherence to the introduction protocols. The research team will monitor the TECH app activity daily to ensure prompt handling of any app functionality problems or participant concerns.

CINI youth will be asked to use the TECH app prototype for 1 month, completing a brief measure of acceptability each week. At the end of the beta-testing period, participants will complete a series of post–beta-testing measures of feasibility and acceptability. The TECH app metadata will be examined on an ongoing basis to monitor use patterns, note which app features are underused, and identify opportunities for targeted intervention to promote increased use. To test urine drug screen administration procedures for phase III's pilot RCT, phase II CINI youth will be asked to complete a Clinical Laboratory Improvement Amendments–waived urine drug screen with eight panels (ie, cannabis, amphetamines, methylenedioxy-methylamphetamine, opiates, oxycodone, methamphetamines, cocaine, and benzodiazepines) during their post–beta-testing appointment. For in-person appointments, a research team member will request a urine sample and administer the dip test. Youth who complete appointments virtually will be mailed necessary screening materials before their appointment and receive step-by-step instructions to self-administer the dip test. Once the results are recorded, all the specimens and testing materials will be destroyed.

Beta-Testing Measures

Baseline

Measures will be identical to those used in phase II.

Feasibility

The feasibility of the TECH app prototype will be assessed using two primary indicators: (1) percentage of CINI youth willing to participate (recruited vs enrolled) with a target of 50% and (2) TECH usability, measured by automatically collected metadata. Available metadata will depend on design decisions following phase I formative work (eg, forum or newsfeed needed to make a post). Ideal app use would involve at least half of phase II participants using the app 2 or more times, spending an average of 30 seconds in the app when using it, posting at least once, and liking a peer's content at least once.

Acceptability

We will measure the acceptability of the TECH app via four quantitative indicators: (1) percentage of CINI youth who withdraw from the beta-testing process, with a target of less than 20%; (2) weekly administration of the user version of the Mobile Application Rating Scale [77], an established, 20-item measure of an app's engagement, functionality, esthetics, and information quality, with a targeted app quality mean score of 2.5, or higher (out of 5); (3) the Consumer Satisfaction Questionnaire [78], with a target of at least 70% of CINI youth indicating that they are satisfied and would recommend TECH to others; and (4) a structured, open-ended questionnaire to solicit feedback about the TECH prototype, including changes to content or features that could improve usability and engagement, administered at the end of the 1-month beta-testing period.

Beta-Testing Analysis

Due to the small sample size, phase II measures will only be used to offer insight into the feasibility of study procedures and guide the TECH app refinement. Data will be presented in

aggregate and analyzed using means, SDs, or similar types of statistical analyses (eg, percentages).

Refining TECH

Similar to the development of the prototype, the TECH app refinement will follow the BIT model [55] and incorporate feedback provided by CINI youth during beta testing. The research team will work with our software developer to troubleshoot usability issues, consider modifications to address acceptability concerns, and finalize the TECH app.

Phase III: Pilot RCT of the Final TECH App

Overview

In the final phase of the study, we will conduct a pilot RCT with CINI youth (n=60) to examine the feasibility, acceptability, and preliminary effectiveness of the TECH app. CINI youth recruitment, caregiver consent, assent, and baseline appointment procedures will be identical to those in phase II. Table 1 depicts the timing of phase III activities.

Table 1. Schedule of enrollment, interventions, and assessments for phase III's pilot randomized controlled trial.

	Enrollment, -t_1	Allocation	Postallocation			Closeout, t_x[a]
			t_1[b]	t_2[c]	t_3[d]	
Enrollment						
Screening	✓					
Informed consent or assent	✓					
User Safety Agreement	✓					
Baseline	✓					
Randomization		✓				
Interventions						
TAU[e] (all study participants)	✓					
TECH[f] app (only TAU+TECH participants)			✓	✓	✓	
Phase III measures for all participants						
Demographics	✓					
Child and Adolescent Services Assessment [64,65]	✓					
Marijuana Use Questionnaire [66,67]	✓					
Timeline Followback Interview [79-81]	✓		✓	✓	✓	
Urine Drug Screens	✓		✓	✓	✓	
Selected items on Global Appraisal of Individual Needs [82]	✓		✓	✓	✓	
Selected items on cannabis use attitudes [83]	✓		✓	✓	✓	
Marijuana Adolescent Problem Inventory [84]	✓		✓	✓	✓	
Marijuana Effect Expectancy Questionnaire [85]	✓		✓	✓	✓	
Cannabis Refusal Self-Efficacy Questionnaire [86,87]	✓		✓	✓	✓	
Adapted Readiness to Change Questionnaire [88]	✓		✓	✓	✓	
Descriptive, Injunctive, and Subjective Peer Norms [33]	✓		✓	✓	✓	
Phase III measures completed by TAU+TECH participants						
Mobile Application Rating Scale (user version) [77]			✓	✓	✓	
Adapted Consumer Satisfaction Questionnaire [78]			✓	✓	✓	
Open-ended TECH app feedback questionnaire			✓	✓	✓	
Other phase III measures not completed by participants						
Percent of CINI[g] youth recruited vs enrolled						✓
Percent of CINI youth who withdraw from the study						✓
TECH app usability metadata						✓

[a]Conclusion of study.

[b]1 month after baseline.

[c]3 months after baseline.

[d]6 months after baseline.

[e]TAU: treatment-as-usual.

[f]TECH: Teen Empowerment through Computerized Health.

[g]CINI: court-involved, nonincarcerated.

Pilot RCT Randomization

Upon completing the computerized ECTG program for cannabis use [56] at the baseline appointment, participants will immediately be randomized to 1 of the 2 treatment conditions. Participants will be randomized using an urn randomization spreadsheet [89], balancing for biological sex and frequency of cannabis use. Youth will be assigned at a ratio of 2 TAU+TECH to 1 TAU-only to ensure sufficient peer presence on the TECH app at any given time.

Pilot RCT Treatment Conditions

TAU-Only

The computerized ECTG program for cannabis use [56] and the family court's standard care services will serve as the TAU-only condition in the pilot RCT.

TAU+TECH

Participants in the TAU+TECH arm will receive all TAU components and access to the TECH app. Immediately after randomization, a research team member will introduce participants to the TECH app using the same procedures used in phase II. TAU+TECH participants will be asked to use the TECH app for 6 months.

Pilot RCT Measures

Baseline, preliminary effectiveness, and putative mediator outcome measures will be administered to all phase III participants at all time points. Feasibility and acceptability measures specific to the TECH app will only be administered to participants randomized to TAU+TECH at the 1-, 3-, and 6-month follow-up appointments.

Baseline

All phase II baseline measures and phase III measures of preliminary effectiveness, secondary effectiveness, and putative mediators will be administered at baseline.

Preliminary Effectiveness Outcomes

These measures will be completed by all phase III participants at all assessments and will assess the preliminary effectiveness of TAU+TECH relative to TAU-only. The primary outcome measure will be CINI youths' self-reported cannabis and other SU, as measured by the Timeline Followback Interview [79-81]. Youth will report their total days of cannabis and other SU, abstinence, high-volume use, co-use, and cannabis grams per day in the past 30 days. Youth self-reported SU will be corroborated by urine drug screens using the same procedures outlined in phase II unless adjustments are deemed necessary.

Secondary Effectiveness Outcomes

We will collect data on delinquent behavior to account for its bidirectional relationship with cannabis and other SU. Furthermore, 9 items adapted from the Global Appraisal of Individual Needs-Core [82] will assess past 30-day delinquency and legal system involvement.

Putative Mediators

We will collect six indicators of putative intrapersonal mechanisms of change: (1) attitudes toward cannabis use will be assessed by two items [83] that capture the perceived acceptability of personal cannabis use on a Likert scale, one positively framed ("Is it ok...?") and one negatively framed ("How wrong is...?"); (2) the Marijuana Adolescent Problem Inventory [84], which includes 23 items on perceived cannabis-related consequences, rated on a 0-4 Likert scale; (3) The 6-item Marijuana Effect Expectancy Questionnaire-Brief [85], which gauges positive and negative expectations of cannabis use [90]; (4) the Cannabis Refusal Self-Efficacy Questionnaire [86,87], which uses 14 items measuring self-confidence to resist or refuse cannabis; (5) an adapted Readiness to Change Questionnaire [88], which will assess motivation to reduce cannabis use with 12-items; and (6) descriptive, injunctive, and subjective peer norms of cannabis use [33,68] as collected at baseline in prior phases.

Feasibility and Acceptability

These measures will be identical to those used in phase II; however, only CINI youth randomized to the TAU+TECH condition will be asked to complete measures regarding their experiences using the TECH app.

Pilot RCT Statistical Analysis

We will conduct preliminary analyses on key variables to examine distributional properties, identify outliers, and transform variables as needed. The conditions (ie, TAU-only vs TAU+TECH) will be compared on baseline demographic variables; differences will be controlled in subsequent analyses. Data analysis will follow intent-to-treat principles [91] and use multiple imputation methods [92] in the event of unplanned missing data. Due to the small sample size, we will be underpowered for significance testing. Instead, our goal is to obtain data on variable distributions, reliability, and effect size estimates for a future large-scale trial. For the same reason, we will not attempt to test for mediation but instead test associations with possible mediators. Findings will inform which variables to evaluate in a future trial.

We anticipate 80% retention [93] of phase III CINI youth (n=60), leaving 48 participants with complete data. This sample will be sufficient to determine the feasibility and acceptability outcomes. Primary preliminary effectiveness outcomes include cannabis and other SU. Putative mediators include intrapersonal and interpersonal mechanisms of change; delinquent behavior is the sole secondary outcome. Brief motivation-enhancing treatments produce short-term decreases in cannabis and SU [40,41], but technological adjuncts can help maintain reductions in SU for up to 6 months [94]. Thus, we expect CINI youth to report the highest rates of cannabis and other SU at baseline, with small to moderate decreases at 1 month that maintain at 6 months.

We will conduct a series of repeated measures analyses of covariance to predict outcomes by condition, test putative mediators, and control for baseline differences. We will also analyze dosage effects on each dependent variable and calculate partial eta squared (η^2) to estimate proportions of variance associated with TECH. Despite effect size stability concerns in small samples [95], detection of small to moderate treatment effects on our primary outcomes alongside strong evidence of feasibility and acceptability could indicate preliminary efficacy of the TECH app as an adjunct to TAU. If such evidence is found, effect size estimates will be used to determine the sample size for a future, fully-powered RCT (eg, for power of 0.8, a small effect size [Cohen d=0.3] would require n=175 per group, whereas a medium effect size [Cohen d=0.5] would require n=64).

Ethical Oversight

The proposed study activities will be reviewed and approved by the institutional review board at Brown University. The

research team has extensive experience conducting intervention research with CINI youth who use cannabis and addressing emergent safety concerns. Study activities will include several safeguards to ensure that the proposed research presents minimal risks of psychological discomfort, coercion, legal, and loss of privacy or confidentiality to CINI youth participants. Psychological discomfort will be minimized by (1) emphasizing that youth will not be obligated to answer distressing questions and can withdraw from the study at their discretion; (2) avoiding direct questions about parental neglect, abuse, or suicidal ideation; and (3) conducting a thorough safety assessment if youth spontaneously disclose these concerns. The risk of coercion will be minimized by (1) reminding CINI youth their participation is voluntary; (2) ensuring that they receive multiple referral options for substance misuse from their court intake worker, of which the study will be one; and (3) offering compensation rates consistent with previous studies and commensurate with the level of effort required. Legal risks will be minimized by obtaining a certificate of confidentiality from the National Institutes of Health and explaining its limitations. Potential loss of privacy or confidentiality will be minimized by (1) using electronic consent and assent forms instead of needing to store and/or transport hard copies from the family court to the university; (2) deidentifying participant data; (3) storing all electronic data, including audio recordings, on a password-protected, secure shared drive; (4) allowing only the principal investigator, coinvestigators, and research team members to access data files; and (5) destroying all audio recordings and any other identifiable information within 6 months of the study's conclusion.

Clinical Trial Registration

Our pilot RCT will be submitted for registration to the ClinicalTrials.gov website within 21 days of enrolling the first phase III participant (anticipated in summer 2023).

Results

The study received institutional review board approval in August 2019 and was funded in March 2020. Participant recruitment began in September 2020, and phase I was completed in December 2021. A total of 11 behavioral health app developers, 8 caregivers, and 14 CINI youth completed the phase I semistructured interviews. We interviewed 3 more behavioral health app developers than initially planned to achieve data saturation. Phase I data analysis and the TECH app development were actively underway at the time of this submission. Data collection is scheduled to begin in spring 2022 and summer 2023 for phases II and III, respectively. Phase III analyses are expected to be completed in spring 2025.

Discussion

Overview

This multiphase research aims to develop and pilot test an adjunctive smartphone app to reduce cannabis and other SU among CINI youth. Given the high prevalence of cannabis use in this population [2] and well-documented associations with health and legal consequences later in life [3-6], preventive and early interventions are vital to disrupt the cycle of SU and criminal justice involvement among CINI youth. We believe that the proposed TECH app will be the first designed specifically for CINI youth, who represent nearly 75% of all youth involved in the juvenile justice system [1]. Although a handful of smartphone apps have been developed to support adolescents in treatment for cannabis use [44,49-52], TECH will be one of the first apps designed for youth who are not currently receiving traditional SU treatment. Treatment rates are generally low across the juvenile justice system: only 28% of detained youth [12] and <30% of CINI youth [13] in need of SU services receive any treatment.

Brief, computerized interventions such as the ECTG program [56] can deliver evidence-based content with high fidelity, making them ideal for self-administration or delivery by a workforce without behavioral health training. However, similar to most brief approaches, gains achieved through brief computerized interventions can fade over time, including in samples of adolescents who use substances [40,41]. Digital adjuncts, like the TECH app, could provide a light touch of ongoing support to help bolster treatment gains. Although this app will be tailored to address the unique needs and preferences of CINI youth who use substances, evidence supporting the feasibility and acceptability of our multimodal digital approach could have broader implications. Fully digital approaches represent a promising, if untested, solution for high-need, low-resource settings, especially those with the existing infrastructure to identify individuals at risk but lacking the workforce, capacity, or overarching mission to deliver behavioral health services, such as juvenile-court settings. Finally, despite overwhelming evidence indicating that peers play a key role in the development of adolescent SU, this study marks one of the first efforts to leverage interpersonal mechanisms of change through a peer networking forum. The proposed platform will connect youth in a monitored, virtual community that limits risky peer behavior so that they can engage and support each other's cannabis-related behavior changes.

Limitations

We note the following limitations to the proposed work. First, although caregiver- and family-based interventions for adolescent SU are well-established [24], this study will intervene solely on CINI youth. If TECH shows promise as an adjunct to brief computerized treatments (ie, the ECTG program for cannabis use [56]), an appropriate next step could be to deliver TECH alongside family-focused digital approaches [96]. Second, all CINI youth will receive TAU, which includes the computerized ECTG program for cannabis use [56] that has been shown to reduce cannabis use among young adults [37,57,58]. This ensures that CINI youth in-need would not be denied standard services. Although this may impact our ability to detect preliminary intervention effects specific to the TECH app, it is also the only way to assess whether the adjunctive TECH app can help sustain cannabis and other SU behavior change among CINI youth. Third, CINI youth and caregivers may be hesitant to disclose youth SU and delinquent behavior in qualitative interviews and surveys. To address this issue, our consent procedures will explain our protocols to maintain

confidentiality and how study data will not lead to reprisals from family courts. Finally, although sufficient to achieve the proposed aims, the small sample in the pilot RCT may provide limited opportunities for in-app peer engagement, which may hinder interpersonal mechanisms of behavior change. We will take several steps to maximize peer presence on the TECH app, including randomizing youth 2:1 to receive the TECH app, phasing recruitment so that multiple participants receive the app concurrently, and populating the app with content collected from previous participants.

Acknowledgments

This study was funded by a National Institute of Drug Abuse grant (K23DA048062) awarded to the first author (SAH). The funder was not involved in any aspect of the study or manuscript publication.

Authors' Contributions

SAH, JG, and DP wrote the initial draft of this manuscript. AS, MAC, NPB, and SJB contributed to the initial grant proposal and assisted in the revision of this manuscript. All authors approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review reports.
[PDF File (Adobe PDF File), 181 KB - resprot_v11i3e35402_app1.pdf]

References

1. Hockenberry S. Delinquency cases in juvenile court, 2018. U.S. Department of Justice: Office of Justice Program. 2020. URL: https://ojjdp.ojp.gov/publications/delinquency-cases-in-juvenile-court-2018.pdf [accessed 2022-03-01]
2. Criminal neglect: substance abuse, juvenile justice and the children left behind. National Center on Addiction and Substance Abuse at Columbia University. 2004. URL: https://www.ojp.gov/ncjrs/virtual-library/abstracts/criminal-neglect-substance-abuse-juvenile-justice-and-children-left [accessed 2017-12-28]
3. Brady SS, Tschann JM, Pasch LA, Flores E, Ozer EJ. Violence involvement, substance use, and sexual activity among Mexican-American and European-American adolescents. J Adolesc Health 2008;43(3):285-295 [FREE Full text] [doi: 10.1016/j.jadohealth.2008.02.007] [Medline: 18710684]
4. Fergusson DM, Horwood LJ, Swain-Campbell N. Cannabis use and psychosocial adjustment in adolescence and young adulthood. Addiction 2002;97(9):1123-1135. [doi: 10.1046/j.1360-0443.2002.00103.x] [Medline: 12199828]
5. Fergusson DM, Swain-Campbell NR, Horwood LJ. Deviant peer affiliations, crime and substance use: a fixed effects regression analysis. J Abnorm Child Psychol 2002;30(4):419-430. [doi: 10.1023/a:1015774125952] [Medline: 12108769]
6. Winters KC, Lee CY. Likelihood of developing an alcohol and cannabis use disorder during youth: association with recent use and age. Drug Alcohol Depend 2008;92(1-3):239-247 [FREE Full text] [doi: 10.1016/j.drugalcdep.2007.08.005] [Medline: 17888588]
7. Tolou-Shams M, Rizzo CJ, Conrad SM, Johnson S, Oliveira C, Brown LK. Predictors of detention among juveniles referred for a court clinic forensic evaluation. J Am Acad Psychiatry Law 2014;42(1):56-65. [Medline: 24618520]
8. Chassin L, Knight G, Vargas-Chanes D, Losoya SH, Naranjo D. Substance use treatment outcomes in a sample of male serious juvenile offenders. J Subst Abuse Treat 2009;36(2):183-194 [FREE Full text] [doi: 10.1016/j.jsat.2008.06.001] [Medline: 18657942]
9. Adelman WP. The need to focus research on adolescent cannabis use interventions. Pediatrics 2020;146(4):e2020012336. [doi: 10.1542/peds.2020-012336] [Medline: 32928987]
10. Young DW, Dembo R, Henderson CE. A national survey of substance abuse treatment for juvenile offenders. J Subst Abuse Treat 2007;32(3):255-266 [FREE Full text] [doi: 10.1016/j.jsat.2006.12.018] [Medline: 17383550]
11. Bath E, Tolou-Shams M, Farabee D. Mobile health (mHealth): building the case for adapting emerging technologies for justice-involved youth. J Am Acad Child Adolesc Psychiatry 2018;57(12):903-905 [FREE Full text] [doi: 10.1016/j.jaac.2018.08.012] [Medline: 30522732]
12. White LM, Aalsma MC, Salyers MP, Hershberger AR, Anderson VR, Schwartz K, et al. Behavioral health service utilization among detained adolescents: a meta-analysis of prevalence and potential moderators. J Adolesc Health 2019;64(6):700-708. [doi: 10.1016/j.jadohealth.2019.02.010] [Medline: 31122506]
13. Dennis ML, Smith CN, Belenko S, Knight D, McReynolds L, Rowan G, et al. Operationalizing a behavioral health services cascade of care model: lessons learned from a 33-site implementation in juvenile justice community supervision. Fed Probat 2019;83(2):52-64 [FREE Full text] [Medline: 34366438]

14. Petraitis J, Flay BR, Miller TQ. Reviewing theories of adolescent substance use: organizing pieces in the puzzle. Psychol Bull 1995;117(1):67-86. [doi: 10.1037/0033-2909.117.1.67] [Medline: 7870864]
15. Feinstein EC, Richter L, Foster SE. Addressing the critical health problem of adolescent substance use through health care, research, and public policy. J Adolesc Health 2012;50(5):431-436. [doi: 10.1016/j.jadohealth.2011.12.033] [Medline: 22525104]
16. Handley ED, Chassin L. Intergenerational transmission of alcohol expectancies in a high-risk sample. J Stud Alcohol Drugs 2009;70(5):675-682 [FREE Full text] [doi: 10.15288/jsad.2009.70.675] [Medline: 19737491]
17. Brown SA, Creamer VA, Stetson BA. Adolescent alcohol expectancies in relation to personal and parental drinking patterns. J Abnorm Psychol 1987;96(2):117-121. [doi: 10.1037//0021-843x.96.2.117] [Medline: 3584659]
18. Ebersole DS, Miller-Day M, Raup-Krieger J. Do actions speak louder than words?: adolescent interpretations of parental substance use. J Fam Commun 2014;14(4):328-351 [FREE Full text] [doi: 10.1080/15267431.2014.945699] [Medline: 25285048]
19. Brook JS, Kessler RC, Cohen P. The onset of marijuana use from preadolescence and early adolescence to young adulthood. Dev Psychopathol 1999;11(4):901-914. [doi: 10.1017/s0954579499002370] [Medline: 10624731]
20. Dishion TJ, Andrews DW. Preventing escalation in problem behaviors with high-risk young adolescents: immediate and 1-year outcomes. J Consult Clin Psychol 1995;63(4):538-548. [doi: 10.1037//0022-006x.63.4.538] [Medline: 7673531]
21. Dishion TJ, Spracklen KM, Andrews DW, Patterson GR. Deviancy training in male adolescent friendships. Behav Ther 1996;27(3):373-390. [doi: 10.1016/s0005-7894(96)80023-2]
22. Hansen WB, Graham JW. Preventing alcohol, marijuana, and cigarette use among adolescents: peer pressure resistance training versus establishing conservative norms. Prev Med 1991;20(3):414-430. [doi: 10.1016/0091-7435(91)90039-7] [Medline: 1862062]
23. Hogue A, Henderson CE, Becker SJ, Knight DK. Evidence base on outpatient behavioral treatments for adolescent substance use, 2014-2017: outcomes, treatment delivery, and promising horizons. J Clin Child Adolesc Psychol 2018;47(4):499-526 [FREE Full text] [doi: 10.1080/15374416.2018.1466307] [Medline: 29893607]
24. Steele DW, Becker SJ, Danko KJ, Balk EM, Saldanha IJ, Adam GP, et al. Interventions for substance use disorders in adolescents: a systematic review. Rockville, MD: Agency for Healthcare Research and Quality; 2020.
25. Steele DW, Becker SJ, Danko KJ, Balk EM, Adam GP, Saldanha IJ, et al. Brief behavioral interventions for substance use in adolescents: a meta-analysis. Pediatrics 2020;146(4):e20200351. [doi: 10.1542/peds.2020-0351] [Medline: 32928988]
26. Stroul BA, Manteuffel BA. The sustainability of systems of care for children's mental health: lessons learned. J Behav Health Serv Res 2007;34(3):237-259. [doi: 10.1007/s11414-007-9065-3] [Medline: 17520373]
27. Walker DD, Stephens R, Roffman R, Demarce J, Lozano B, Towe S, et al. Randomized controlled trial of motivational enhancement therapy with nontreatment-seeking adolescent cannabis users: a further test of the teen marijuana check-up. Psychol Addict Behav 2011;25(3):474-484 [FREE Full text] [doi: 10.1037/a0024076] [Medline: 21688877]
28. Godley SH, Garner BR, Passetti LL, Funk RR, Dennis ML, Godley MD. Adolescent outpatient treatment and continuing care: main findings from a randomized clinical trial. Drug Alcohol Depend 2010;110(1-2):44-54 [FREE Full text] [doi: 10.1016/j.drugalcdep.2010.02.003] [Medline: 20219293]
29. Dennis M, Godley SH, Diamond G, Tims FM, Babor T, Donaldson J, et al. The cannabis youth treatment (CYT) study: main findings from two randomized trials. J Subst Abuse Treat 2004;27(3):197-213. [doi: 10.1016/j.jsat.2003.09.005] [Medline: 15501373]
30. Amodeo M, Lundgren L, Cohen A, Rose D, Chassler D, Beltrame C, et al. Barriers to implementing evidence-based practices in addiction treatment programs: comparing staff reports on Motivational Interviewing, Adolescent Community Reinforcement Approach, Assertive Community Treatment, and Cognitive-behavioral Therapy. Eval Program Plann 2011;34(4):382-389. [doi: 10.1016/j.evalprogplan.2011.02.005] [Medline: 21420171]
31. Gil AG, Wagner EF, Tubman JG. Culturally sensitive substance abuse intervention for Hispanic and African American adolescents: empirical examples from the Alcohol Treatment Targeting Adolescents in Need (ATTAIN) Project. Addiction 2004;99 Suppl 2:140-150. [doi: 10.1111/j.1360-0443.2004.00861.x] [Medline: 15488112]
32. Dauria EF, McWilliams MA, Tolou-Shams M. Substance use prevention and treatment interventions for court-involved, non-incarcerated youth. In: Monti PM, Colby SM, Tevyaw TO, editors. Brief interventions for adolescent alcohol and substance abuse. New York, NY: The Guilford Press; 2018:213-241.
33. Blevins CE, Walker DD, Stephens RS, Banes KE, Roffman RA. Changing social norms: the impact of normative feedback included in motivational enhancement therapy on cannabis outcomes among heavy-using adolescents. Addict Behav 2018;76:270-274 [FREE Full text] [doi: 10.1016/j.addbeh.2017.08.030] [Medline: 28886575]
34. Linke S, Murray E, Butler C, Wallace P. Internet-based interactive health intervention for the promotion of sensible drinking: patterns of use and potential impact on members of the general public. J Med Internet Res 2007;9(2):e10 [FREE Full text] [doi: 10.2196/jmir.9.2.e10] [Medline: 17513281]
35. Campbell AN, Nunes EV, Miele GM, Matthews A, Polsky D, Ghitza UE, et al. Design and methodological considerations of an effectiveness trial of a computer-assisted intervention: an example from the NIDA Clinical Trials Network. Contemp Clin Trials 2012;33(2):386-395 [FREE Full text] [doi: 10.1016/j.cct.2011.11.001] [Medline: 22085803]

36. Rooke S, Thorsteinsson E, Karpin A, Copeland J, Allsop D. Computer-delivered interventions for alcohol and tobacco use: a meta-analysis. Addiction 2010;105(8):1381-1390. [doi: 10.1111/j.1360-0443.2010.02975.x] [Medline: 20528806]
37. Riggs NR, Conner BT, Parnes JE, Prince MA, Shillington AM, George MW. Marijuana eCHECKUPTO GO: effects of a personalized feedback plus protective behavioral strategies intervention for heavy marijuana-using college students. Drug Alcohol Depend 2018;190:13-19. [doi: 10.1016/j.drugalcdep.2018.05.020] [Medline: 29960918]
38. Prince MA, Tyskiewicz AJ, Conner BT, Parnes JE, Shillington AM, George MW, et al. Mechanisms of change in an adapted marijuana e-CHECKUP TO GO intervention on decreased college student cannabis use. J Subst Abuse Treat 2021;124:108308. [doi: 10.1016/j.jsat.2021.108308] [Medline: 33771289]
39. Grove L, King CM, Bomysoad R, Vasquez L, Kois LE. Technology for assessment and treatment of justice-involved youth: a systematic literature review. Law Hum Behav 2021;45(5):413-426. [doi: 10.1037/lhb0000467] [Medline: 34871014]
40. Doumas DM, Esp S, Turrisi R, Hausheer R, Cuffee C. A test of the efficacy of a brief, web-based personalized feedback intervention to reduce drinking among 9th grade students. Addict Behav 2014;39(1):231-238. [doi: 10.1016/j.addbeh.2013.10.011] [Medline: 24148137]
41. Doumas DM, Hausheer R, Esp S, Cuffee C. Reducing alcohol use among 9th grade students: 6 month outcomes of a brief, web-based intervention. J Subst Abuse Treat 2014;47(1):102-105. [doi: 10.1016/j.jsat.2014.02.006] [Medline: 24666810]
42. Rideout V, Robb MB. The common sense census: media use by tweens and teens. Common Sense Media. 2019. URL: https://www.commonsensemedia.org/ [accessed 2020-10-26]
43. Anderson M, Jiang J. Teens' social media habits and experiences. Pew Research Center. 2018. URL: https://www.pewresearch.org/internet/2018/11/28/teens-social-media-habits-and-experiences/ [accessed 2019-02-17]
44. Dennis ML, Scott CK, Funk RR, Nicholson L. A pilot study to examine the feasibility and potential effectiveness of using smartphones to provide recovery support for adolescents. Subst Abus 2015;36(4):486-492 [FREE Full text] [doi: 10.1080/08897077.2014.970323] [Medline: 25310057]
45. Black SK, de Moor C, Kendall AD, Shrier LA. Feasibility of momentary sampling assessment of cannabis use in adolescents and young adults. J Child Adolesc Subst Abuse 2014;23(3):177-184. [doi: 10.1080/1067828x.2013.786923]
46. Ramo DE, Popova L, Grana R, Zhao S, Chavez K. Cannabis mobile apps: a content analysis. JMIR Mhealth Uhealth 2015;3(3):e81 [FREE Full text] [doi: 10.2196/mhealth.4405] [Medline: 26268634]
47. Tofighi B, Chemi C, Ruiz-Valcarcel J, Hein P, Hu L. Smartphone apps targeting alcohol and illicit substance use: systematic search in in commercial app stores and critical content analysis. JMIR Mhealth Uhealth 2019;7(4):e11831 [FREE Full text] [doi: 10.2196/11831] [Medline: 31008713]
48. Kazemi DM, Li S, Levine MJ, Auten B, Granson M. Systematic review of smartphone apps as a mHealth intervention to address substance abuse in adolescents and adults. J Addict Nurs 2021;32(3):180-187. [doi: 10.1097/JAN.0000000000000416] [Medline: 34473447]
49. Shrier LA, Rhoads AM, Fredette ME, Burke PJ. "Counselor in Your Pocket": youth and provider perspectives on a mobile motivational intervention for marijuana use. Subst Use Misuse 2014;49(1-2):134-144. [doi: 10.3109/10826084.2013.824470] [Medline: 24000892]
50. Kells M, Shrier L. Reducing cannabis use with a real-time intervention using mobile technology. In: Preedy VR, editor. Handbook of cannabis and related pathologies: biology, pharmacology, diagnosis, and treatment. Cambridge, MA: Elsevier Academic Press; 2017:1101-1110.
51. Shrier LA, Rhoads A, Burke P, Walls C, Blood EA. Real-time, contextual intervention using mobile technology to reduce marijuana use among youth: a pilot study. Addict Behav 2014;39(1):173-180. [doi: 10.1016/j.addbeh.2013.09.028] [Medline: 24139665]
52. Shrier LA, Burke PJ, Kells M, Scherer EA, Sarda V, Jonestrask C, et al. Pilot randomized trial of MOMENT, a motivational counseling-plus-ecological momentary intervention to reduce marijuana use in youth. Mhealth 2018;4:29 [FREE Full text] [doi: 10.21037/mhealth.2018.07.04] [Medline: 30148142]
53. Lyon AR, Koerner K. User-centered design for psychosocial intervention development and implementation. Clin Psychol (New York) 2016;23(2):180-200 [FREE Full text] [doi: 10.1111/cpsp.12154] [Medline: 29456295]
54. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017;19(5):e153 [FREE Full text] [doi: 10.2196/jmir.7725] [Medline: 28490417]
55. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
56. Marijuana eCHECKUP TO GO (e-TOKE). echeckup TO GO: San Diego State University Research Foundation. URL: https://www.echeckuptogo.com [accessed 2018-01-24]
57. Elliott JC, Carey KB, Vanable PA. A preliminary evaluation of a web-based intervention for college marijuana use. Psychol Addict Behav 2014;28(1):288-293. [doi: 10.1037/a0034995] [Medline: 24731118]
58. Elliott JC, Carey KB. Correcting exaggerated marijuana use norms among college abstainers: a preliminary test of a preventive intervention. J Stud Alcohol Drugs 2012;73(6):976-980. [doi: 10.15288/jsad.2012.73.976] [Medline: 23036216]

59. Knight JR, Sherritt L, Shrier LA, Harris SK, Chang G. Validity of the CRAFFT substance abuse screening test among adolescent clinic patients. Arch Pediatr Adolesc Med 2002;156(6):607-614. [doi: 10.1001/archpedi.156.6.607] [Medline: 12038895]
60. Grisso T, Fusco S, Paiva-Salisbury M, Perraout R, Williams V, Barnum R. The Massachusetts youth screening instrument-version 2 (MAYSI-2): comprehensive research review. National Youth Screening & Assessment Partners. Worcester: Univeristy of Massachusetts; 2012. URL: http://www.nysap.us/MAYSI-2%20Review.pdf [accessed 2021-09-23]
61. Rodriguez CA, Gopalakrishnan L, Del Cid M, Folk JB, Yonek J, Tolou-Shams M. The ethical implications of using social media to engage and retain justice-involved youth in behavioral health research. J Empir Res Hum Res Ethics 2021;16(4):356-363. [doi: 10.1177/15562646211039701] [Medline: 34533383]
62. Bonar EE, Schneeberger DM, Bourque C, Bauermeister JA, Young SD, Blow FC, et al. Social media interventions for risky drinking among adolescents and emerging adults: protocol for a randomized controlled trial. JMIR Res Protoc 2020;9(5):e16688 [FREE Full text] [doi: 10.2196/16688] [Medline: 32401225]
63. Yurasek AM, Kemp K, Otero J, Tolou-Shams M. Substance use screening and rates of treatment referral among justice-involved youth. Addict Behav 2021;122:107036. [doi: 10.1016/j.addbeh.2021.107036] [Medline: 34274872]
64. Burns BJ, Angold A, Magruder-Habib K, Costello EJ, Patrick MK. The child and adolescent services assessment (CASA) Version 4.2. Duke University Medical Centre. 1997. URL: https://devepi.duhs.duke.edu/measures/the-child-and-adolescent-services-assessment-casa/ [accessed 2018-02-19]
65. Farmer EM, Angold A, Burns BJ, Costello EJ. Reliability of self-reported service use: test-retest consistency of children's responses to the Child and Adolescent Services Assessment (CASA). J Child Fam Stud 1994;3(3):307-325. [doi: 10.1007/bf02234688]
66. Metrik J, Rohsenow DJ, Monti PM, McGeary J, Cook TA, de Wit H, et al. Effectiveness of a marijuana expectancy manipulation: piloting the balanced-placebo design for marijuana. Exp Clin Psychopharmacol 2009;17(4):217-225 [FREE Full text] [doi: 10.1037/a0016502] [Medline: 19653787]
67. Cuttler C, Spradlin A. Measuring cannabis consumption: psychometric properties of the daily sessions, frequency, age of onset, and quantity of cannabis use inventory (DFAQ-CU). PLoS One 2017;12(5):e0178194 [FREE Full text] [doi: 10.1371/journal.pone.0178194] [Medline: 28552942]
68. Eisenberg ME, Toumbourou JW, Catalano RF, Hemphill SA. Social norms in the development of adolescent substance use: a longitudinal analysis of the International Youth Development Study. J Youth Adolesc 2014;43(9):1486-1497 [FREE Full text] [doi: 10.1007/s10964-014-0111-1] [Medline: 24633850]
69. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
70. Morse JM. The significance of saturation. Qual Health Res 1995;5(2):147-149. [doi: 10.1177/104973239500500201]
71. O'Reilly M, Parker N. 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. Qual Res 2012;13(2):190-197. [doi: 10.1177/1468794112446106]
72. Dishion TJ, Tipsord JM. Peer contagion in child and adolescent social and emotional development. Annu Rev Psychol 2011;62:189-214 [FREE Full text] [doi: 10.1146/annurev.psych.093008.100412] [Medline: 19575606]
73. Peute LW, Spithoven R, Bakker PJ, Jaspers MW. Usability studies on interactive health information systems; where do we stand? Stud Health Technol Inform 2008;136:327-332. [Medline: 18487752]
74. Bull S, McFarlane M. Technology-based health promotion. Thousand Oaks, CA: Sage Publications; 2011.
75. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17 Suppl 1:11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
76. Ybarra ML, Prescott TL, Holtrop JS. Steps in tailoring a text messaging-based smoking cessation program for young adults. J Health Commun 2014;19(12):1393-1407. [doi: 10.1080/10810730.2014.901441] [Medline: 24766267]
77. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). JMIR Mhealth Uhealth 2016;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
78. McMahon RJ, Forehand RL. Helping the noncompliant child: family-based treatment for oppositional behavior. 2nd edition. New York, NY: The Guilford Press; 2005.
79. Sobell LC, Sobell MB. Timeline follow-back: a technique for assessing self-reported alcohol consumption. In: Litten RZ, Allen JP, editors. Measuring alcohol consumption: psychosocial and biochemical methods. Totowa, NJ: Humana Press; 1992:41-72.
80. Sobell MB, Maisto SA, Sobell LC, Cooper AM, Cooper T, Sanders B. Developing a prototype for evaluating alcohol treatment effectiveness. In: Inobell LC, Sobell LC, Ward E, editors. Evaluating alcohol and drug abuse treatment effectiveness: recent advances. Oxford, UK: Pergamon Press; 1980:129-150.
81. Sobell LC, Sobell MB. Alcohol timeline follow-back users' manual. Toronto, ON: Addiction Research Foundation; 1995.
82. Dennis M, Titus J, White M, Unsicker J, Hodgkins D. Global appraisal of individual needs (GAIN): administration guide for the GAIN and related measures (Version 5). Chestnut Health Systems. 2008. URL: https://www.chestnut.org/li/gain/gadm1299.pdf [accessed 2022-03-02]

83. Foster DW, Ye F, Chung T, Hipwell AE, Sartor CE. Longitudinal associations between marijuana-related cognitions and marijuana use in African-American and European-American girls from early to late adolescence. Psychol Addict Behav 2018;32(1):104-114 [FREE Full text] [doi: 10.1037/adb0000323] [Medline: 29094955]
84. Knapp AA, Babbin SF, Budney AJ, Walker DD, Stephens RS, Scherer EA, et al. Psychometric assessment of the marijuana adolescent problem inventory. Addict Behav 2018;79:113-119 [FREE Full text] [doi: 10.1016/j.addbeh.2017.12.013] [Medline: 29288984]
85. Torrealday O, Stein LA, Barnett N, Golembeske C, Lebeau R, Colby SM, et al. Validation of the marijuana effect expectancy questionnaire-brief. J Child Adolesc Subst Abuse 2008;17(4):1-17 [FREE Full text] [doi: 10.1080/15470650802231861] [Medline: 22058648]
86. Young RM, Gullo MJ, Feeney GF, Connor JP. Development and validation of the cannabis refusal self-efficacy questionnaire (CRSEQ) in adult cannabis users in treatment. Drug Alcohol Depend 2012;125(3):244-251. [doi: 10.1016/j.drugalcdep.2012.02.018] [Medline: 22445623]
87. Papinczak ZE, Connor JP, Feeney GF, Young RM, Gullo MJ. Treatment seeking in cannabis dependence: the role of social cognition. Drug Alcohol Depend 2017;170:142-146. [doi: 10.1016/j.drugalcdep.2016.11.005] [Medline: 27894043]
88. Stephens RS, Roffman RA, Fearer SA, Williams C, Burke RS. The marijuana check-up: promoting change in ambivalent marijuana users. Addiction 2007;102(6):947-957. [doi: 10.1111/j.1360-0443.2007.01821.x] [Medline: 17523990]
89. Stout RL, Wirtz PW, Carbonari JP, Del Boca FK. Ensuring balanced distribution of prognostic factors in treatment outcome research. J Stud Alcohol Suppl 1994;12:70-75. [doi: 10.15288/jsas.1994.s12.70] [Medline: 7723001]
90. Aarons GA, Brown SA, Stice E, Coe MT. Psychometric evaluation of the marijuana and stimulant effect expectancy questionnaires for adolescents. Addict Behav 2001;26(2):219-236. [doi: 10.1016/s0306-4603(00)00103-9] [Medline: 11316378]
91. Little R, Yau L. Intent-to-treat analysis for longitudinal studies with drop-outs. Biometrics 1996;52(4):1324-1333. [Medline: 8962456]
92. Baraldi AN, Enders CK. An introduction to modern missing data analyses. J Sch Psychol 2010;48(1):5-37. [doi: 10.1016/j.jsp.2009.10.001] [Medline: 20006986]
93. Spirito A, Hernandez L, Cancilliere MK, Graves HR, Rodriguez AM, Operario D, et al. Parent and adolescent motivational enhancement intervention for substance-using, truant adolescents: a pilot randomized trial. J Clin Child Adolesc Psychol 2018;47(Sup1):S467-S479 [FREE Full text] [doi: 10.1080/15374416.2017.1399402] [Medline: 29252011]
94. Carroll KM, Ball SA, Martino S, Nich C, Babuscio TA, Rounsaville BJ. Enduring effects of a computer-assisted training program for cognitive behavioral therapy: a 6-month follow-up of CBT4CBT. Drug Alcohol Depend 2009;100(1-2):178-181 [FREE Full text] [doi: 10.1016/j.drugalcdep.2008.09.015] [Medline: 19041197]
95. Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. Biol Psychiatry 2006;59(11):990-996. [doi: 10.1016/j.biopsych.2005.09.014] [Medline: 16368078]
96. Becker SJ, Helseth SA, Janssen T, Kelly LM, Escobar KI, Souza T, et al. Parent SMART (substance misuse in adolescents in residential treatment): pilot randomized trial of a technology-assisted parenting intervention. J Subst Abuse Treat 2021;127:108457. [doi: 10.1016/j.jsat.2021.108457] [Medline: 34134877]

Abbreviations

BIT: Behavior Intervention Technology
CINI: court-involved, nonincarcerated
ECTG: eCHECKUP-TO-GO
RCT: randomized controlled trial
SU: substance use
TAU: treatment-as-usual
TECH: Teen Empowerment through Computerized Health

Edited by T Derrick; submitted 31.12.21; this is a non–peer-reviewed article;accepted 06.01.22; published 11.03.22.

Please cite as:
Helseth SA, Guigayoma J, Price D, Spirito A, Clark MA, Barnett NP, Becker SJ
Developing a Smartphone-Based Adjunct Intervention to Reduce Cannabis Use Among Juvenile Justice-Involved Adolescents: Protocol for a Multiphase Study
JMIR Res Protoc 2022;11(3):e35402
URL: https://www.researchprotocols.org/2022/3/e35402
doi:10.2196/35402
PMID:35275086

Protocol

Feasibility of a Home-Based Exercise Program for Managing Posttransplant Metabolic Syndrome in Lung and Liver Transplant Recipients: Protocol for a Pilot Randomized Controlled Trial

Dmitry Rozenberg[1,2], MD, PhD; Daniel Santa Mina[3,4], PhD; Sahar Nourouzpour[1,2], MSc; Encarna Camacho Perez[1,5], MSc; Brooke Lyn Stewart[6], RD; Lisa Wickerson[2,7], BScPT, MSc, PhD; Cynthia Tsien[2,8], MPH, MD; Nazia Selzner[2,8], MD, PhD; Josh Shore[1], MSc; Meghan Aversa[1,2], MSc, MD; Minna Woo[9], MD, PhD; Sandra Holdsworth[10], Dip; Karina Prevost[10], BSc; Jeff Park[1], BSc; Amirhossein Azhie[2,8], MD; Ella Huszti[11], PhD; Elizabeth McLeod[6], RD; Sarah Dales[6], RD; Mamatha Bhat[2,8], MD, PhD

[1]Respirology and Lung Transplantation, Temerty Faculty of Medicine, Toronto General Hospital Research Institute, University Health Network, Toronto, ON, Canada

[2]Ajmera Transplant Program, University Health Network, Toronto, ON, Canada

[3]Faculty of Kinesiology and Physical Education, University of Toronto, Toronto, ON, Canada

[4]Department of Anesthesia and Pain Management, University Health Network, Toronto, ON, Canada

[5]GoodHope Ehlers-Danlos Syndrome Program, University Health Network, Toronto, ON, Canada

[6]Nutrition, Ajmera Transplant Centre, University Health Network, Toronto, ON, Canada

[7]Department of Physical Therapy, University of Toronto, Toronto, ON, Canada

[8]Gastroenterology and Liver Transplantation, Temerty Faculty of Medicine, University Health Network, Toronto, ON, Canada

[9]Endocrinology, Temerty Faculty of Medicine, Toronto General Hospital Research Institute, University Health Network, Toronto, ON, Canada

[10]Canadian Donation and Transplantation Research Program, Edmonton, ON, Canada

[11]Biostatistics Research Unit, University Health Network, Toronto, ON, Canada

Corresponding Author:
Dmitry Rozenberg, MD, PhD
Respirology and Lung Transplantation, Temerty Faculty of Medicine
Toronto General Hospital Research Institute
University Health Network
200 Elizabeth Street, 13 EN-229
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 340 4800 ext 7358
Fax: 1 416 340 4531
Email: Dmitry.Rozenberg@uhn.ca

Abstract

Background: Posttransplant metabolic syndrome (PTMS) is a common contributor to morbidity and mortality among solid organ transplant recipients in the late posttransplant period (≥1 year). Patients diagnosed with PTMS are at a higher risk of cardiovascular disease and frequently experience decreased physical function and health-related quality of life (HRQL). Studies in the early posttransplant period (<1 year) have shown the benefits of facility-based exercise training on physical function and HRQL, but have not evaluated the effects on metabolic risk factors. It remains unclear whether home-based exercise programs are feasible and can be delivered at a sufficient exercise dose to have effects on PTMS. This protocol outlines the methodology of a randomized controlled trial of a partly supervised home-based exercise program in lung transplant (LTx) and orthotopic liver transplant (OLT) recipients.

Objective: This study aims to evaluate the feasibility (ie, recruitment rate, program adherence, attrition, safety, and participant satisfaction) of a 12-week individualized, home-based aerobic and resistance training program in LTx and OLT recipients initiated 12 to 18 months after transplantation, and to assess estimates of intervention efficacy on metabolic risk factors, exercise self-efficacy, and HRQL.

Methods: In total, 20 LTx and 20 OLT recipients with ≥2 cardiometabolic risk factors at 12 to 18 months after transplantation will be randomized to an intervention (home-based exercise training) or control group. The intervention group will receive an

individualized exercise prescription comprising aerobic and resistance training, 3 to 5 times a week for 12 weeks. Participants will meet on a weekly basis (via videoconference) with a qualified exercise professional who will supervise exercise progression, provide support, and support exercise self-efficacy. Participants in both study groups will receive a counseling session on healthy eating with a dietitian at the beginning of the intervention. For the primary aim, feasibility will be assessed through recruitment rate, program adherence, satisfaction, attrition, and safety parameters. Secondary outcomes will be measured at baseline and 12 weeks, including assessments of metabolic risk factors (ie, insulin resistance, abdominal obesity, blood pressure, and cholesterol), HRQL, and exercise self-efficacy. Descriptive statistics will be used to summarize program feasibility and effect estimates (means and 95% CIs) for sample size calculations in future trials.

Results: Enrollment started in July 2021. It is estimated that the study period will be 18 months, with data collection to be completed by December 2022.

Conclusions: A partly supervised home-based, individually tailored exercise program that promotes aerobic and resistance training and exercise self-efficacy may be an important intervention for improving the metabolic profile of LTx and OLT recipients with cardiometabolic risk factors. Thus, characterizing the feasibility and effect estimates of home-based exercise constitutes the first step in developing future clinical trials designed to reduce the high morbidity associated with PTMS.

Trial Registration: ClinicalTrials.gov NCT04965142; https://clinicaltrials.gov/ct2/show/NCT04965142

International Registered Report Identifier (IRRID): DERR1-10.2196/35700

(JMIR Res Protoc 2022;11(3):e35700) doi:10.2196/35700

KEYWORDS

lung transplant; liver transplant; posttransplant metabolic syndrome; exercise training; randomized controlled trial; pilot study

Introduction

Background

Posttransplant metabolic syndrome (PTMS) is prevalent among solid organ transplant recipients with >25% of lung transplant (LTx) and 50% of orthotopic liver transplant (OLT) recipients developing PTMS within 12-18 months [1,2]. PTMS comprises impaired glucose tolerance, obesity, hypercholesterolemia, and hypertension and has been associated with an increased risk of hospital readmissions and cardiovascular morbidity accompanied by decreased long-term survival [3-5]. LTx and OLT recipients frequently present with risk factors for PTMS, including increased appetite and weight gain along with a reversal of the pretransplant catabolic state, immunosuppression, and physical inactivity [6]. Our research group has shown that in a sample of >2000 OLT recipients in the Toronto Liver Program (1990-2015), 35% had pre-existing, new onset, or transient posttransplant diabetes mellitus, a PTMS risk factor associated with decreased long-term survival [7]. A retrospective evaluation of 227 LTx recipients conducted by our group also demonstrated a significant increase in the prevalence of metabolic risk factors (hypertension, hyperlipidemia, diabetes mellitus, and obesity) by 1 year after transplantation (≥3 factors: 26%) compared with before transplantation (11%) [8].

LTx and OLT recipients have been observed to engage in less physical activity than the general population and report barriers to exercise, including lack of specific exercise guidelines, medication side effects, comorbidities, and inability to access fitness facilities [9,10]. In a cross-sectional survey of 656 solid organ transplant recipients, van Adrichem et al [11] found that only 56% of patients met the physical activity guidelines and engaged in ≥150 minutes per week of moderate aerobic activity after transplantation. A 2019 consensus statement [12] highlighted the importance of strategies to mitigate posttransplant metabolic risk factors with physical training. The recommendation for solid organ transplant recipients is to practice moderate to vigorous intensity physical training (either aerobic or aerobic plus resistance training) 3 to 5 times a week for a minimum of 8 weeks [12].

The posttransplant period offers an opportunity for transplant recipients to derive significant metabolic and functional benefits from exercise, as most of them can train at a greater volume and intensity than in the pretransplant period [13]. In this regard, 24 sessions of 30-minute treadmill exercise sessions were associated with a significant increase in resting energy expenditure in OLT recipients [14]. Moreover, a meta-analysis of 15 randomized controlled trials (RCTs) showed a significant reduction in body fat percentage in OLT recipients (1 trial, 119 patients) who participated in physical exercise training (median −5.40%, 95% CI −8.03% to −2.77%; P<.001) [15]. Most of the exercise programs included in this meta-analysis implemented exercise interventions starting a year after transplantation (program duration between 8 and 24 weeks) and focused on aerobic and strength training.

To date, most exercise interventions in LTx and OLT recipients have been delivered as supervised, facility-based programs within the first 12 months after transplantation [16-19]. However, facility-based programs present challenges for patients [12,20], especially in the late posttransplant period. Given the decreased frequency of clinic visits, relocation requirements, return to work [21], and the interruption of hospital and community-based programs during the COVID-19 pandemic [22], home-based exercise programs can overcome some of these challenges and have been recognized as an important strategy that requires further research [20].

In light of the current challenges to health care delivery because of the COVID-19 pandemic, telerehabilitation and remote exercise interventions are gaining attention as the preferred modalities for promoting physical activity and exercise behaviors among solid organ transplant recipients [23-26]. Only

a few studies on home-based exercise programming in the early and late posttransplant periods have been conducted in both LTx [27-29] and OLT [30,31] recipients, and they have suggested benefits to aerobic capacity, quadriceps strength, and health-related quality of life (HRQL). In one of these studies, program adherence in OLT recipients was poor (only 37% of participants completed ≥50% of exercises with bimonthly phone calls) [30], whereas adherence was not assessed in a home-based program with 12 LTx recipients (mean 36 months, SD 33 months after transplantation) [27]. Furthermore, the effects of home-based training on metabolic risks were evaluated in only 1 OLT study, which used a combination of a few personalized sessions and group telehealth classes for exercise and nutritional counseling for a 3-month period and showed a modest improvement in metabolic syndrome and HRQL [32]. However, the optimal structure of counseling, exercise prescription, and effects on metabolic risk factors among solid organ transplant groups remains to be defined [12].

There is limited literature surrounding the experiences of transplant recipients concerning knowledge of, motivation for, and barriers to exercise training, which limits our understanding of the feasibility of real-world implementation. Several strategies to engage participants in exercise interventions and optimize adherence have been explored for chronic diseases [9]. The literature on chronic disease management highlights that adherence can be increased by providing close support by an exercise professional, simplifying the number of exercises, and fostering self-efficacy for exercise [33]. The support and coaching by the exercise professional can be enhanced with the use of telecommunication and the ability of the participants to readily contact the health coach as needed [26,27]. In addition, focusing on four sources of self-efficacy built into an exercise program (mastery, vicarious experiences, verbal feedback, physiological and emotional state) has been shown to significantly improve exercise adherence in patients with cardiometabolic risk factors [34,35] but has not been applied to transplant recipients.

These gaps in knowledge are important to address as optimal effects on PTMS are likely achieved if exercise is performed with sufficient volume (at least moderate intensity, accumulating ≥150 min/week) [36]. However, it remains unclear whether this training dose is optimal or if it can be achieved with a home-based program targeting metabolic risk factors starting at 12-18 months after transplantation, a critical period prognostic of long-term PTMS and cardiovascular morbidity [37]. Before these efficacy questions can be addressed with an appropriately powered RCT, a pilot study evaluating the feasibility of a home-based exercise program is needed.

Aims

The specific aims of this study are (1) to evaluate the feasibility of a 12-week individualized, home-based aerobic and resistance training program in LTx and OLT recipients at 12-18 months after transplantation and (2) to assess estimates of intervention efficacy on PTMS risk factors, exercise self-efficacy, and HRQL. We hypothesize that we will achieve a recruitment rate of 30% of eligible LTx and OLT recipients, participants in the intervention group will demonstrate an adherence rate of ≥70% to the prescribed exercise dose (including attainment of the prescribed exercise progression over time) and that PTMS risk factors, self-efficacy, and HRQL will be improved in the intervention group compared with the control group.

Methods

Design

This paper describes a pilot RCT to assess the feasibility of a phase 3 RCT examining the effects of an individualized, partly supervised, home-based exercise intervention in LTx and OLT recipients. The study is being conducted at the Toronto General Hospital in Toronto, Canada, and ethical approval was obtained from the research ethics board at the University Health Network (study ID: 20-5185). The study was registered at ClinicalTrials.gov NCT04965142. Informed written consent will be obtained from all participants before starting any research activities.

This protocol follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [38], the CONSORT (Consolidated Standards of Reporting Trials) statement [39], and the Consensus on Exercise Reporting Template [40]. Two patient partners are part of the research team, have provided input into the study design, and will contribute support and their lived experience throughout the study. The research team will hold meetings with patient partners every 3 to 6 months to collect feedback on study progress, assess concerns, and work on knowledge translation tools once the study is completed. Modifications to the study protocol will be approved by the research ethics board before implementation, communicated to funding agencies, and outlined in future publications.

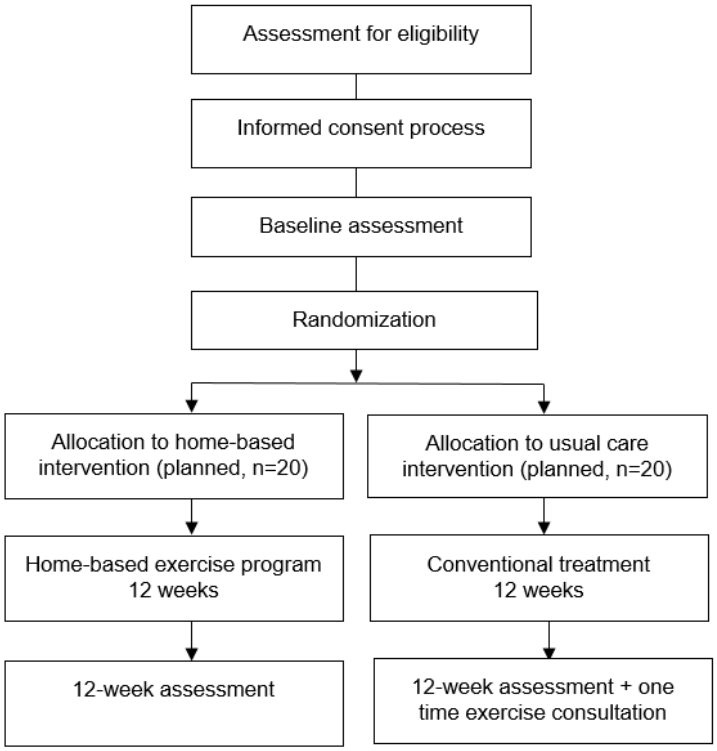
Participants

The study sample will consist of adult LTx and OLT recipients (≥18 years) with at least two or more metabolic risk factors (ie, hypertension, hyperlipidemia, diabetes, and obesity) at 12 to 18 months after transplantation. Exclusion criteria include (1) active cardiovascular disease (eg, recent heart attack, significant coronary artery disease on cardiac catheterization, heart failure, uncontrolled arrhythmias, chest pain, dizziness, or fainting in the last 3 months); (2) neuromuscular disease or orthopedic limitations; and (3) a self-reported active lifestyle (ie, patients achieving ≥150 min/week of moderate-intensity aerobic physical activity).

Recruitment and Screening

Potential participants will be recruited from posttransplant lung and liver clinics. The study team will identify potential participants at 12 to 18 months after transplantation with ≥2 metabolic risk factors. Participants will be asked by members of their circle of care if they are interested in participating in the study. If agreeable, the research team will contact the prospective participant to confirm eligibility criteria, discuss study participation, and conduct the informed consent process. The study flow of participants is shown in Figure 1.

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



Randomization

After the baseline assessment, participants will be randomly allocated using a 1:1 ratio to either the intervention or the control group, stratified by transplant type (ie, liver or lung). The allocation of participants to their groups will be randomized using shuffled, opaque envelopes by a biostatistician with no involvement in data collection in the study. The research team, exercise professionals, and study participants will be aware of the group allocation.

Study Assessments

Overview

Study assessments will be conducted at baseline and 2, 6, and 12 weeks. For a detailed list of assessment procedures by time point, refer to Tables 1 and 2. For participants unable to come on-site, assessments will be performed remotely from participants' homes, and blood screenings and electrocardiogram assessments will be conducted at local laboratories. During remote assessments, participants will be supervised by the research team using videoconferencing. To ensure the safety and confidentiality of participants undergoing remote assessments, the research team will ascertain the participant's identity, home address, and phone number. An assessment of the suitability of the participant to perform physical assessments will also be performed before starting any interaction with participants (eg, self-reported breathlessness, dizziness, and chest pain). Participants will be instructed to discontinue exercise testing if they experience any symptoms or feel unwell.

Table 1. Assessment of primary study outcomes by time point.

Primary outcomes	Assessment measures	Time point
Recruitment	Consented patients over total of eligible patients approached	During study period
Retention	Measured by attrition throughout the intervention period	During study period
Adherence	Completion of at least 70% of intervention elements	During study period
Satisfaction	Satisfaction survey	Weeks 2, 6, and 12[a]

[a]For usual care, satisfaction will be evaluated at baseline and 12 weeks.

Table 2. Assessment of secondary study outcomes by time point.

Secondary outcomes	Assessment measures	Time point
Metabolic risk factors		
• Glucose tolerance (fasting glucose and HbA1c[a] test), insulin sensitivity, (fasting insulin and HOMA-IR[b]), cholesterol panel, and C-reactive protein • Diabetes, cholesterol, and hypertension medications • Abdominal obesity	Blood work and medical charts • Blood pressure monitor, charts, and home measurement	Baseline and 12 weeks
Liver fibrosis and steatosis (OLT[c] patients attending assessments on-site)	Transient elastography (liver stiffness and controlled attenuation parameter for steatosis assessment)	Baseline and 12 weeks
Clinical parameters		
• Model for end-stage liver disease at transplant • Lung allocation score • Medication history • Electrocardiogram • Liver enzymes • Hospital admissions • Graft function and comorbidities	Medical charts	Baseline and 12 weeks
Physical function	Short performance physical battery	Baseline and 12 weeks
Body composition	• Bioelectrical impedance • Waist circumference • BMI	Baseline and 12 weeks
Self-reported outcomes		
• Quality of life • Barriers to exercise • Physical activity behaviors • Dietary habits	• Short-Form Survey-36 • Lifestyle questionnaire • Physical Activity Scale for the Elderly • Rapid Eating Assessment 3-day nutritional intake	Baseline and 12 weeks
• Exercise self-efficacy	• Exercise Self-efficacy Scale[d]	Baseline, 2, 6, and 12 weeks[d]

[a]HBA_{1c}: hemoglobin A_{1c}.

[b]HOMA-IR: homeostatic model assessment for insulin resistance.

[c]OLT: orthotopic liver transplant.

[d]For usual care, exercise self-efficacy will be evaluated at baseline and 12 weeks.

Study Arms

Following randomization, all study participants will be provided with a physical activity tracker (Fitbit) and exercise or physical activity logs to monitor exercise adherence and capture lifestyle changes in the intervention group and any contamination in the control group. Given the importance of diet on metabolic risk factors, both study groups will receive a counseling session on healthy eating consistent with the Canadian Food Guide by a registered dietitian, a review of the 3-day food record, and a nutritional handout. Nutritional advice will be tailored to specific metabolic concerns (eg, weight management, diabetes, and hypercholesterolemia). The dietitian will review all participants' medical histories and pertinent bloodwork as routine care during dietary assessments.

Intervention Group

Overview

Our exercise program was informed by expert guidelines on exercise prescription in solid organ transplant recipients [12,20,41] and will be delivered by an exercise professional with experience in chronic disease management. Following the baseline assessment and randomization, intervention participants will receive an individualized, home-based exercise program consisting of aerobic and resistance training exercises for 12 weeks. The exercise prescription targets the Canadian Physical Activity guidelines [42], which recommends ≥150 minutes of at least moderate-intensity aerobic training per week and at least two days per week of resistance training for major muscle groups. Daily exercise practice will be unsupervised; however, the exercise professional will provide the exercise prescription during the first training session via videoconference, including the following: an assessment of the participant's ability to exercise independently in their home environment (eg, review

of the participants' heart rate, blood pressure, and history of any pertinent injuries), demonstration and review of training techniques, and proper use of their Fitbit and exercise equipment.

Participants will also receive a detailed manual describing the importance of exercise and self-management skills for exercise behaviors based on the theory of planned behavior and social cognitive theory adapted from previous studies [43,44]. The manual will highlight the basics of exercise training, the importance of routines, principles related to exercise progression, and photographs with descriptions of the prescribed exercises. After the initial training session, the exercise professional will follow up weekly (via videoconference) with study participants to address potential barriers to exercise, provide motivational support, review symptoms or adverse events during independent training at home, and prescribe training progression. The exercise program will be tailored to the specific and changing needs of the participant, taking into consideration comorbidities, exercise capacity, physical function, and availability of training equipment. Details on exercise prescription for the intervention group are briefly provided and described in Table 3.

Table 3. Exercise prescription details for participants in the intervention group.

Modality and frequency	Intensity	Time and volume of exercise	Type of exercise	Progression
Aerobic; 3-5 days/week	• Moderate to vigorous (target Borg RPE[a]: 12-14/20) • 65% to 85% of age-predicted maximum heart rate	• 30 minutes • Continuous or interval training • ≥150 minutes of at least moderate-intensity aerobic training per week	• Walking • Running • Alternating walking and running bouts • Biking	• Gradual progression of training duration (up to 60 min/session) • Increase in walking and running speed or inclination • If cycling, increase in revolutions per minute or resistance • Increase in prescribed percentage of maximum heart rate

Resistance; 2 days/week	Target Borg RPE: 12-14/20	• 1-3 sets
		• 6 exercises
		• 10-12 repetitions
			• Resistance band exercises:
			  • Biceps curl
			  • Triceps extension
			  • Low row
			  • Shoulder abduction
			• Body weight exercises:
			  • Wall squats
			  • Calf raises
				• Resistance intensity will be increased when 3 sets of 12 repetitions can be completed without difficulty to maintain moderate exertion
				• Increase in resistance band tension
				• Increase in the number of sets, reduction in resting time

[a]RPE: rating of perceived exertion.

Aerobic Training

The aerobic exercise prescription comprises three to five 30-minute sessions per week at a moderate intensity of 65% to 85% of the estimated maximal heart rate or 12-14/20 Borg rating of perceived exertion (RPE) scale [45]. The maximum heart rate will be estimated using the formula 207-(0.7×age in years) [46]. Training modalities may include walking outdoors, using a treadmill, cycling with a stationary bicycle, or other modalities available to the participant. The duration of the aerobic training session will be adjusted or divided (eg, 2 sessions of 15 minutes) if the participant has limitations because of comorbidities.

Resistance Training

Resistance training will be prescribed for at least two days per week and will focus on general conditioning comprising 6 to 10 exercises targeting the major muscle groups (1-3 sets of 8-12 repetitions). Prescribed resistance training exercises will include a combination of resistance bands with different tension and body-weight exercises. Resistance intensity will be increased when 3 sets of 12 repetitions can be completed without difficulty to maintain moderate exertion despite muscular adaptation (target Borg RPE=12-14/20) [43,44].

Safety Precautions for Remote Exercise Training

Participants will be instructed to discontinue exercise training if they experience any symptoms or feel unwell and will be strongly encouraged to immediately contact a member of the research team upon noticing any adverse events. Participants will be advised to reduce the intensity of their training if they reach ≥15 on the Borg RPE scale, if oxygen saturation is less than 85%, or if they experience any symptoms (eg, severe shortness of breath). If participants report chest pain or discomfort (ie, uncomfortable feeling of pressure, pain, squeezing, or heaviness in the chest spreading to the shoulder, arms, neck, and back), they will be instructed to stop and rest and seek emergency assistance if the discomfort persists for 5 minutes.

Participants who communicate medically concerning adverse events (eg, cardiovascular event or disease exacerbation) or contraindications to the exercise program will have their training suspended until clearance is provided by the study physician and the exercise professional. If participants sustain an injury because of participating in the study, the study team will ensure that appropriate medical treatment is received.

Before initiating any interaction with study participants, the study team will perform an objective assessment (ie, by measuring oxygen saturation, blood pressure, and heart rate if participants have a pulse oximeter or a blood pressure monitor at home). A subjective assessment of the study participants (eg, self-reported breathlessness, dizziness, and chest pain) will ascertain their suitability to proceed with testing, training, and progress with the exercise prescription. The study team will discuss with participants aspects of their home environment, availability and suitability of exercise training equipment, home walking space, and any potential safety concerns (eg, footwear, rugs, clutter, and pets). During the internet-based training sessions, participants will be supervised by the study team via a videoconference platform. Participants will be able to book additional sessions with the exercise professional if any concerns or questions arise.

Control Group

The control group will meet with the exercise professional at the start of the study to highlight the benefits of physical activity and will be provided with physical activity monitors (Fitbits) to track their daily physical activity levels (minutes and steps) along with physical activity logs to record daily physical activity. The study team will correspond with the control group every 2 to 4 weeks during the study period to see if any questions arise regarding the physical activity logs and trackers. After 12 weeks, the control group will have the option of receiving an exercise session with the exercise professional. During this session, study participants will receive general physical activity and exercise training recommendations and an exercise manual tailored to the needs of LTx and OLT recipients.

Outcomes

A schedule of study outcome measurements is provided in Tables 1 and 2.

Primary Outcomes

Feasibility

The feasibility of a larger RCT to assess intervention efficacy (ie, phase 3 clinical trial) will be assessed through recruitment rates, program adherence, contamination, attrition, and safety at the end of the 12 weeks.

Recruitment

The recruitment rate will be defined as the proportion of randomized patients relative to all eligible patients approached for study participation. A consent rate of 30% or greater has been established as our criteria to determine study feasibility for future projects. The reasons for study nonparticipation will be collected.

Retention

Retention for the whole study and per group will be determined by the number of participants completing the final study assessment compared with the number of randomized participants. A retention rate of 80% or greater has been established as our criteria to determine study feasibility for future projects. The reasons for dropping out or interruptions to training will be collected.

Adherence

Adherence to the prescribed exercise program will be defined as completion of at least 70% of intervention components by the intervention group. This will be ascertained during weekly communication with participants and through a review of completed daily exercise logs, indicating exercise days, duration, intensity, and frequency of exercise sessions. Similarly, contamination in the control group will be assessed using physical activity logs.

Safety

The safety of the intervention will be monitored via daily logs (record of adverse events) and will be ascertained by the qualified exercise professional and reviewed by the principal investigator during weekly meetings.

Satisfaction

Satisfaction with the exercise program and other study interventions (eg, nutritional counseling) will be assessed at weeks 2 and 6 and at the end of the 12-week intervention via a satisfaction survey (multiple-choice questions) created by the research team. The control group will also complete a satisfaction survey in which participants will provide feedback on their experiences at the end of the study.

Secondary Outcomes

Preliminary efficacy data will be collected at baseline and 12-week assessments (or the last recorded day of program completion). Details on the study measures by time point are presented in Table 2.

Metabolic Risk Factors

Impaired glucose tolerance, abdominal obesity, cholesterol, and blood pressure will be assessed and classified according to the National Cholesterol Education Program Adult Treatment Panel III criteria [47]. More specifically, the following markers will be used to classify metabolic risks: (1) increased waist circumference (>102 cm [>40 in] for men, >88 cm [>35 in] for women); (2) elevated triglycerides of ≥1.7 mmol/L (≥150 mg/dL); (3) low levels of high-density lipoprotein: <1 mmol/L (40 mg/dL, male), <1.3 (50 mg/dL, female), or on cholesterol medication; (4) a resting blood pressure of ≥130/85 mm Hg or on hypertensive medication; and (5) fasting plasma glucose levels of 6.1 mmol/L (110 mg/dL) or on pharmacotherapy for diabetes. Participants will have the option of completing bloodwork either on-site or externally in a medical laboratory. Blood will be collected in a fasting state (no eating and drinking for 8 hours before testing; however, water is allowed) and includes a total cholesterol panel (ie, total cholesterol, triglycerides, low-density lipoprotein, and high-density lipoprotein), fasting blood glucose level, hemoglobin A_{1c} test, and C-reactive protein levels. Insulin resistance will be captured using the homeostatic model assessment for insulin resistance protocol (fasting insulin multiplied by fasting blood glucose). Finally, C-peptide levels will be analyzed in participants receiving exogenous insulin therapy.

Clinical Parameters

A liver transient elastography (FibroScan) assessment will be performed in liver transplant recipients to assess the degree of liver fibrosis (thickening and scarring of tissues) and steatosis (the amount of fat in liver tissue). Medical charts will be reviewed to collect data on age, sex, transplant type, the model for end-stage liver disease and lung allocation score, and detailed medication history through pharmacy records with cumulative and average daily doses of prednisone (mg/day) and calcineurin inhibitors, mTOR inhibitors, and antimetabolites (mycophenolate mofetil). Furthermore, liver enzymes along with total bilirubin, hemoglobin, and renal function (creatinine) levels will be abstracted from chart review. Data on posttransplant hospital admissions, allograft rejection, and cardiovascular comorbidities will also be ascertained from medical charts from the time of transplant to the end of the study period. Lifestyle factors, such as current alcohol consumption and smoking, will be ascertained from the chart review.

Physical Activity and Exercise Behaviors

Previous physical activity and exercise training habits will be collected at baseline by self-report (after discussion with the research team) and by completion of the Physical Activity Scale for the Elderly questionnaire [48], a short survey created to assess physical activity levels in older adults commonly applied in solid organ transplant recipients [9].

Physical Function

The short performance physical battery [49] test is used to assess gait, balance, and lower extremity performance. Participants will be asked to stand in one position holding their balance, rise from a chair 5 times, and walk for 4 m while being timed at their usual pace. For remote assessments (via videoconference), the balance test may be excluded if participants report a history of balance impairments or falls.

Body Composition

Weight, BMI, and body fat percentage will be measured via bioelectrical impedance analysis (Tanita, DC-430U) for participants coming on-site for study assessments; otherwise, participants will be asked to measure their weight at home if they have a scale. Participants will be asked to measure and record their body weight once a week. Height data will be obtained from the medical charts. For participants coming on-site for study assessments, waist circumference will be measured by the research team according to the World Health Organization standardized protocol. For participants completing the assessment remotely, waist circumference will be self-assessed using a measuring tape via videoconference under the direct supervision of the research team.

Nutritional Outcomes

The dietary habits of our study participants will be assessed using the Rapid Eating Assessment for Patients [50] questionnaire and a 3-day food record.

Health-Related Quality of Life

The 36-Item Short-Form Survey [51] will be used to measure mental and physical domains of quality of life in our participants (scores range from 0 to 100, with higher scores representing higher HRQL). The Short-Form 36 has been routinely applied in both OLT [30] and LTx recipients [52].

Self-efficacy

The Exercise Self-efficacy Scale [53], a Likert scale with 4-point rating in which participants rate their confidence levels with regard to carrying out regular physical activities and exercise, will be implemented to measure self-efficacy toward exercise.

Barriers to and Facilitators of Exercise

Familiarity and comfort levels surrounding technology, barriers to exercise, and evaluation of previous experience with exercise, especially within the home environment, will be captured using a questionnaire developed by our research team.

Statistical Procedures

Sample Size

The sample size of 40 (10 per study group and transplant type) is in keeping with the recommendations for a pilot study to assess feasibility [54]. A power calculation was not performed for the secondary outcomes, as the main purpose of this study is to determine feasibility and obtain point estimates for a larger RCT. The liver transplant and LTx program at the University Health Network each has approximately 150 to 200 recipients per year. Furthermore, the rates of PTMS among LTx and OLT recipients range from 30% to 50%; thus, we anticipate an approximate eligible pool of 50 transplant recipients per year for each organ group. Our previous consent rates for rehabilitation studies in these populations have been approximately 30% to 50% [55,56].

Proposed Data Analyses

Descriptive statistics (means, SDs, or medians and IQRs) will be used for continuous variables, and frequencies will be used for categorical variables to summarize program adherence, participant satisfaction, contamination, attrition, and safety in the intervention group, and to compare any differences between LTx and OLT recipients across these parameters using 2-tailed *t* tests and tests of proportion. Estimates at baseline and 12 weeks within- and between-group will be reported using means and 95% CIs to derive effect estimates for sample size calculations in subsequent trials. Statistical analysis will be performed by a statistician who will be blinded to the group assignments. All eligible participants (N=40) will be included in the data analyses based on intention-to-treat, depending on the group to which they are randomized.

Data Management and Quality Assurance

Our research team is trained in the study requirements, measurement protocols, and has expertise in conducting physical assessments in solid organ transplant recipients. The research team has critically appraised the external peer-review comments from the Canadian Donation Transplant Research Program (Multimedia Appendix 1) and has incorporated them into the study protocol. Standard operating procedures have been generated for all the protocol elements. The accuracy of the data entered in our research database will be periodically audited by the research coordinator. The research team will attempt to have participants complete the entirety of the study assessments, and

if there are incomplete data for a specific participant or missing responses to questionnaires, missing data will not be imputed. Participants will be closely monitored by the research team and the exercise professional to maximize study retention and assess exercise adherence rates during the 12-week intervention. Participants in the intervention group will be educated on the importance of adhering to the intervention, logging progress, and completing study assessments. The research team will send reminders to intervention participants for weekly meetings, and study assessments will be booked at convenient times for participants.

The study participants will have a random identifier generated with only the coded identifier attached to the participant data. The study data will be stored on a password-protected server following the currently established institutional and national research ethics and privacy guidelines. The study information will be accessible only to approved study members. In the event of an inappropriate release of personal health information to an unauthorized party, we will take the appropriate steps to minimize any potential harm: (1) further release of information will be stopped, (2) an attempt to retrieve all inappropriately released information will be made, and (3) the sponsor's privacy office and research ethics board will be immediately notified.

Results

This research was funded in July 2020, and enrollment began in July 2021. It is estimated that the study period will be 18 months and that study assessment and data collection will be completed by December 2022. The study results will be submitted for publication in the first half of 2023.

Discussion

Overview

PTMS is a common sequela of solid organ transplantation and is associated with increased cardiovascular morbidity and mortality among LTx and OLT recipients [57,58]. Studies in the early posttransplant period (<1 year) have shown the benefits of facility-based exercise training on physical function and HRQL, but to date, only a recent study has evaluated the effects of 3 months of telehealth sessions on metabolic risk factors in OLT recipients [32]. These teleconferencing group sessions comprised alternating exercise and nutrition-based interventions with a few personalized sessions, which demonstrated improvement in PTMS, diet adherence, and mental HRQL. However, it remains unclear whether home-based exercise programs can be delivered with a sufficient exercise dose and adequate adherence to have effects on metabolic risk factors and exercise self-efficacy in both LTx and OLT recipients. This is an important consideration given the significant differences in routine perioperative rehabilitation practices across solid organ transplant recipients. For instance, LTx recipients participate in both pre- and posttransplant rehabilitation [41], whereas many centers do not have a dedicated exercise program for OLT recipients and may impact self-efficacy, adherence, and perception related to exercise training [12,59]. Thus, this protocol outlines the methodology of a pilot study implementing a partly supervised personalized home-based exercise program for LTx and OLT recipients as the first key step in developing future clinical trials to offset the high morbidity associated with PTMS.

Home-Based Exercise Training

With fewer patients attending nonessential visits to health care centers during the COVID-19 pandemic, providing patients with tools to facilitate independent training has become a priority. There is increasing evidence that remote exercise interventions across a range of clinical populations have emerged in the last year [60-62], but more importantly in LTx and OLT recipients, showing promising results in training volumes and HRQL, respectively [26,32]. Furthermore, home-based interventions have become an important strategy in a number of chronic conditions and have the potential to be incorporated into the standard of care treatment plan for LTx and OLT recipients [63]. The novelty of the present home-based program lies in the promotion of exercise behaviors by increasing patients' self-efficacy to exercise in a familiar environment with minimal equipment, resources, and supervision. Furthermore, home-based exercise training allows for reduced travel and increased convenience for participants to exercise in the comfort of their homes [26,32,64].

Strengths and Limitations

Our study protocol has several strengths. First, we aimed to design a standardized, high-quality intervention grounded on theoretical and practical considerations specifically tailored for LTx and OLT recipients. The 12-18–month posttransplant period was chosen given the high rates of PTMS at this time point and the barriers experienced by patients [9,10]. We acknowledge that implementation of an exercise rehabilitation program early in the posttransplant period (eg, 3 months) may help prevent the occurrence of some of these metabolic risk factors, but we are specifically interested in addressing the knowledge gap of whether PTMS may be attenuated with a home-based exercise program in the late posttransplant period. Second, the home-based delivery format of this intervention constitutes a novel approach in the LTx and OLT populations. We prioritized implementation feasibility in the design of our protocol by creating an exercise program that can be replicated at home with minimal equipment and supervision while meeting safety parameters. Our intervention protocol was also designed to promote the sustainability of behaviors after the study period with exercises and activities that can be performed every day, such as walking, which also promotes transference to daily routines. Other highlights of our research include the inclusion of a control group, the rigor in our methodology (detailed and accurate reporting of methods according to the CONSORT, SPIRIT, and the Consensus on Exercise Reporting Template statements), and the inclusion of nutritional counseling for all study participants. Despite our trial being focused on exercise, we also acknowledge the importance of nutritional support in managing PTMS and made available a consultation with a dietitian as part of the educational component of our research.

We would also like to acknowledge the potential limitations of the study methodology. Given the complex profile of some of the patients diagnosed with PTMS, a 12-week intervention may not be sufficient to elicit a noticeable improvement in the

metabolic profile of these patients. An extended intervention may have a greater impact on PTMS risk factors, insulin resistance, and physical function and will be considered in a future RCT. There is a potential risk of bias in our research given the lack of blinding of the exercise professional and study participants (double-blinding is often not achievable in exercise intervention studies) [65]. However, our study will aim to satisfy other aspects of the Revised Cochrane risk-of-bias tool for randomized trials [66]. Furthermore, we excluded patients diagnosed with unstable cardiovascular or neuromuscular disease to ensure participant safety and understand that this may reduce some of the generalizability across transplant recipients. We are also aware of the challenges to exercise training in this population (eg, osteoporosis and diabetes mellitus), and although we designed an exercise protocol achievable by most transplant patients, personalized adaptations of the intervention cannot be avoided and represent a key feature of our protocol. Finally, to facilitate the assessment of patients during the COVID-19 pandemic, we took steps to adopt our assessment protocol to provide participants with the possibility of performing study assessments from the participant's home environment and bloodwork locally at community laboratories. We would like to acknowledge the potential threat to consistency in our assessment protocols, given that some assessments may be conducted on-site, whereas others will be performed remotely in the participant's home.

Conclusions

Exercise has become an important pillar in the management of comorbidities associated with LTx and OLT. A home-based exercise program may prove to be an effective posttransplant strategy for improving physical function and the metabolic profile of transplant recipients. Characterizing the feasibility, adherence, and effect estimates of home-based exercise training constitutes the first step in the promotion of a healthy lifestyle in transplant recipients and the establishment of long-term sustained change. This will be the first study to investigate the effects of exercise on PTMS risk factors, self-efficacy, and HRQL. The results of this trial can provide a greater understanding of behavioral strategies aimed at increasing exercise and physical activity in LTx and OLT recipients at risk of PTMS. Given the lack of patient education resources in the LTx and OLT populations, we hope that our results will provide greater insights regarding patients' exercise preferences and exercise modalities to create patient-directed materials aimed at promoting healthy living after transplantation. Our research will create opportunities for future collaborations and initiatives to gain greater insights into the mechanisms of PTMS and reduce its high prevalence through lifestyle interventions.

Acknowledgments

The research team would like to acknowledge the collaboration of study participants and the clinical staff at the University Health Network. This research is funded by the Canadian Donation Transplant Research Program, the Godfrey S Pettit Block Term Grant from the University of Toronto, and the Sandra Faire and Ivan Fecan Professorship in rehabilitation medicine. The study sponsors had no role in the design of the study protocol and will have no involvement in the data collection process, data analysis, interpretation of results, or manuscript preparation. This study is sponsored by the University Health Network.

Authors' Contributions

DR, MB, BLS, and DSM conceived the study and participated in the design of the protocol and the development of the intervention. DR, MB, BLS, and DSM are grant holders. EH provides statistical expertise in clinical trial design and conducting the primary statistical analysis. SN coordinates the study under the direct supervision of DR. SN, ECP, and JS are responsible for the implementation of the study. ECP and JS supervise the home-based exercise program. DR, SN, and ECP helped draft this manuscript. EM provides dietary counseling to study participants. BLS, LW, CT, NS, MA, MW, JP, AA, SD, SH, KP, and MB provided expert advice and reviewed the intervention protocol. All authors were involved in planning and writing the study protocol and approved the final manuscript. The plan is for all authors to be involved in the final manuscript preparation and submission process.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer-review report from the Canadian Donation Transplant Research Program competition.
[PDF File (Adobe PDF File), 277 KB - resprot_v11i3e35700_app1.pdf]

References

1. Thoefner LB, Rostved AA, Pommergaard H, Rasmussen A. Risk factors for metabolic syndrome after liver transplantation: a systematic review and meta-analysis. Transplant Rev (Orlando) 2018 Jan;32(1):69-77. [doi: 10.1016/j.trre.2017.03.004] [Medline: 28501338]
2. Savioli G, Surbone S, Giovi I, Salinaro F, Preti P, Meloni F, et al. Early development of metabolic syndrome in patients subjected to lung transplantation. Clin Transplant 2013;27(3):E237-E243. [doi: 10.1111/ctr.12098] [Medline: 23414418]

3. Bhat V, Tazari M, Watt KD, Bhat M. New-onset diabetes and preexisting diabetes are associated with comparable reduction in long-term survival after liver transplant: a machine learning approach. Mayo Clin Proc 2018 Dec;93(12):1794-1802. [doi: 10.1016/j.mayocp.2018.06.020] [Medline: 30522594]
4. Davidson J, Wilkinson A, Dantal J, Dotta F, Haller H, Hernández D, International Expert Panel. New-onset diabetes after transplantation: 2003 International consensus guidelines. Proceedings of an international expert panel meeting. Barcelona, Spain, 19 February 2003. Transplantation 2003 May 27;75(10 Suppl):SS3-S24. [doi: 10.1097/01.TP.0000069952.49242.3E] [Medline: 12775942]
5. Moon JI, Barbeito R, Faradji RN, Gaynor JJ, Tzakis AG. Negative impact of new-onset diabetes mellitus on patient and graft survival after liver transplantation: long-term follow up. Transplantation 2006 Dec 27;82(12):1625-1628. [doi: 10.1097/01.tp.0000250361.60415.96] [Medline: 17198248]
6. Painter P, Ward K, Nelson RD. Self-reported physical activity in patients with end stage renal disease. Nephrol Nurs J 2011;38(2):139-47; quiz 148. [Medline: 21520692]
7. Aravinthan AD, Fateen W, Doyle AC, Venkatachalapathy SV, Issachar A, Galvin Z, et al. The impact of preexisting and post-transplant diabetes mellitus on outcomes following liver transplantation. Transplantation 2019 Dec;103(12):2523-2530. [doi: 10.1097/TP.0000000000002757] [Medline: 30985734]
8. Park J, Chohan K, Nourouzpour S, Bhat M, Santa Mina D, Singer LG, et al. Clinical implications of metabolic risk factors in lung transplant recipients. Am J Resp Critical Care 2021 May;203:A4138. [doi: 10.1164/ajrccm-conference.2021.203.1_MeetingAbstracts.A4138]
9. Gustaw T, Schoo E, Barbalinardo C, Rodrigues N, Zameni Y, Motta VN, et al. Physical activity in solid organ transplant recipients: participation, predictors, barriers, and facilitators. Clin Transplant 2017 Apr;31(4). [doi: 10.1111/ctr.12929] [Medline: 28185297]
10. van Adrichem EJ, van de Zande SC, Dekker R, Verschuuren EA, Dijkstra PU, van der Schans CP. Perceived barriers to and facilitators of physical activity in recipients of solid organ transplantation, a qualitative study. PLoS One 2016;11(9):e0162725 [FREE Full text] [doi: 10.1371/journal.pone.0162725] [Medline: 27622291]
11. van Adrichem EJ, Dekker R, Krijnen WP, Verschuuren EA, Dijkstra PU, van der Schans CP. Physical activity, sedentary time, and associated factors in recipients of solid-organ transplantation. Phys Ther 2018 Aug 01;98(8):646-657. [doi: 10.1093/ptj/pzy055] [Medline: 29757444]
12. Janaudis-Ferreira T, Mathur S, Deliva R, Howes N, Patterson C, Räkel A, et al. Exercise for solid organ transplant candidates and recipients: a joint position statement of the Canadian society of transplantation and CAN-RESTORE. Transplantation 2019 Sep;103(9):e220-e238. [doi: 10.1097/TP.0000000000002806] [Medline: 31461743]
13. Trájer E, Bosnyák E, Komka ZS, Kováts T, Protzner A, Szmodis M, et al. Retrospective study of the Hungarian national transplant team's cardiorespiratory capacity. Transplant Proc 2015;47(6):1600-1604. [doi: 10.1016/j.transproceed.2015.02.022] [Medline: 26293020]
14. Garcia AM, Veneroso CE, Soares DD, Lima AS, Correia MI. Effect of a physical exercise program on the functional capacity of liver transplant patients. Transplant Proc 2014;46(6):1807-1808. [doi: 10.1016/j.transproceed.2014.05.023] [Medline: 25131042]
15. Didsbury M, McGee RG, Tong A, Craig JC, Chapman JR, Chadban S, et al. Exercise training in solid organ transplant recipients: a systematic review and meta-analysis. Transplantation 2013 Mar 15;95(5):679-687. [doi: 10.1097/TP.0b013e31827a3d3e] [Medline: 23364480]
16. Moya-Nájera D, Moya-Herraiz A, Compte-Torrero L, Hervás D, Borreani S, Calatayud J, et al. Combined resistance and endurance training at a moderate-to-high intensity improves physical condition and quality of life in liver transplant patients. Liver Transpl 2017 Oct;23(10):1273-1281. [doi: 10.1002/lt.24827] [Medline: 28749550]
17. Mitchell MJ, Baz MA, Fulton MN, Lisor CF, Braith RW. Resistance training prevents vertebral osteoporosis in lung transplant recipients. Transplantation 2003 Aug 15;76(3):557-562. [doi: 10.1097/01.TP.0000076471.25132.52] [Medline: 12923444]
18. Langer D, Burtin C, Schepers L, Ivanova A, Verleden G, Decramer M, et al. Exercise training after lung transplantation improves participation in daily activity: a randomized controlled trial. Am J Transplant 2012 Jun;12(6):1584-1592 [FREE Full text] [doi: 10.1111/j.1600-6143.2012.04000.x] [Medline: 22390625]
19. Basha MA, Mowafy ZE, Morsy EA. Sarcopenic obesity and dyslipidemia response to selective exercise program after liver transplantation. Egyptian J Med Human Genetics 2015 Jul;16(3):263-268. [doi: 10.1016/j.ejmhg.2014.12.006]
20. Mathur S, Janaudis-Ferreira T, Wickerson L, Singer LG, Patcai J, Rozenberg D, et al. Meeting report: consensus recommendations for a research agenda in exercise in solid organ transplantation. Am J Transplant 2014 Oct;14(10):2235-2245 [FREE Full text] [doi: 10.1111/ajt.12874] [Medline: 25135579]
21. Vieux L, Simcox AA, Mediouni Z, Wild P, Koller M, Swiss Transplant Cohort Study, et al. Predictors of return to work 12 months after solid organ transplantation: results from the swiss transplant cohort study. J Occup Rehabil 2019 Jun;29(2):462-471. [doi: 10.1007/s10926-018-9804-8] [Medline: 30145704]
22. Radtke T, Haile SR, Dressel H, Benden C. COVID-19 pandemic restrictions continuously impact on physical activity in adults with cystic fibrosis. PLoS One 2021;16(9):e0257852 [FREE Full text] [doi: 10.1371/journal.pone.0257852] [Medline: 34555108]

23. Barnett A, Campbell KL, Mayr HL, Keating SE, Macdonald GA, Hickman IJ. Liver transplant recipients' experiences and perspectives of a telehealth-delivered lifestyle programme: a qualitative study. J Telemed Telecare 2021 Oct;27(9):590-598. [doi: 10.1177/1357633X19900459] [Medline: 31986966]
24. Wu Y, Chien C, Chou N, Wang S, Lai J, Wu Y. Efficacy of a home-based exercise program for orthotopic heart transplant recipients. Cardiology 2008;111(2):87-93. [doi: 10.1159/000119695] [Medline: 18376119]
25. Painter PL, Hector L, Ray K, Lynes L, Dibble S, Paul SM, et al. A randomized trial of exercise training after renal transplantation. Transplantation 2002 Jul 15;74(1):42-48. [doi: 10.1097/00007890-200207150-00008] [Medline: 12134097]
26. Wickerson L, Helm D, Gottesman C, Rozenberg D, Singer LG, Keshavjee S, et al. Telerehabilitation for lung transplant candidates and recipients during the COVID-19 pandemic: program evaluation. JMIR Mhealth Uhealth 2021 Jun 17;9(6):e28708 [FREE Full text] [doi: 10.2196/28708] [Medline: 34048354]
27. Vivodtzev I, Pison C, Guerrero K, Mezin P, Maclet E, Borel J, et al. Benefits of home-based endurance training in lung transplant recipients. Respir Physiol Neurobiol 2011 Jul 31;177(2):189-198. [doi: 10.1016/j.resp.2011.02.005] [Medline: 21333761]
28. Fuller LM, Button B, Tarrant B, Steward R, Bennett L, Snell G, et al. Longer versus shorter duration of supervised rehabilitation after lung transplantation: a randomized trial. Arch Phys Med Rehabil 2017 Feb;98(2):220-6.e3. [doi: 10.1016/j.apmr.2016.09.113] [Medline: 27697429]
29. Choi J, Hergenroeder AL, Burke L, Dabbs AD, Morrell M, Saptono A, et al. Delivering an in-home exercise program via telerehabilitation: a pilot study of lung transplant go (LTGO). Int J Telerehabil 2016;8(2):15-26 [FREE Full text] [doi: 10.5195/ijt.2016.6201] [Medline: 28775798]
30. Krasnoff JB, Vintro AQ, Ascher NL, Bass NM, Paul SM, Dodd MJ, et al. A randomized trial of exercise and dietary counseling after liver transplantation. Am J Transplant 2006 Aug;6(8):1896-1905 [FREE Full text] [doi: 10.1111/j.1600-6143.2006.01391.x] [Medline: 16889545]
31. Tomás MT, Santa-Clara H, Bruno PM, Monteiro E, Carrolo M, Barroso E, et al. The impact of exercise training on liver transplanted familial amyloidotic polyneuropathy (FAP) patients. Transplantation 2013 Jan 27;95(2):372-377. [doi: 10.1097/TP.0b013e31827220e7] [Medline: 23211558]
32. Hickman IJ, Hannigan AK, Johnston HE, Elvin-Walsh L, Mayr HL, Staudacher HM, et al. Telehealth-delivered, cardioprotective diet and exercise program for liver transplant recipients: a randomized feasibility study. Transplant Direct 2021 Mar;7(3):e667 [FREE Full text] [doi: 10.1097/TXD.0000000000001118] [Medline: 33564717]
33. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, Del Coso J, Leyton-Román M, Luque-Casado A, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021 Feb 19;18(4):2023 [FREE Full text] [doi: 10.3390/ijerph18042023] [Medline: 33669679]
34. Mishali M, Omer H, Heymann AD. The importance of measuring self-efficacy in patients with diabetes. Fam Pract 2011 Feb;28(1):82-87 [FREE Full text] [doi: 10.1093/fampra/cmq086] [Medline: 21047940]
35. Zheng X, Yu H, Qiu X, Chair SY, Wong EM, Wang Q. The effects of a nurse-led lifestyle intervention program on cardiovascular risk, self-efficacy and health promoting behaviours among patients with metabolic syndrome: randomized controlled trial. Int J Nurs Stud 2020 Sep;109:103638 [FREE Full text] [doi: 10.1016/j.ijnurstu.2020.103638] [Medline: 32553996]
36. Johnson JL, Slentz CA, Houmard JA, Samsa GP, Duscha BD, Aiken LB, et al. Exercise training amount and intensity effects on metabolic syndrome (from Studies of a Targeted Risk Reduction Intervention through Defined Exercise). Am J Cardiol 2007 Dec 15;100(12):1759-1766 [FREE Full text] [doi: 10.1016/j.amjcard.2007.07.027] [Medline: 18082522]
37. García-Pajares F, Peñas-Herrero I, Sánchez-Ocaña R, Torrres-Yuste R, Cimavilla-Román M, Carbajo-López A, et al. Metabolic syndrome after liver transplantation: five-year prevalence and risk factors. Transplant Proc 2016 Nov;48(9):3010-3012. [doi: 10.1016/j.transproceed.2016.07.038] [Medline: 27932133]
38. Chan A, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
39. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010 Jun 1;152(11):726-732. [doi: 10.7326/0003-4819-152-11-201006010-00232] [Medline: 20335313]
40. Slade SC, Dionne CE, Underwood M, Buchbinder R. Consensus on Exercise Reporting Template (CERT): explanation and elaboration statement. Br J Sports Med 2016 Dec 05;50(23):1428-1437. [doi: 10.1136/bjsports-2016-096651] [Medline: 27707738]
41. Wickerson L, Rozenberg D, Janaudis-Ferreira T, Deliva R, Lo V, Beauchamp G, et al. Physical rehabilitation for lung transplant candidates and recipients: an evidence-informed clinical approach. World J Transplant 2016 Sep 24;6(3):517-531 [FREE Full text] [doi: 10.5500/wjt.v6.i3.517] [Medline: 27683630]
42. Canadian 24-hour movement guidelines: an integration of physical activity, sedentary behaviour, and sleep. 24-Hour Movement Guidelines. URL: https://csepguidelines.ca/ [accessed 2022-02-14]

43. Santa Mina D, Hilton WJ, Matthew AG, Awasthi R, Bousquet-Dion G, Alibhai SM, et al. Prehabilitation for radical prostatectomy: a multicentre randomized controlled trial. Surg Oncol 2018 Jun;27(2):289-298. [doi: 10.1016/j.suronc.2018.05.010] [Medline: 29937184]
44. Santa MD, Alibhai SM, Matthew AG, Guglietti CL, Pirbaglou M, Trachtenberg J, et al. A randomized trial of aerobic versus resistance exercise in prostate cancer survivors. J Aging Phys Act 2013 Oct;21(4):455-478. [Medline: 23238110]
45. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [Medline: 7154893]
46. Gellish RL, Goslin BR, Olson RE, McDonald A, Russi GD, Moudgil VK. Longitudinal modeling of the relationship between age and maximal heart rate. Med Sci Sports Exerc 2007 May;39(5):822-829. [doi: 10.1097/mss.0b013e31803349c6] [Medline: 17468581]
47. Grundy SM, Brewer HB, Cleeman JI, Smith SC, Lenfant C, American HA, et al. Definition of metabolic syndrome: report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. Circulation 2004 Jan 27;109(3):433-438 [FREE Full text] [doi: 10.1161/01.CIR.0000111245.75752.C6] [Medline: 14744958]
48. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. J Clin Epidemiol 1993 Feb;46(2):153-162. [Medline: 8437031]
49. 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]
50. Segal-Isaacson CJ, Wylie-Rosett J, Gans KM. Validation of a short dietary assessment questionnaire: the Rapid Eating and Activity Assessment for Participants short version (REAP-S). Diabetes Educ 2004;30(5):774, 776, 778 passim. [doi: 10.1177/014572170403000512] [Medline: 15510530]
51. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
52. Singer LG, Chowdhury NA, Faughnan ME, Granton J, Keshavjee S, Marras TK, et al. Effects of recipient age and diagnosis on health-related quality-of-life benefit of lung transplantation. Am J Respir Crit Care Med 2015 Oct 15;192(8):965-973. [doi: 10.1164/rccm.201501-0126OC] [Medline: 26131729]
53. Everett B, Salamonson Y, Davidson PM. Bandura's exercise self-efficacy scale: validation in an Australian cardiac rehabilitation setting. Int J Nurs Stud 2009 Jun;46(6):824-829. [doi: 10.1016/j.ijnurstu.2009.01.016] [Medline: 19261281]
54. 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]
55. Rozenberg D, Singer LG, Herridge M, Goldstein R, Wickerson L, Chowdhury NA, et al. Evaluation of skeletal muscle function in lung transplant candidates. Transplantation 2017 Sep;101(9):2183-2191. [doi: 10.1097/TP.0000000000001754] [Medline: 28376036]
56. Wickerson L, Mathur S, Singer LG, Brooks D. Physical activity levels early after lung transplantation. Phys Ther 2015 Apr;95(4):517-525. [doi: 10.2522/ptj.20140173] [Medline: 25504488]
57. Laryea M, Watt KD, Molinari M, Walsh MJ, McAlister VC, Marotta PJ, et al. Metabolic syndrome in liver transplant recipients: prevalence and association with major vascular events. Liver Transpl 2007 Aug;13(8):1109-1114 [FREE Full text] [doi: 10.1002/lt.21126] [Medline: 17663411]
58. Li C, Xu J, Qin W, Hu Y, Lu H. Meta-analysis of the effects of exercise training on markers of metabolic syndrome in solid organ transplant recipients. Prog Transplant 2018 Sep;28(3):278-287. [doi: 10.1177/1526924818781576] [Medline: 29898634]
59. Trojetto T, Elliott RJ, Rashid S, Wong S, Dlugosz K, Helm D, et al. Availability, characteristics, and barriers of rehabilitation programs in organ transplant populations across Canada. Clin Transplant 2011;25(6):E571-E578. [doi: 10.1111/j.1399-0012.2011.01501.x] [Medline: 21955056]
60. Langer A, Gassner L, Flotz A, Hasenauer S, Gruber J, Wizany L, et al. How COVID-19 will boost remote exercise-based treatment in Parkinson's disease: a narrative review. NPJ Parkinsons Dis 2021 Mar 08;7(1):25 [FREE Full text] [doi: 10.1038/s41531-021-00160-3] [Medline: 33686074]
61. Ai X, Yang J, Lin Z, Wan X. Mental health and the role of physical activity during the COVID-19 pandemic. Front Psychol 2021;12:759987 [FREE Full text] [doi: 10.3389/fpsyg.2021.759987] [Medline: 34744938]
62. Narici M, Vito GD, Franchi M, Paoli A, Moro T, Marcolin G, et al. Impact of sedentarism due to the COVID-19 home confinement on neuromuscular, cardiovascular and metabolic health: physiological and pathophysiological implications and recommendations for physical and nutritional countermeasures. Eur J Sport Sci 2020 May 12;21(4):614-635. [doi: 10.1080/17461391.2020.1761076]
63. Li J, Wei D, Liu S, Li M, Chen X, Chen L, et al. Efficiency of an mHealth app and chest-wearable remote exercise monitoring intervention in patients with type 2 diabetes: a prospective, multicenter randomized controlled trial. JMIR Mhealth Uhealth 2021 Feb 09;9(2):e23338 [FREE Full text] [doi: 10.2196/23338] [Medline: 33560244]
64. Ramadi A, Haennel RG, Stone JA, Arena R, Threlfall TG, Hitt E, et al. The sustainability of exercise capacity changes in home versus center-based cardiac rehabilitation. J Cardiopulm Rehabil Prev 2015;35(1):21-28. [doi: 10.1097/HCR.0000000000000084] [Medline: 25313452]

65. Karanicolas PJ, Farrokhyar F, Bhandari M. Practical tips for surgical research: blinding: who, what, when, why, how? Can J Surg 2010 Oct;53(5):345-348 [FREE Full text] [Medline: 20858381]
66. Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
HRQL: health-related quality of life
LTx: lung transplant
OLT: orthotopic liver transplant
PTMS: posttransplant metabolic syndrome
RCT: randomized controlled trial
RPE: rating of perceived exertion
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by T Derrick; submitted 16.12.21; this is a non–peer-reviewed article;accepted 21.12.21; published 23.03.22.

Please cite as:
Rozenberg D, Santa Mina D, Nourouzpour S, Camacho Perez E, Stewart BL, Wickerson L, Tsien C, Selzner N, Shore J, Aversa M, Woo M, Holdsworth S, Prevost K, Park J, Azhie A, Huszti E, McLeod E, Dales S, Bhat M
Feasibility of a Home-Based Exercise Program for Managing Posttransplant Metabolic Syndrome in Lung and Liver Transplant Recipients: Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2022;11(3):e35700
URL: https://www.researchprotocols.org/2022/3/e35700
doi:10.2196/35700
PMID:35319467

Protocol

Impact-Oriented Dialogue for Culturally Safe Adolescent Sexual and Reproductive Health in Bauchi State, Nigeria: Protocol for a Codesigned Pragmatic Cluster Randomized Controlled Trial

Anne Cockcroft[1,2], MBBS, MD; Khalid Omer[2], MBBS, MSc; Yagana Gidado[3], HND, RN, RM; Rilwanu Mohammed[4], MBBS; Loubna Belaid[1], PhD; Umaira Ansari[2], BA, LLB; Claudia Mitchell[5], BA, MA, PhD; Neil Andersson[1,2], MD, PhD

[1]Community Information for Empowerment and Transparency-Participatory Research at McGill, Department of Family Medicine, McGill University, Montreal, QC, Canada

[2]Centro de Investigacion de Enfermedades Tropicales, Universidad Autonoma de Guerrero, Acapulco, Mexico

[3]Federation of Muslim Women's Associations of Nigeria, Bauchi, Nigeria

[4]Bauchi State Primary Health Care Development Agency, Bauchi, Nigeria

[5]Department of Integrated Studies in Education, McGill University, Montreal, QC, Canada

Corresponding Author:
Anne Cockcroft, MBBS, MD
Community Information for Empowerment and Transparency-Participatory Research at McGill
Department of Family Medicine
McGill University
5858 Côte-des-Neiges, Suite 300
Montreal, QC, H3S 1Z1
Canada
Phone: 1 450 626 8432
Email: anne.cockcroft@mcgill.ca

Abstract

Background: Adolescents (10-19 years) are a big segment of the Nigerian population, and they face serious risks to their health and well-being. Maternal mortality is very high in Nigeria, and rates of pregnancy and maternal deaths are high among female adolescents. Rates of HIV infection are rising among adolescents, gender violence and sexual abuse are common, and knowledge about sexual and reproductive health risks is low. Adolescent sexual and reproductive health (ASRH) indicators are worse in the north of the country.

Objective: In Bauchi State, northern Nigeria, the project will document the nature and extent of ASRH outcomes and risks, discuss the findings and codesign solutions with local stakeholders, and measure the short-term impact of the discussions and proposed solutions.

Methods: The participatory research project is a sequential mixed-methods codesign of a pragmatic cluster randomized controlled trial. Focus groups of local stakeholders (female and male adolescents, parents, traditional and religious leaders, service providers, and planners) will identify local priority ASRH concerns. The same stakeholder groups will map their knowledge of factors causing these concerns using the fuzzy cognitive mapping (FCM) technique. Findings from the maps and a scoping review will inform the contextualization of survey instruments to collect information about ASRH from female and male adolescents and parents in households and from local service providers. The survey will take place in 60 Bauchi communities. Adolescents will cocreate materials to share the findings from the maps and survey. In 30 communities, randomly allocated, the project will engage adolescents and other stakeholders in households, communities, and services to discuss the evidence and to design and implement culturally acceptable actions to improve ASRH. A follow-up survey in communities with and without the intervention will measure the short-term impact of these discussions and actions. We will also evaluate the intervention process and use narrative techniques to assess its impact qualitatively.

Results: Focus groups to explore ASRH concerns of stakeholders began in October 2021. Baseline data collection in the household survey is expected to take place in mid-2022. The study was approved by the Bauchi State Health Research Ethics Committee, approval number NREC/03/11/19B/2021/03 (March 1, 2021), and by the Faculty of Medicine and Health Sciences Institutional Review Board McGill University (September 13, 2021).

Conclusions: Evidence about factors related to ASRH outcomes in Nigeria and implementation and testing of a dialogic intervention to improve these outcomes will fill a gap in the literature. The project will document and test the effectiveness of a participatory approach to ASRH intervention research.

Trial Registration: ISRCTN Registry ISRCTN18295275; https://www.isrctn.com/ISRCTN18295275

International Registered Report Identifier (IRRID): DERR1-10.2196/36060

(JMIR Res Protoc 2022;11(3):e36060) doi:10.2196/36060

KEYWORDS

adolescents; sexual and reproductive health; participatory research; mixed methods research; dialogic intervention; co-design; cultural safety; Nigeria

Introduction

Background

Pregnancy and childbirth complications are the leading cause of death in girls aged 15 to 19 years in low- and middle-income countries [1]. Adolescent sexual and reproductive health (ASRH) risks in sub-Saharan Africa are extremely high, with important intersectional inequalities related to gender, education, economic status, and urban/rural residence [2].

One-fifth of Nigeria's population are adolescents aged between 10 and 19 years, and the government recognizes the need to protect their sexual and reproductive health [3]. Adolescent girls are at higher risk. By the age of 18 years, 43% of Nigerian women are married, 56% have had sexual debut, and 29% have given birth [4]. In 2018, 19% of girls aged 15 to 19 years had begun childbearing, and 40% of deaths in this age group were related to pregnancy and childbirth [4]. Poor, uneducated women in rural areas are more likely to marry and to give birth early [4-6]. Pregnant teenagers may risk unsafe abortions [7,8]. Nigerian adolescents also face risks of HIV and gender violence [9-11]. Nearly all youth aged 15 to 19 years have heard of AIDS, but few have functional knowledge about prevention [4]. Nationally, 18% of women aged 15 to 19 years reported physical violence in the last year and 5% during their last pregnancy [4].

To protect adolescents, the Government of Nigeria introduced the Family Life and HIV Education program in schools in 2003. The program only reaches a small proportion of in-school adolescents, and teachers are uncomfortable dealing with the topics [12,13]. Programs reach even fewer out-of-school youth [14].

Small studies describe the knowledge and experiences of Nigerian adolescents of sexual and reproductive health risks [15-21]. They rarely use available reproductive health services [22,23]. A recent scoping review of 1302 articles on ASRH in sub-Saharan Africa found studies most frequently focused on HIV, sexual behavior, and access to services. Few included younger adolescents (10-14 years), 53% used quantitative methods only, 44% were cross-sectional studies, and only 13% used mixed methods [24]. Many ASRH programs are ineffective [25], and there is a need to measure the impact of interventions and to engage adolescents in designing and implementing programs [26]. We are not aware of published studies that involved Nigerian adolescents in codesigning interventions to address their own sexual and reproductive health needs.

This project aims to fill this evidence gap. In a participatory approach, we will collect qualitative and quantitative evidence about ASRH and share it with adolescents and other stakeholders, who will codesign interventions to improve ASRH. The work builds on an existing strong collaboration between the research team and the state government. This collaboration completed two projects in Bauchi State under the Innovating for Maternal and Child Health in Africa initiative [27]: a trial of universal home visits to improve maternal and early childhood health [28], and a study of causes and prevention of short birth interval [29]. These projects included married adolescent girls aged 14 to 19 years.

Conceptual Framework

The project considers influences on the sexual and reproductive health of adolescent girls and boys at individual, family, and broader structural levels [30]. At the individual and family level, the CASCADA (Conscious knowledge, Attitude, Subjective norms, intention to Change, Agency to change, Discussion of issues, and Action to change) results chain [31,32] expands the knowledge-attitudes-practice model [33], criticized for its lack of detail between attitudes and practice [34]. A modified theory of planned behavior [35], CASCADA is an acronym for a partial order of intermediate outcomes between knowledge and action. We have used CASCADA in resource-poor settings to support analysis of studies of health behaviors [36,37], in the design of interventions [38], and as a framework for analysis of intervention effects [39,40].

Many risks for ASRH are structural; potential solutions need to address factors that constrain individual choice, including patriarchy and unhelpful gender norms [30,41-43]. We developed a concept of *choice disability* among marginalized young women at high risk of gender violence and HIV in Southern Africa [44], showed its association with HIV risk [45], and implemented interventions combining individual empowerment, more responsive services, and the creation of an enabling environment to support individual choices [46,47]. Choice disability likely also applies to at-risk adolescents in Bauchi, although the cultural context is different.

The concept of *cultural safety* [48] requires that the people concerned—in this case, adolescents and other stakeholders in Bauchi—decide whether a service or intervention respects their cultural identity and values [49]. Beyond being a moral and ethical issue, cultural safety makes an intervention more likely to succeed.

Participatory research partners with and values the knowledge of stakeholders [50]. The project in Bauchi will prioritize and systematize the concerns and knowledge of local stakeholders, especially adolescents, about ASRH risks and engage them to codesign and implement solutions that work within their cultural context. Our recent work on *kunika* (short birth interval) in Bauchi demonstrates how an approach of integrated knowledge translation and exchange can result in locally identified and culturally attuned interventions and communication materials for a sensitive issue (U Ansari, unpublished data, January 2022) [29,51].

Research Goal and Objectives

Goal

Codesign culturally appropriate interventions to improve ASRH and pilot their implementation in Bauchi State, Nigeria.

Objectives

1. Explore priority stakeholder concerns about ASRH in Bauchi, collate the knowledge of female and male adolescents and other stakeholders about causes and protective factors for these concerns, and compare their knowledge with documented associations in the literature.
2. Quantify ASRH-related knowledge, attitudes, experiences, and behaviors of female and male adolescents, parents, and service providers, using data collection instruments informed by the collated local knowledge and literature review.
3. Engage adolescents, parents, service providers, and decision-makers in dialogue about the evidence on ASRH outcomes and causes, to identify and implement locally appropriate interventions at different levels to improve ASRH.
4. Evaluate the intervention process and measure impact on ASRH knowledge, attitudes, experiences, and behaviors of female and male adolescents and other stakeholders quantitatively and qualitatively.

Study Design

The participatory research project is a sequential mixed-methods codesign of a pragmatic cluster randomized controlled trial. It will begin with qualitative data collection and a scoping review, supporting codesign of the dialogic intervention to share local findings about ASRH risks and codesign solutions with adolescents and other stakeholders. The impact evaluation of the intervention will use both quantitative and qualitative methods.

Methods

Setting

The study will take place in communities in the Toro Local Government Area of Bauchi State in the North East of Nigeria. Ninety-five percent of the state population are Muslim. Health and education indicators are worse than elsewhere in Nigeria. Only 26% of women and 48% of men aged 15 to 49 years are literate in Bauchi State, compared with 53% and 72% nationally [4]. Women in Bauchi start childbearing particularly young; 41% of girls aged 15 to 19 years have given birth or are pregnant [4]. In a recent program of home visits in Bauchi State [28], 56% of 7282 women aged 14 to 49 years were pregnant during one year, and 38% of those aged 14 to 19 years. Gender violence is common in Bauchi. In one study, 23% of women aged 14 to 19 years experienced domestic violence in the last year and 15% during their last pregnancy [52].

Qualitative Exploration of ASRH Concerns and Scoping Review

Focus Group Discussions

Local researchers will facilitate focus group discussions with female and male adolescents and other stakeholders about priority concerns for ASRH. In each participating community, they will conduct 10 groups: 5 adolescent groups, 2 groups of parents of adolescents (male and female), 2 groups of traditional and religious leaders (male and female), and 1 group of service providers (mixed male and female). The adolescent groups will comprise 3 female groups (10-14 years, 15-19 years unmarried, and 15-19 years married) and 2 male groups (10-14 years and 15-19 years). All groups will include adolescents in and out of school. Female team members will facilitate female groups, and male team members will facilitate male groups. Young facilitators (less than 25 years) will facilitate the groups of adolescents. The 6 participating communities will be 2 urban, 2 rural, and 2 rural-remote, spread across 6 wards in Toro Local Government Area. The research team will also conduct 2 groups with male and female ward-level leaders in each of the 6 wards, 1 group with government officers in the local government area, and 3 groups at the state level, with health planners, traditional and religious leaders, and nongovernment organizations (NGOs). In total, there will be 76 focus groups and 380 to 456 participants (5-6 per group).

Analysis

An inductive thematic analysis will identify priority ASRH concerns [53], using the approach of Braun and Clarke [54] and applying criteria for trustworthiness proposed by Guba [55,56]. Without pre-empting the findings, we expect that the priority ASRH concerns will likely include adverse effects of early pregnancy, sexually transmitted infections, gender violence, sexual abuse, and poor access to responsive services.

Fuzzy Cognitive Mapping

Fuzzy cognitive mapping (FCM) is the graphic representation of knowledge about causality in a system [57,58]. FCM depicts factors that stakeholders consider to be causes of an outcome, in this case, adverse ASRH outcomes. The stakeholders rate the strength of associations between these factors and the outcome. We will compare maps between different stakeholder groups using the mathematical technique of transitive closure [59,60].

Local fieldworkers will facilitate stakeholder groups to create maps of the factors they believe cause the priority ASRH concerns identified in the focus groups. The stakeholders will estimate the strength of each association in the map. They will use a scale of 1 to 5, with 5 representing the strongest association. The FCM groups will be the same as the focus groups.

Analysis

We will digitize the maps using YEd software [61]. A thematic analysis [53,54] of factors in the maps will create broader themes. Transitive closure analysis [59,60] will identify the most influential factors. The analysis will give the knowledge of adolescent girls and boys at least equal weight to that of other stakeholders. We will compare maps of different stakeholder groups and maps from stakeholder groups with the map from the scoping review [62]. Causes of adverse ASRH outcomes identified on the maps will inform the contextualization of the survey instruments. Based on our experience of FCM of causes of short birth interval in Bauchi, we expect they will include causes at individual, family, community, and service levels [29].

Scoping Review of the Literature

Systematic reviews have examined the effectiveness of interventions to improve ASRH, including in low- and middle-income countries [25,26,63]. A 2020 review of studies and demographic and health survey data from sub-Saharan Africa confirmed high ASRH risks, with variation between countries [2]. There remains a need to collate quantitative and qualitative evidence of associations with ASRH in low- and middle-income countries. Supported by a specialist librarian, we will conduct a scoping review of quantitative and qualitative studies of factors associated with ASRH outcomes in low- and middle-income countries. In addition to a standard reporting of the review, we will create a fuzzy cognitive map (see as follows) of factors related to ASRH outcomes, marking the strength of associations as odds ratios or regression coefficients, and compare the literature map with stakeholder-created maps [62,64,65].

Cluster Randomized Controlled Trial of Dialogic Intervention to Promote ASRH

Overview

The trial will begin with a baseline household survey in all study communities. Half the communities will participate in evidence-based dialogues to plan and implement local solutions to ASRH concerns. Normal health and other services will continue in all communities. A follow-up household survey in all communities will document the quantitative impact of the intervention on priority ASRH outcomes, and narratives of change will explore perceived experiences of the intervention.

Participants

The household sample for measurement of impact in intervention and control communities will comprise all adolescents and their parents in 100 households in each community. Eligible households will have at least one adolescent girl. We will adjust survey timing to ensure we reach in-school as well as out-of-school adolescents. No adolescent or parent who agrees to participate in the survey will be excluded. For the intervention, all adolescents, adults, service providers, and traditional and religious community leaders in the intervention communities will be eligible to participate in activities developed in each community after discussing the local evidence in dialogue groups.

The Intervention: SEPA

We have developed the socializing evidence for participatory action (SEPA) approach over 25 years [32] to support participatory action on health concerns [38,66-69]. The SEPA approach includes deliberative dialogue where small groups of stakeholders discuss local evidence and decide on actions to tackle a common concern [70]. SEPA is itself an intervention, and we can measure its impact. Stakeholders decide what actions to take; these vary from place to place, but the protocol of sharing and discussing evidence can be standardized and randomized [71].

Evidence Materials

We will prepare summarized outputs from the transitive closure analysis of the cognitive maps [29]. With support from the research team, groups of approximately 15 adolescent girls and boys respectively, identified through a local NGO, the Federation of Muslim Women's Associations in Nigeria (FOMWAN), will produce materials to share evidence from the maps and baseline survey, including short video docudramas in the style of popular local soap operas. We recently codesigned with community groups video docudramas about short birth intervals in Bauchi (U Ansari, unpublished data, January 2022). We will work with the female and male adolescents and a local production company to prepare the video docudramas and other communication materials in about four months at the end of year two.

Implementation of the Intervention

SEPA will take place over 18 months, in years three and four. Local researchers and adolescents will share the findings about ASRH, including sharing the video docudramas, with groups in the 30 SEPA communities: adolescent girls, adolescent boys, male and female adults, community and religious leaders, and relevant service providers. The groups will plan and implement local actions to improve ASRH in their communities, engaging other adolescents and other stakeholders in these community actions. The research team will provide logistic and administrative support for the community groups and their actions. Research team members will visit the communities monthly to liaise with community leaders, document progress, and help to resolve challenges. Follow up will be virtual if necessary; we have successfully used cellular teleconferencing in Bauchi communities [72]. Adolescent girls and boys will meet separately; we will explore ways to share ideas between female and male groups, for example, creating visual media such as cellphilms [73]. In combined dialogue groups, there is a risk that powerful stakeholders could silence the voice of adolescents, especially females. Adolescents and other stakeholders will initially meet separately; our trained facilitators will only convene combined groups if the adolescents are confident, and we will give them a protected space to speak in any combined groups. We expect that actions proposed by the SEPA groups will engage larger numbers of adolescents and other stakeholders. To support changes in service provision and community norms, we will include service providers and powerful stakeholders; some may change their views through their engagement. The research team will broker dialogues between adolescents, adults, and service providers in ways that

recognize the steep power gradients involved and take steps to ensure that the voices of adolescents, especially female adolescents, are heard and respected. About 500 adolescents will participate in community SEPA groups, and we expect perhaps 3000 to participate in community activities led by the SEPA groups.

Process Evaluation of the SEPA Intervention

The process evaluation will consider implementation, mechanisms, and context [74]. Towards the end of the implementation period, adolescents and other stakeholders implementing SEPA in each community will review their achievements and document the challenges they faced. Using reports from these meetings and documentation of the SEPA activities over the period, a researcher not involved in implementation will "score" the level of implementation in each community, based on level of participation in groups, planned activities, and activities that were implemented. This will include adolescent-led activities as well as the support or resistance they received from other stakeholders in the community and services. This implementation score will be a factor in the quantitative analysis of the impact of the intervention.

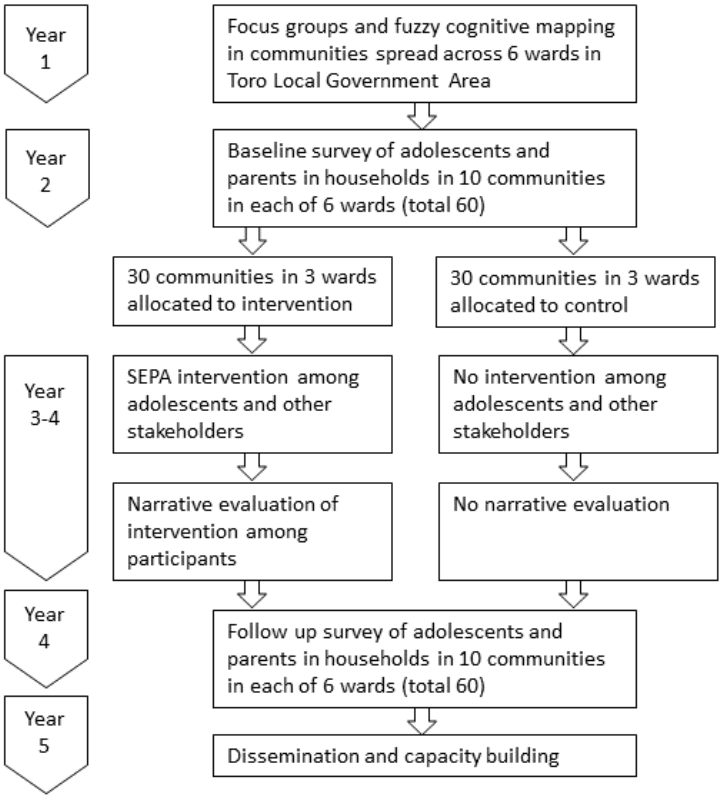
Outcomes

The initial focus groups and FCM with adolescents and other stakeholders will inform the specific priority ASRH outcomes to be addressed and measured. The outcomes are likely to include sexually transmitted diseases, experience and perpetration of physical and sexual violence, emotional distress, and use and experience of health services. Intermediate outcomes will cover steps in the CASCADA sequence. Questionnaires administered to adolescents and parents in the baseline household survey and the follow-up household survey 18 to 24 months later will collect data to measure the outcomes.

Timing

Figure 1 shows the timeline for the project, including the initial qualitative phase and the trial of the intervention.

Figure 1. Timeline of the project. SEPA: socializing evidence for participatory action.



Sample Size

The sample for the baseline and follow-up household survey will include 6000 adolescent girls and 4000 adolescent boys, plus 4000 mothers and 2500 fathers. The clinical trials simulator of Taylor and Bosch [75] estimates study power for the indicative outcome of the experience of violence by adolescent girls. With an expected prevalence of about 20% at baseline, and a sample of 3000 girls in 30 SEPA communities and 3000 in 30 non-SEPA communities, a 15% reduction of gender violence (to 17%) would be detectable with 80% power at the 5% level (κ=0.05). We anticipate a household response rate of around 95%. In a previous Bauchi household survey, the response rate was 96% (3.4% were not available, and 0.4% declined) [76].

Allocation of the Intervention

We will stratify the wards by size and proportion of urban communities, and an epidemiologist not involved in the fieldwork will randomly allocate (using a computer-generated random sequence) 3 of the 6 wards to receive the SEPA intervention. All the 30 study communities in the 3 SEPA wards will participate in SEPA. We do not expect any community to

decline to participate; they all participated in and strongly valued our recent home visits program [40,77].

Blinding

It is not possible to blind participating communities to the SEPA intervention. The dialogue groups and resulting community activities will be apparent and are intended to be. The interviewers for the household surveys will not be involved in supporting intervention activities. They may become aware of the intervention status of some communities when undertaking the follow-up survey; they will use a standard questionnaire in all communities, and there is no reason to believe they will conduct interviews differently in intervention and control communities.

Quantitative Data Collection and Analysis

The baseline survey will take place in year two of the project in 60 communities in the 6 wards of the Toro Local Government Area. In each community, a team of interviewers will cover a cluster of 100 households, recording questionnaire responses on android handsets using ODK software (Get ODK) [78]. They will upload records to a Cloud server via the cellular network, and we will download data sets for analysis. The analysis will use CIETmap open-source software (version 2.2.21; CIET group) [79], which interfaces with the R programming language (R Core Team). The household survey of adolescents and parents will provide data on the frequency of the specified ASRH outcomes and individual and family factors potentially related to these. Institutional reviews of health facilities serving the communities will provide data about potential supply-side determinants. Key informant interviews with community leaders and opinion-makers and community profiles will record factors at the community level potentially relevant to ASRH outcomes.

The follow-up survey will cover the same 60 communities as the baseline survey, but not necessarily the same households.

Data Collection Instruments

1. ASRH questionnaires for adolescents and adults will draw on existing validated instruments [80-82], contextualized by findings of the focus groups and FCM. They will include intermediate outcomes of knowledge, attitudes, and experience amenable to change stimulated by the intervention in a short time scale.
2. A guide for institutional reviews of health facilities will gather information on available ASRH services in one clinic per community
3. A questionnaire for key informants and a community profile proforma will enquire about factors relevant to ASRH at the community level. Key informants will include traditional and religious leaders, headteachers, health facility heads, and social workers.

Quantitative Data Analysis

Bivariate and then multivariate analysis of responses to the baseline survey will examine associations between ASRH outcomes and potential determinants at individual, family, and community/services level, using the Mantel-Haenszel procedure [83], with the Lamothe cluster adjustment [84]. The analysis will include gender as a key variable in the examination of ASRH outcomes and associations with these outcomes.

To measure the impact of the intervention, we will compare the pre-specified ASRH outcomes at follow-up between adolescents in SEPA and non-SEPA wards and compare the change from baseline to follow-up between SEPA and non-SEPA wards. We will use generalized estimating equations to account for clustering (at ward and community levels), differences at baseline, known potential confounders, and any other community-level changes in the period [85]. A supplementary analysis will look at changes in outcomes in relation to the SEPA implementation 'score' (see Process Evaluation of the SEPA Intervention) in individual SEPA communities.

Qualitative Evaluation of Impact

Drawing on the most significant change technique [86], we will collect stories of life changes that participants attribute to the SEPA intervention. These stories can reveal both expected and unexpected effects of the intervention. This narrative approach complements quantitative evaluations, shedding light on possible mechanisms of effect. Sampling is usually purposive and should aim to include a range of experiences and views [87,88].

In intervention communities, local researchers will identify a purposive sample of 60 storytellers, including at least 25 adolescent girls and 15 adolescent boys. The sample will include people likely to have had a range of different experiences depending on their age, gender, urban or rural residence, role in the community, and extent of involvement in SEPA. Fieldworkers will collect stories by taking notes as the storytellers speak, reading back the stories to them to check for accuracy.

A hybrid thematic analysis [53,54] of the narratives will include a deductive analysis using the CASCADA results chain as a framework, supplemented by an inductive analysis of other themes emerging in the narratives.

Auditing Trial Conduct

A project steering committee led by the Bauchi State Primary Health Care Development Agency and including relevant government bodies and the research team will meet twice-yearly for oversight and decision-making.

Ethical Considerations

Ethical Approval

On 1 March 2021, the Bauchi State Health Research Ethics Committee approved the overall project (NREC/03/11/19B/2021/03). On 13 September 2021, the Faculty of Medicine and Health Sciences Institutional Review Board at McGill University (A09-B51-21B) approved the initial qualitative phase of focus groups and FCM and will be asked to approve the baseline and follow-up survey and implementation of the intervention of evidence-based dialogues.

Consent

Local research supervisors will seek consent from community leaders to work in each community. Facilitators will seek oral informed consent from group participants and parental/guardian consent for participants under 18 years old. Facilitators will ask

participants to respect each other's confidentiality and stress that they should not share personal information in the group setting. For the household survey, interviewers will seek and record on the handset oral informed consent from all respondents. They will also seek parental consent for adolescents under 18 years old.

Confidentiality and Data Access

Parents or guardians will not be present during adolescent group activities or individual interviews. We will not record any names or identifying information alongside responses from individuals. Group reports will not identify individuals. Fieldworkers will conduct group sessions in a private location. They will not proceed with household interviews unless they can establish and maintain privacy. Data on the server are password protected; only designated research team members will have access.

Minimizing Potential Harms

Survey respondents, especially females, who disclose sensitive issues, like domestic violence, might face retribution if other household members hear of this. Ensuring privacy minimizes this risk. Discussing topics like experiencing violence or abuse might cause distress. Interviewers will carry details of local support services. Facilitators will refer any group participant who is disturbed by the discussion to pre-arranged support in the community.

Fieldworkers, particularly women, potentially face security threats. All fieldworkers will be from the local area. Each fieldwork team will include male members, part of whose role is to ensure the safety of their female colleagues. Government focal points in each ward will advise on any current security risks, and the teams will not visit insecure communities.

Dissemination and Capacity Building

The Bauchi State Primary Health Care Development Agency will convene meetings at state and zonal levels to discuss findings and policy implications with planners and decision-makers in government, NGOs, and development partners. In year three, the project will share findings from the FCM and baseline survey, and in year five, discuss the impact of the SEPA intervention.

Evidence about factors related to ASRH outcomes in Nigeria and how these might be improved will fill a gap in the literature. We will also publish articles about participatory methods and advances in their analysis, of interest beyond the field of ASRH research. We plan to publish at least five open access papers in peer-reviewed journals and present findings in two international conferences.

Capacity Building

The mainly female research team will consolidate their skills in participatory methods and learn about recent advances. The senior team members will support junior members and knowledge users to analyze and present findings and write articles for publication. Officers from the State Ministry of Health and Primary Health Care Development Agency will work on the project and learn about participatory, qualitative methods and collecting reliable survey data. In year five, the research team will facilitate an analysis and interpretation workshop over 18 days for 20 women and men from the government and FOMWAN. Participants will practice analysis techniques using project data. The research team will support them in preparing articles for publication. Adolescents will design materials to share the evidence from FCM and the baseline survey, including video docudramas. We will help them to create videos and other materials, and they will build skills as they do so.

Results

Focus groups exploring ASRH concerns of stakeholders began in October 2021. Baseline data collection in the household survey is expected to take place in mid-2022.

Discussion

The project will demonstrate the feasibility of implementing ASRH interventions in a Muslim and conservative culture. It tests a participatory approach: adolescents and other stakeholders share their knowledge, codesign instruments to collect data about ASRH and share the findings, participate in evidence-informed dialogue, and codesign and implement culturally appropriate solutions.

Methodological innovations will have wider relevance. We will test FCM with adolescents and continue our advances in the analysis of these maps. We will refine our use of transitive closure to analyze shifts in intermediate outcomes after an intervention [89].

During the project, adolescents and other stakeholders will identify interventions at different levels, from individual to policy level, to improve ASRH in Bauchi, that can be implemented by adolescents, by communities, and by services. Some will be implemented and have a measurable impact during the project; others are long-term interventions, and measurement of their impact is beyond the scope of this project. The interventions will be specific to the culture of Bauchi, but the way we develop them with stakeholders could have much wider resonance. The project could pave the way for a full-scale randomized controlled trial of the participatory intervention across different contexts.

By training and working with planners and decision-makers, the project will support the adoption of evidence-based policies to improve ASRH in Bauchi. It will contribute to a culture of evidence-based planning of health services, building on previous work in the State by the research team.

Evidence about factors related to ASRH outcomes in Nigeria and implementation and testing of a dialogic intervention to improve these outcomes will fill a gap in the literature. The project will document and test the effectiveness of a participatory approach to ASRH intervention research.

Acknowledgments

This study is supported by the Canadian Institutes of Health Research (grant numbers P14-175355 and PJT-178066). The funding body had no role in the study design, and it will have no role in the collection, management, analysis, and interpretation of data, writing of the study report, and the decision to submit the study report for publication.

We thank colleagues in Bauchi, especially in the Bauchi State Primary Health Care Development Agency and FOMWAN, for their support and constructive comments as we developed the project protocol.

Authors' Contributions

AC designed the project and drafted the manuscript. YG, KO, RM, LB, UA, and CM contributed to the project design and reviewed the manuscript. NA codesigned the project and supported the drafting of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from the CIHR/ISRC (Canada) Project Grant Committee.
[PDF File (Adobe PDF File), 78 KB - resprot_v11i3e36060_app1.pdf]

References

1. Patton GC, Coffey C, Sawyer SM, Viner RM, Haller DM, Bose K, et al. Global patterns of mortality in young people: a systematic analysis of population health data. The Lancet 2009 Sep;374(9693):881-892. [doi: 10.1016/s0140-6736(09)60741-8]
2. Melesse DY, Mutua MK, Choudhury A, Wado YD, Faye CM, Neal S, et al. Adolescent sexual and reproductive health in sub-Saharan Africa: who is left behind? BMJ Glob Health 2020 Dec 13;5(1):e002231 [FREE Full text] [doi: 10.1136/bmjgh-2019-002231] [Medline: 32133182]
3. Federal MOH, Nigeria. National policy on the health & development of adolescents & young people in Nigeria. First revision. URL: http://www.health.gov.ng/doc/policy.pdf [accessed 2021-12-28]
4. National Population Commission (NPC)ICF International. Nigeria Demographic and Health Survey. 2018. URL: https://www.dhsprogram.com/publications/publication-FR359-DHS-Final-Reports.cfm [accessed 2021-12-28]
5. Cortez R, Saadat S, Marinda E, Odutolu O. Adolescent fertility and sexual health in Nigeria: determinants and implications. Health, Nutrition and Population (HNP) Discussion Paper, World Bank. 2016. URL: https://openknowledge.worldbank.org/bitstream/handle/10986/24041/Adolescent0fer0al0health0in0Nigeria.pdf;sequence=1 [accessed 2021-12-28]
6. Adebowale SA, Fagbamigbe FA, Okareh TO, Lawal GO. Survival analysis of timing of first marriage among women of reproductive age in Nigeria: regional differences. Afr J Reprod Health 2012 Dec;16(4):95-107. [Medline: 23444547]
7. Okereke CI. Assessing the prevalence and determinants of adolescents' unintended pregnancy and induced abortion in Owerri, Nigeria. J Biosoc Sci 2010 Sep;42(5):619-632. [doi: 10.1017/S0021932010000179] [Medline: 20522271]
8. Aderibigbe SA, Araoye MO, Akande TM, Monehin JO, Musa OI, Babatunde OA. Teenage Pregnancy and Prevalence of Abortion among In-school Adolescents in North Central, Nigeria. ASS 2010 Dec 21;7(1):122-127. [doi: 10.5539/ass.v7n1p122]
9. Aderinto A. Sexual abuse of the girl-child in urban Nigeria and implications for the transmission of HIV/AIDS. Gen & Behav 2010 May 24;8(1):2735-2761 [FREE Full text] [doi: 10.4314/gab.v8i1.54690]
10. Okoro F, Obozokhai O. Sexual harassment: the experience of out-of-school teenagers in Benin City, Nigeria. Afr J Reprod Health 2006;9(3):118-127. [doi: 10.2307/3583418]
11. Olley BO. Social and health behaviors in youth of the streets of Ibadan, Nigeria. Child Abuse Negl 2006 Mar;30(3):271-282. [doi: 10.1016/j.chiabu.2005.10.011] [Medline: 16519940]
12. Aransiola JO, Asa S, Obinjuwa P, Olarewaju O, Ojo OO, Fatusi AO. Teachers' perspectives on sexual and reproductive health interventions for in-school adolescents in Nigeria. Afr J Reprod Health 2013 Dec;17(4):84-92. [Medline: 24558785]
13. Dlamini N, Okoro F, Ekhosuehi U, Esiet A, Lowik A, Metcalfe K. Empowering teachers to change youth practices: evaluating teacher delivery and responses to the FLHE programme in Edo State, Nigeria. Afr J Reprod Health 2012 Jun;16(2):87-102. [Medline: 22916546]
14. Isiugo-Abanihe UC, Olajide R, Nwokocha E, Fayehun F, Okunola R, Akingbade R. Adolescent Sexuality and Life Skills Education in Nigeria: To What Extent have Out-of-School Adolescents Been Reached? Afr J Reprod Health 2015 Mar;19(1):101-111. [Medline: 26103700]
15. Adedimeji AA, Omololu FO, Odutolu O. HIV risk perception and constraints to protective behaviour among young slum dwellers in Ibadan, Nigeria. J Health Popul Nutr 2007 Jun;25(2):146-157 [FREE Full text] [Medline: 17985816]

16. Oladepo O, Fayemi MM. Perceptions about sexual abstinence and knowledge of HIV/AIDS prevention among in-school adolescents in a western Nigerian city. BMC Public Health 2011 May 12;11:304 [FREE Full text] [doi: 10.1186/1471-2458-11-304] [Medline: 21569416]
17. Onyeonoro UU, Oshi DC, Ndimele EC, Chuku NC, Onyemuchara IL, Ezekwere SC, et al. Sources of sex information and its effects on sexual practices among in-school female adolescents in Osisioma Ngwa LGA, south east Nigeria. J Pediatr Adolesc Gynecol 2011 Oct;24(5):294-299. [doi: 10.1016/j.jpag.2011.05.002] [Medline: 21763163]
18. Owoaje ET, Uchendu OC. Sexual risk behaviour of street youths in south west Nigeria. East Afr J Public Health 2009 Dec;6(3):274-279. [Medline: 20803918]
19. Owolabi AT, Onayade AA, Ogunlola IO, Ogunniyi SO, Kuti O. Sexual behaviour of secondary school adolescents in Ilesa, Nigeria: implications for the spread of STIs including HIV/AIDS. J Obstet Gynaecol 2005 Feb;25(2):174-178. [doi: 10.1080/01443610500051460] [Medline: 15814400]
20. Oyediran KA, Feyisetan OI, Akpan T. Predictors of condom-use among young never-married males in Nigeria. J Health Popul Nutr 2011 Jun 22;29(3):273-285 [FREE Full text] [doi: 10.3329/jhpn.v29i3.7875] [Medline: 21766563]
21. Sunmola AM, Dipeolu M, Babalola S, Otu AD. Reproductive, sexual and contraceptive behaviour of adolescents in Niger State, Nigeria. Afr J Reprod Health 2002 Dec;6(3):82-92. [Medline: 12685412]
22. Loto O, Ezechi O, Kalu B, Loto AB, Ezechi LO, Ogunniyi S. Poor obstetric performance of teenagers: is it age- or quality of care-related? J Obstet Gynaecol 2004 Jun 02;24(4):395-398. [doi: 10.1080/01443610410001685529] [Medline: 15203579]
23. Rai RK, Singh PK, Singh L. Utilization of maternal health care services among married adolescent women: insights from the Nigeria Demographic and Health Survey, 2008. Womens Health Issues 2012;22(4):e407-e414. [doi: 10.1016/j.whi.2012.05.001] [Medline: 22749200]
24. Ajayi AI, Otukpa EO, Mwoka M, Kabiru CW, Ushie BA. Adolescent sexual and reproductive health research in sub-Saharan Africa: a scoping review of substantive focus, research volume, geographic distribution and Africa-led inquiry. BMJ Glob Health 2021 Feb 10;6(2):e004129 [FREE Full text] [doi: 10.1136/bmjgh-2020-004129] [Medline: 33568395]
25. Chandra-Mouli V, Lane C, Wong S. What Does Not Work in Adolescent Sexual and Reproductive Health: A Review of Evidence on Interventions Commonly Accepted as Best Practices. Glob Health Sci Pract 2015 Aug 31;3(3):333-340. [doi: 10.9745/ghsp-d-15-00126]
26. Salam RA, Faqqah A, Sajjad N, Lassi ZS, Das JK, Kaufman M, et al. Improving Adolescent Sexual and Reproductive Health: A Systematic Review of Potential Interventions. J Adolesc Health 2016 Oct;59(4S):S11-S28 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.05.022] [Medline: 27664592]
27. Innovating for Maternal and Child Health in Africa. International Development Research Centre. URL: https://www.idrc.ca/en/initiative/innovating-maternal-and-child-health-africa [accessed 2021-12-28]
28. Cockcroft A, Omer K, Gidado Y, Gamawa AI, Andersson N. Impact of universal home visits on maternal and infant outcomes in Bauchi state, Nigeria: protocol of a cluster randomized controlled trial. BMC Health Serv Res 2018 Jul 03;18(1):510 [FREE Full text] [doi: 10.1186/s12913-018-3319-z] [Medline: 29970071]
29. Sarmiento I, Ansari U, Omer K, Gidado Y, Baba MC, Gamawa AI, et al. Causes of short birth interval (kunika) in Bauchi State, Nigeria: systematizing local knowledge with fuzzy cognitive mapping. Reprod Health 2021 Apr 06;18(1):74 [FREE Full text] [doi: 10.1186/s12978-021-01066-2] [Medline: 33823874]
30. George AS, Amin A, de Abreu Lopes CM, Ravindran TKS. Structural determinants of gender inequality: why they matter for adolescent girls' sexual and reproductive health. BMJ 2020 Jan 27:l6985. [doi: 10.1136/bmj.l6985]
31. Andersson N, Ledogar RJ. The CIET Aboriginal Youth Resilience Studies: 14 Years of Capacity Building and Methods Development in Canada. Pimatisiwin 2008;6(2):65-88 [FREE Full text] [Medline: 20862230]
32. Andersson N. Building the community voice into planning: 25 years of methods development in social audit. BMC Health Serv Res 2011 Dec 21;11 Suppl 2:S1-S3 [FREE Full text] [doi: 10.1186/1472-6963-11-S2-S1] [Medline: 22376121]
33. Bentler PM, Speckart G. Models of attitude-behavior relations. Psychological Review 1979;86(5):452-464. [doi: 10.1037/0033-295x.86.5.452]
34. Baranowski T, Cullen KW, Nicklas T, Thompson D, Baranowski J. Are current health behavioral change models helpful in guiding prevention of weight gain efforts? Obes Res 2003 Oct;11 Suppl:23S-43S [FREE Full text] [doi: 10.1038/oby.2003.222] [Medline: 14569036]
35. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
36. Cockcroft A, Andersson N, Omer K, Ansari NM, Khan A, Chaudhry UU, et al. One size does not fit all: local determinants of measles vaccination in four districts of Pakistan. BMC Int Health Hum Rights 2009 Oct 14;9 Suppl 1:S4 [FREE Full text] [doi: 10.1186/1472-698X-9-S1-S4] [Medline: 19828062]
37. Mitchell S, Cockcroft A, Lamothe G, Andersson N. Equity in HIV testing: evidence from a cross-sectional study in ten Southern African countries. BMC Int Health Hum Rights 2010 Sep 13;10:23 [FREE Full text] [doi: 10.1186/1472-698X-10-23] [Medline: 20836859]
38. Andersson N, Cockcroft A, Ansari N, Omer K, Baloch M, Ho Foster A, et al. Evidence-based discussion increases childhood vaccination uptake: a randomised cluster controlled trial of knowledge translation in Pakistan. BMC Int Health Hum Rights 2009 Oct 14;9 Suppl 1:S8 [FREE Full text] [doi: 10.1186/1472-698X-9-S1-S8] [Medline: 19828066]

39. Cameron M, Cockcroft A, Waichigo GW, Marokoane N, Laetsang D, Andersson N. From knowledge to action: participant stories of a population health intervention to reduce gender violence and HIV in three southern African countries. AIDS Care 2014 Jul 03;26(12):1534-1540 [FREE Full text] [doi: 10.1080/09540121.2014.931560] [Medline: 24991886]
40. Belaid L, Ansari U, Omer K, Gidado Y, Baba MC, Daniel LE, et al. "I had to change my attitude": narratives of most significant change explore the experience of universal home visits to pregnant women and their spouses in Bauchi State, Nigeria. Arch Public Health 2021 Nov 18;79(1):202 [FREE Full text] [doi: 10.1186/s13690-021-00735-9] [Medline: 34794488]
41. Odok GE. Disruption of patriarchy in northern Islamic Nigeria. Gender, Place and Culture 2019 Dec 17;27(12):1663-1681. [doi: 10.1080/0966369x.2019.1693346]
42. Maina BW, Orindi BO, Sikweyiya Y, Kabiru CW. Gender norms about romantic relationships and sexual experiences among very young male adolescents in Korogocho slum in Kenya. Int J Public Health 2020 May 09;65(4):497-506 [FREE Full text] [doi: 10.1007/s00038-020-01364-9] [Medline: 32270236]
43. Moreau C, Li M, De Meyer S, Vu Manh L, Guiella G, Acharya R, et al. Measuring gender norms about relationships in early adolescence: Results from the global early adolescent study. SSM Popul Health 2019 Apr;7:014-014 [FREE Full text] [doi: 10.1016/j.ssmph.2018.10.014] [Medline: 30581959]
44. Andersson N. Prevention for those who have freedom of choice--or among the choice-disabled: confronting equity in the AIDS epidemic. AIDS Res Ther 2006 Sep 25;3:23 [FREE Full text] [doi: 10.1186/1742-6405-3-23] [Medline: 16999869]
45. Andersson N, Cockcroft A. Choice-disability and HIV infection: a cross sectional study of HIV status in Botswana, Namibia and Swaziland. AIDS Behav 2012 Jan;16(1):189-198 [FREE Full text] [doi: 10.1007/s10461-011-9912-3] [Medline: 21390539]
46. Andersson N, Cockcroft A, Thabane L, Marokoane N, Laetsang D, Masisi M. HIV prevention in favour of the choice-disabled in southern Africa: study protocol for a randomised controlled trial. Trials 2013 Aug 29;14(1):274 [FREE Full text] [doi: 10.1186/1745-6215-14-274] [Medline: 23987126]
47. Cockcroft A, Marokoane N, Kgakole L, Kefas J, Andersson N. The Inter-ministerial National Structural Intervention trial (INSTRUCT): protocol for a parallel group cluster randomised controlled trial of a structural intervention to reduce HIV infection among young women in Botswana. BMC Health Serv Res 2018 Oct 30;18(1):822 [FREE Full text] [doi: 10.1186/s12913-018-3638-0] [Medline: 30376834]
48. Williams R. Cultural safety--what does it mean for our work practice? Aust N Z J Public Health 1999 Apr;23(2):213-214. [doi: 10.1111/j.1467-842x.1999.tb01240.x] [Medline: 10330743]
49. Curtis E, Jones R, Tipene-Leach D, Walker C, Loring B, Paine S, et al. Why cultural safety rather than cultural competency is required to achieve health equity: a literature review and recommended definition. Int J Equity Health 2019 Nov 14;18(1):174 [FREE Full text] [doi: 10.1186/s12939-019-1082-3] [Medline: 31727076]
50. Andersson N. Participatory research-A modernizing science for primary health care. J Gen Fam Med 2018 Sep 11;19(5):154-159 [FREE Full text] [doi: 10.1002/jgf2.187] [Medline: 30186727]
51. Ansari U, Pimentel J, Omer K, Gidado Y, Baba M, Andersson N, et al. "Kunika women are always sick": views from community focus groups on short birth interval (kunika) in Bauchi state, northern Nigeria. BMC Womens Health 2020 May 24;20(1):113 [FREE Full text] [doi: 10.1186/s12905-020-00970-2] [Medline: 32448373]
52. Ansari U, Cobham B, Etim EM, Ahamad HM, Owan NO, Tijani Y, et al. Insights Into Intimate Partner Violence in Pregnancy: Findings From a Cross-Sectional Study in Two States in Nigeria. Violence Against Women 2017 Mar 09;23(4):469-481. [doi: 10.1177/1077801216644072] [Medline: 27121216]
53. Thomas DR. A General Inductive Approach for Analyzing Qualitative Evaluation Data. American Journal of Evaluation 2006 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
54. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
55. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. ECTJ 1981 Jun;29(2):75. [doi: 10.1007/bf02766777]
56. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. International Journal of Qualitative Methods 2017 Oct 02;16(1):1-13. [doi: 10.1177/1609406917733847]
57. FCMappers. URL: http://www.fcmappers.net/joomla/ [accessed 2021-12-28]
58. Özesmi U, Özesmi SL. Ecological models based on people's knowledge: a multi-step fuzzy cognitive mapping approach. Ecological Modelling 2004 Aug;176(1-2):43-64. [doi: 10.1016/j.ecolmodel.2003.10.027]
59. Giles BG, Findlay CS, Haas G, LaFrance B, Laughing W, Pembleton S. Integrating conventional science and aboriginal perspectives on diabetes using fuzzy cognitive maps. Soc Sci Med 2007 Feb;64(3):562-576. [doi: 10.1016/j.socscimed.2006.09.007] [Medline: 17084952]
60. Giles BG, Haas G, Šajna M, Findlay CS. Exploring Aboriginal Views of Health Using Fuzzy Cognitive Maps and Transitive Closure. Can J Public Health 2008 Sep 1;99(5):411-417. [doi: 10.1007/bf03405252]
61. yEd. URL: https://www.yworks.com/products/yed [accessed 2021-12-28]

62. Sarmiento I, Paredes-Solís S, Dion A, Silver H, Vargas E, Cruz P, et al. Maternal health and Indigenous traditional midwives in southern Mexico: contextualisation of a scoping review. BMJ Open 2021 Dec 23;11(12):e054542 [FREE Full text] [doi: 10.1136/bmjopen-2021-054542] [Medline: 34949629]
63. Desrosiers A, Betancourt T, Kergoat Y, Servilli C, Say L, Kobeissi L. A systematic review of sexual and reproductive health interventions for young people in humanitarian and lower-and-middle-income country settings. BMC Public Health 2020 May 12;20(1):666 [FREE Full text] [doi: 10.1186/s12889-020-08818-y] [Medline: 32398129]
64. Dion A, Joseph L, Jimenez V, Gutierrez AC, Ben Ameur A, Robert E, et al. Grounding evidence in experience to support people-centered health services. Int J Public Health 2019 Jun;64(5):797-802 [FREE Full text] [doi: 10.1007/s00038-018-1180-9] [Medline: 30542753]
65. Dion A, Carini-Gutierrez A, Jimenez V, Ben Ameur A, Robert E, Joseph L, et al. Weight of Evidence: Participatory Methods and Bayesian Updating to Contextualize Evidence Synthesis in Stakeholders' Knowledge. Journal of Mixed Methods Research 2021 Oct 13:1-26. [doi: 10.1177/15586898211037412]
66. Omer K, Mhatre S, Ansari N, Laucirica J, Andersson N. Evidence-based training of frontline health workers for door-to-door health promotion: a pilot randomized controlled cluster trial with Lady Health Workers in Sindh Province, Pakistan. Patient Educ Couns 2008 Aug;72(2):178-185. [doi: 10.1016/j.pec.2008.02.018] [Medline: 18395396]
67. Arrizón AV, Andersson N, Ledogar RJ. Micro-regional planning: evidence-based community buy-in for health development in five of Mexico's poorest rural districts. BMC Health Serv Res 2011 Dec 21;11 Suppl 2:S2 [FREE Full text] [doi: 10.1186/1472-6963-11-S2-S2] [Medline: 22375532]
68. Andersson N, Nava-Aguilera E, Arosteguí J, Morales-Perez A, Suazo-Laguna H, Legorreta-Soberanis J, et al. Evidence based community mobilization for dengue prevention in Nicaragua and Mexico (Camino Verde, the Green Way): cluster randomized controlled trial. BMJ 2015 Jul 08;351:h3267 [FREE Full text] [doi: 10.1136/bmj.h3267] [Medline: 26156323]
69. Ledogar RJ, Arosteguí J, Hernández-Alvarez C, Morales-Perez A, Nava-Aguilera E, Legorreta-Soberanis J, et al. Mobilising communities for Aedes aegypti control: the SEPA approach. BMC Public Health 2017 May 30;17(Suppl 1):403 [FREE Full text] [doi: 10.1186/s12889-017-4298-4] [Medline: 28699561]
70. Boyko JA, Lavis JN, Abelson J, Dobbins M, Carter N. Deliberative dialogues as a mechanism for knowledge translation and exchange in health systems decision-making. Soc Sci Med 2012 Dec;75(11):1938-1945. [doi: 10.1016/j.socscimed.2012.06.016] [Medline: 22938912]
71. Andersson N. Community-led trials: Intervention co-design in a cluster randomised controlled trial. BMC Public Health 2017 May 30;17(Suppl 1):397 [FREE Full text] [doi: 10.1186/s12889-017-4288-6] [Medline: 28699556]
72. Omer K, Ansari U, Aziz A, Hassan K, Bgeidam LA, Baba MC, et al. Participatory health research under COVID-19 restrictions in Bauchi State, Nigeria: Feasibility of cellular teleconferencing for virtual discussions with community groups in a low-resource setting. Digit Health 2022;8:20552076211070386 [FREE Full text] [doi: 10.1177/20552076211070386] [Medline: 35003757]
73. Thompson J, Mitchell C, Starr L. Cellphilming: a tool for addressing gender equality, facilitators guide. Ottawa, ON: CODE. 2019. URL: https://www.mcgill.ca/morethanwords/files/morethanwords/cellphilm_guidebook.final_.pdf [accessed 2021-12-28]
74. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]
75. Taylor DW, Bosch EG. CTS: a clinical trials simulator. Stat Med 1990 Jul;9(7):787-801. [doi: 10.1002/sim.4780090708] [Medline: 2218181]
76. Omer K, Afi NJ, Baba MC, Adamu M, Malami SA, Oyo-Ita A, et al. Seeking evidence to support efforts to increase use of antenatal care: a cross-sectional study in two states of Nigeria. BMC Pregnancy Childbirth 2014 Nov 20;14:380 [FREE Full text] [doi: 10.1186/s12884-014-0380-4] [Medline: 25410003]
77. Cockcroft A, Omer K, Gidado Y, Baba MC, Aziz A, Ansari U, et al. The impact of universal home visits with pregnant women and their spouses on maternal outcomes: a cluster randomised controlled trial in Bauchi State, Nigeria. BMJ Glob Health 2019 Jul 03;4(1):e001172 [FREE Full text] [doi: 10.1136/bmjgh-2018-001172] [Medline: 30899560]
78. Open Data Kit. URL: https://opendatakit.org/ [accessed 2021-12-28]
79. Andersson N, Mitchell S. Epidemiological geomatics in evaluation of mine risk education in Afghanistan: introducing population weighted raster maps. Int J Health Geogr 2006 Jan 03;5:1 [FREE Full text] [doi: 10.1186/1476-072X-5-1] [Medline: 16390549]
80. Global Early Adolescent Study Questionnaires. URL: https://www.geastudy.org/download-measures-redesign [accessed 2021-12-28]
81. Adolescent sexual and reproductive health stigma scale. URL: https://emerge.ucsd.edu/r_2yc4kivbzdy04ws/ [accessed 2021-12-28]
82. Hall KS, Manu A, Morhe E, Harris LH, Loll D, Ela E, et al. Development and Validation of a Scale to Measure Adolescent Sexual and Reproductive Health Stigma: Results From Young Women in Ghana. J Sex Res 2018 Jan 07;55(1):60-72 [FREE Full text] [doi: 10.1080/00224499.2017.1292493] [Medline: 28266874]

83. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959 Apr;22(4):719-748. [Medline: 13655060]
84. Andersson N, Lamothe G. Clustering and meso-level variables in cross-sectional surveys: an example of food aid during the Bosnian crisis. BMC Health Serv Res 2011 Dec 21;11 Suppl 2:S15 [FREE Full text] [doi: 10.1186/1472-6963-11-S2-S15] [Medline: 22376353]
85. Hox J, Moerbeek M, van de Schoot R. Multilevel analysis: techniques and applications, Third Edition (Quantitative Methodology Series). In: Multilevel analysis: techniques and applications,. New York: Routledge; 2018. [doi: 10.4324/9781315650982]
86. Dart J, Davies R. A Dialogical, Story-Based Evaluation Tool: The Most Significant Change Technique. American Journal of Evaluation 2016 Jun 30;24(2):137-155. [doi: 10.1177/109821400302400202]
87. Willetts J, Crawford P. The most significant lessons about the Most Significant Change technique. Development in Practice 2007 Jun;17(3):367-379. [doi: 10.1080/09614520701336907]
88. Tonkin K, Silver H, Pimentel J, Chomat AM, Sarmiento I, Belaid L, et al. How beneficiaries see complex health interventions: a practice review of the Most Significant Change in ten countries. Arch Public Health 2021 Feb 08;79(1):18 [FREE Full text] [doi: 10.1186/s13690-021-00536-0] [Medline: 33557938]
89. Andersson N, Beauchamp M, Nava-Aguilera E, Paredes-Solís S, Šajna M. The women made it work: fuzzy transitive closure of the results chain in a dengue prevention trial in Mexico. BMC Public Health 2017 May 30;17(Suppl 1):408 [FREE Full text] [doi: 10.1186/s12889-017-4301-0] [Medline: 28699563]

Abbreviations

ASRH: adolescent sexual and reproductive health
CASCADA: conscious knowledge, attitude, subjective norms, intention to change, agency to change, discussion of issues, and action to change
FCM: fuzzy cognitive mapping
FOMWAN: Federation of Muslim Women's Associations in Nigeria
NGO: nongovernment organization
SEPA: socializing evidence for participatory action

Edited by T Derrick; submitted 30.12.21; this is a non–peer-reviewed article;accepted 05.01.22; published 15.03.22.

Please cite as:
Cockcroft A, Omer K, Gidado Y, Mohammed R, Belaid L, Ansari U, Mitchell C, Andersson N
Impact-Oriented Dialogue for Culturally Safe Adolescent Sexual and Reproductive Health in Bauchi State, Nigeria: Protocol for a Codesigned Pragmatic Cluster Randomized Controlled Trial
JMIR Res Protoc 2022;11(3):e36060
URL: https://www.researchprotocols.org/2022/3/e36060
doi:10.2196/36060
PMID:35289762

Original Paper

Feasibility and Acceptability of an Online WhatsApp Support Group on Breastfeeding: Protocol for a Randomized Controlled Trial

Kris YW Lok[1], PhD; Rachel WT Ko[1], MPhil; Heidi SL Fan[1], PhD; PH Chau[1], PhD; Janet YH Wong[1], PhD; MP Wang[1], PhD; Vicky Tsang[2]

[1]School of Nursing, Li Ka Shing Faculty of Medicine, University of Hong Kong, Pokfulam, Hong Kong

[2]Natural Parenting Network, Hong Kong, Hong Kong

Corresponding Author:
Kris YW Lok, PhD
School of Nursing, Li Ka Shing Faculty of Medicine
University of Hong Kong
4/F, William MW Mong Block, 21 Sassoon Road
Pokfulam
Hong Kong
Phone: 852 39176690
Email: krislok@hku.hk

Abstract

Background: Mobile health, the use of mobile technology in delivering health care, has been found to be effective in changing health behaviors, including improving breastfeeding practices in postpartum women. With the widespread use of smartphones and instant messaging apps in Hong Kong, instant messaging groups could be a useful channel for delivering breastfeeding peer support.

Objective: The aim of this paper is to study the feasibility and acceptability of an online instant messaging peer support group by trained peer counselors on improving breastfeeding outcome in primiparous women in Hong Kong.

Methods: A two-arm, assessor-blind, randomized controlled feasibility study will be conducted on 40 primiparous women with the intention to breastfeed. Participants are recruited from the antenatal obstetrics and gynecology clinic of a public hospital in Hong Kong and randomly assigned at a 1:1 ratio to either intervention or control group. The intervention group receives peer support in an online instant messaging group with trained peer counselors on top of standard care, whereas the control group receives standard care. Breastfeeding outcome will be assessed for 6 months post partum or until weaned. The breastfeeding status, the proportion and duration of exclusive and any breastfeeding in each group, and the self-efficacy and attitude of participants will be assessed. The feasibility and acceptability of the study would also be assessed in preparation for a full randomized controlled trial.

Results: This study (protocol version 1 dated January 5, 2021) has been reviewed and approved by the institutional review board of the University of Hong Kong, Hospital Authority Hong Kong West Cluster (reference UW 21-039), on January 26, 2021. Data collection is ongoing and expected to be completed in December 2021. The findings will be updated on clinical trial registry and disseminated in peer-reviewed journals.

Conclusions: This study aims to assess the feasibility and effectiveness of an online instant messaging peer support group in improving the breastfeeding outcome of primiparous women in Hong Kong. Its findings could inform the feasibility of a full-scale trial with this intervention design.

Trial Registration: ClinicalTrials.gov NCT04826796; https://clinicaltrials.gov/ct2/show/NCT04826796

International Registered Report Identifier (IRRID): DERR1-10.2196/32338

(JMIR Res Protoc 2022;11(3):e32338) doi:10.2196/32338

KEYWORDS

mHealth; breastfeeding; peer support; mobile health; parenting; instant messaging; online support; women's health; postpartum health; postpartum support

Introduction

Mobile health (mHealth) is the delivery of medical and public health practices using mobile devices [1]. With the advances and increasing ubiquity in mobile technologies, there is an increase in the use of mHealth interventions for promoting healthy infant feeding practices and breastfeeding. A recent meta-analysis has reported that mHealth significantly improved exclusive breastfeeding initiation, breastfeeding attitude, and knowledge [2]. However, the effectiveness of the more interactive and adaptive platforms, such as smartphone apps and social networking tools, in improving breastfeeding outcomes remains understudied.

Smartphones and social networking apps have emerged in recent years as the new tool for information acquisition and exchanges. In 2019, the number of smartphone users in Hong Kong was estimated to reach 6.5 million. It is predicted that 93% of the population in Hong Kong will be smartphone users by 2025 [3]. Previous studies have found the use of these online discussion and social networking platforms to be an acceptable, affordable, and accessible method for a range of health promotion intervention, including promoting physical activities [4], healthy diet [5], smoking cessation [6], and reducing alcohol use [7]. As these communication tools allow for more personalized and instant response without time and location limitations, interventions delivered through online social networking platforms tend to have a high uptake. These platforms could be a promising tool for breastfeeding promotion. A recent integrated review has found that women tend to seek online breastfeeding support if they were isolated or do not receive sufficient professional support, and that online support could be an accessible and easily available channel for them to learn from others' experience [8]. Nevertheless, most studies on the topic to date have focused on the qualitative aspects of the online support experience [8], and there remains a lack of intervention studies on online peer support intervention. Moreover, existing data on whether online breastfeeding peer support could help improve breastfeeding outcome remain scarce and inconclusive [8].

According to the most recent survey, more than 87.5% of women initiated breastfeeding in Hong Kong in 2018 [9]. However, half of the breastfeeding women have never exclusively breastfed their babies [9]. The exclusive breastfeeding rate tailed off during the first 2 months and dropped to 26.3% at 6 months [9]. Considering the importance of exclusive breastfeeding in the first 6 months, this low exclusive breastfeeding rate is of public health concern. While practices vary, many women in Hong Kong continued to adopt postnatal rituals that Chinese women adhere to in the first month post partum, such as staying at home, eating prescribed foods (eg, chicken and Chinese ginger), and avoiding bathing or washing their hair. This would consequently mean that Chinese women would stay homebound for 1 month after giving birth and often follow specific restrictions on their diets and activities. The existing antenatal and postnatal breastfeeding support is mainly provided by hospitals and maternal and child health centers, which due to the pandemic, have been further limited. Postpartum women in Hong Kong, thus, may face additional challenges in seeking timely breastfeeding support. Unresolved problems and inadequate support are risk factors for women to cease exclusive breastfeeding early [10,11]. There is a need for innovative approaches to engage and support breastfeeding women in the first 6 months after childbirth.

Thus, a randomized controlled feasibility study is proposed with the aim to examine the practicality and feasibility of supporting Chinese breastfeeding mothers in Hong Kong. In a previous study conducted in Hong Kong, it is suggested that peer support delivered via instant text messages could be a viable way of promoting breastfeeding [12]. The instant messaging smartphone app, WhatsApp (WhatsApp Inc), is one of the most used social networking platforms in Hong Kong [13]. It is a free all-in-one app for sending text and voice messages, multimedia contents, and video calls. The proposed intervention aims to address an important service gap in Hong Kong to promote and sustain exclusive breastfeeding. While WhatsApp has been used as a tool to deliver professional breastfeeding support in countries such as Turkey [14], this will be the first WhatsApp breastfeeding support intervention provided by trained peer supporters in Hong Kong. However, there might be potential barriers to a successful implementation of the intervention and evaluation of its effectiveness. Therefore, a feasibility pilot study is proposed to identify if a WhatsApp online group on breastfeeding by peer counselors, delivered through instant messaging support and designed to improve breastfeeding outcomes, is feasible and acceptable.

This trial was registered on ClinicalTrials.gov (NCT04826796) on April 1, 2021.

Methods

Study Aim

The main objective of the study is to determine the feasibility and acceptability of an online messaging support group hosted by trained peer counselors on breastfeeding outcome in women with intention to breastfeed in Hong Kong over a 6-month period.

Trial Design

This trial adopts a randomized controlled superiority design with 2 parallel groups and a primary end point of cessation of breastfeeding in the first 6 months post partum. Participants will be randomly allocated to the intervention or control group on a 1:1 ratio.

Study Setting

Participant recruitment will be conducted at the obstetrics and gynecology outpatient clinic at a public hospital in Hong Kong serving an urban population.

Sample Size

The sample size of the trial is 40 participants, with 20 participants in each arm. As this is a feasibility and pilot study that aims to gauge the rate of recruitment, adherence, and retention levels, as well as to identify unanticipated issues of the trial design, this recommended sample size for feasibility

and pilot study is used [15]. Recruitment will continue until target sample size is reached.

Inclusion and Exclusion Criteria

To be eligible for inclusion in the study, participants must (1) be primiparous, (2) intend to breastfeed, (3) have a singleton pregnancy, (4) have term infant (37-42 weeks gestational), (5) be Cantonese speakers, (6) reside in Hong Kong, and (7) have no serious medical or obstetrical complications.

Individuals will be excluded from the study if they fail to meet the inclusion criteria or if their newborn (1) are <37 week gestation; (2) have an Apgar score <8 at 5 minutes; (3) have a birthweight <2500 grams; (4) have any severe medical conditions or congenital malformations; (5) are placed in the special care baby unit for more than 48 hours after birth; or (6) are placed in the neonatal intensive care unit at any time after birth. Participants who are, due to mental or physical reasons, unable to provide written informed consent are also excluded from the study.

Recruitment, Randomization, and Allocation of Intervention

Recruitment will be conducted at the obstetrics and gynecology outpatient antenatal clinics in a public hospital in Hong Kong. Research assistants will have obtained written informed consent from participants. All eligible participants who consented to participate will be randomized. They will be randomly assigned to either the intervention or control group at a 1:1 ratio. The randomization sequence was computer generated using STATA (Stata Corp) with simple randomization by a researcher prior to study recruitment. The sequence will be concealed in a password-encoded excel file, which will not be assessed by nor disclosed to a second researcher responsible for participant recruitment and data collection. After the participants are recruited and the baseline assessments have been completed, a third researcher with access to the randomized sequence will notify the participants of their group assignment.

Due to the nature of the study, a single-blind design is used. The participants will be notified of their group assignment after the completion of baseline assessment. The researcher responsible for study recruitment and assessments will be blinded to group allocation.

Intervention

The participant timeline in the study is outlined in Table 1. The control group will continue to receive standard care, whereas the intervention group will, in addition to standard care, be included in a breastfeeding support group on a popular online messaging mobile app, WhatsApp ("WhatsApp group"), with trained peer counselors. The participants will be notified before they were added to the group and will receive a standard welcome message in the group afterward. The message will introduce the peer counselors and encourage participants to raise any question or concern they may have. Peer counselors will send prompts asking for questions or send information related to breastfeeding every week for 6 months. No harm to the participants is anticipated. The participants could leave the WhatsApp group on their own if they wish to discontinue the intervention. The participants are permitted to receive other professional breastfeeding support.

Table 1. Participant timeline in the WhatsApp breastfeeding peer support study.

Point of contact	Study Period					
	Enrollment	Allocation	Postallocation			
	Study entry	Allocation	1 month post partum	2 months post partum	3 months post partum	6 months post partum
Enrollment						
Eligibility screen	✓					
Informed consent	✓					
Allocation		✓				
Interventions						
Control group			Standard care			
Intervention group			WhatsApp peer support group and standard care			
Assessments						
Demographics	✓					
Infant feeding status			✓	✓	✓	✓
Breastfeeding self-efficacy	✓			✓		
Breastfeeding attitude	✓			✓		

In this trial, the WhatsApp group includes 4 trained peer counselors, 1 of whom is an experienced trainer, and all participants are allocated to the intervention group. The trained peer counselors are all women living in Hong Kong with at least 4 months of breastfeeding experience. All of them have been trained under the Breastfeeding Peer Support Scheme organized by Natural Parenting Network and have received at least 16 hours of training, attended 2 practicums, and passed a

standardized assessment. The training program covered topics including (1) why breastfeeding is important, (2) how to assure good start of breastfeeding, (3) how to help mother breastfeeding, (4) communication skills, (5) common breastfeeding problems, (6) diet and hygiene, (7) maternal illness and needs, and (8) local support and role of peer counselors. The peer counselors were assessed by their trainers after completing the training program.

Outcomes

The feasibility of the study will be examined based on the proportion of participants agreed to participate in the study and the acceptance of women to be randomized; completion of the intervention and follow-up at 6 months post partum will also be measured. Data on the number of women approached and screened and the reasons for rejection or exclusion will be collected. In addition, the women's views on the intervention, including strengths, weaknesses, and room for improvements, will be assessed at each follow-up time points. The reason for dropping out of intervention would be recorded. Women in the intervention group will also be invited to a qualitative interview assessing their views on the intervention at the end of the study.

In terms of intervention efficacy, the primary outcome of the study is the infant feeding status. We will collect data on the participants' infant feeding status at 1, 2, 4, and 6 months post partum. The infant feeding status will be classified according to the World Health Organization definitions in the 24 hours prior to each data collection period [16]. We will compare the proportion of any and exclusive breastfeeding among participants in the intervention and control groups. The secondary outcomes are the breastfeeding self-efficacy and attitude of the participants. Breastfeeding self-efficacy will be measured using the Hong Kong Chinese version of the Breastfeeding Self-Efficacy Scale-Short Form (BSES-SF) [17]. BSES-SF is a 14-item scale with total score ranging from 14 to 70, with a higher score indicating higher breastfeeding self-efficacy [18]. Attitude toward breastfeeding will be measured using the Chinese version of the 17-item Iowa Infant Feeding Attitude Scale (IIFAS) [19]. The total score ranges from 17 to 85, with a higher score indicating a more favorable attitude toward breastfeeding [20]. The breastfeeding self-efficacy and attitude of participants will be measured at study entry and 2 months post partum. All participants will be followed for 6 months post partum or until weaned.

Data Collection Method

Demographic information and breastfeeding self-efficacy and attitude will be collected from the participants at study entry. The participants will be asked to complete the questionnaires after providing informed consent to participate. Demographic information such as age, educational level, household income, intention to exclusively breastfeed, and intended duration of any and exclusive breastfeeding, whether they were breastfed as a child, and whether they know someone with breastfeeding experience, were collected. Maternal and birth data will be collected by the research assistant from the participants' medical records. Follow-up assessments will be conducted via telephone interviews at 1, 2, 4, and 6 months post partum by the research assistant. To promote retention, the participants will be asked their preferred time of contact for the phone interviews. Infant feeding status will be assessed at all 4 follow-up time points, whereas the breastfeeding self-efficacy and attitude will be assessed at 2 months post partum. Any professional breastfeeding support received will also be recorded.

Data Management Plan

Data containing personal information will be stored separately in a locked cabinet at the School of Nursing or on password-encoded files. Data entry will be completed by the research assistants. Range check for data value will be conducted to promote data quality. Only study investigators and research assistants will have access to the data set.

Data Analysis

To assess the feasibility of the study, descriptive statistics of the proportion of individuals who agreed to participate and to be randomized and who completed the study, as well as any loss to follow-up will be reported. The main reasons for rejection and exclusion will also be reported. To compare the baseline characteristics of the participants across the 2 study groups, *t* tests will be used for continuous variables and chi-square tests for categorical variables. Moreover, we will compare and report the breastfeeding outcomes in the 2 study groups and will check for evidence of harm. Intention-to-treat analysis will also be conducted. Multiple imputation will be used to account for missing data, and reporting will follow the guideline published in the British Medical Journal [21]. Where appropriate, each estimate will be accompanied by a 95% confidence interval, and a 5% level of significance will be used in all statistical tests. Data analysis will be performed using the Stata (version 16.0) statistical software [22].

Data Monitoring

As the intervention poses minimal risks to the participants, no data monitoring committee is needed. However, the dialogue and responses from the peer supporter and participants are recorded and monitored in the intervention to ensure that queries are being responded to appropriately, and to ensure that any negativity that may be generated by difficulties experienced by new mothers is handled skillfully. Investigators will have the final decision to terminate the trial if needed. Any serious adverse events will be reported to the ethics committee within 1 week, and trial record will be made available for audit by the ethics committee.

Ethics Approval

This study has been reviewed and approved by the institutional review board of the University of Hong Kong, Hospital Authority Hong Kong West Cluster (Reference UW 21-039) on January 26, 2021. Any amendment to the protocol will be submitted to the institutional review board for review and approval. This study complies with the Declaration of Helsinki and its later amendments. All participants have provided written informed consent to participate.

Results

Study recruitment has commenced in March 2021. Data collection is ongoing. The projected end date of intervention

and data collection is at the end of December 2021. The result from this study will be updated on the clinical trial registry and submitted to suitable peer-reviewed publications within 12 months of study completion. Anonymized data will be available upon reasonable request within 24 months of study completion.

Discussion

This protocol for a randomized controlled trial aims to study the feasibility and potential efficacy of a WhatsApp group with trained peer counselors in improving breastfeeding outcomes in primiparous women in Hong Kong. The small sample size may limit the power and generalizability of the study; however, as a feasibility study, this could provide valuable information for conducting a full-scale randomized controlled trial.

Acknowledgments

This study is funded by Seed Fund for Basic Research, University of Hong Kong (Reference 202011159239).

Authors' Contributions

KL and RK drafted the manuscript. KL is the principal investigator of the study and is responsible for conducting the study overall. KL and VT conceived the study. KL, RK, HF, JW, PHC, MPW, and VT contributed to the design of the study. All authors contributed to the data acquisition, critically appraised and approved the manuscript, and assume responsibility for the contents of the manuscript.

Conflicts of Interest

None declared.

References

1. Premji S. Mobile health in maternal and newborn care: fuzzy logic. Int J Environ Res Public Health 2014 Jun 20;11(6):6494-6503 [FREE Full text] [doi: 10.3390/ijerph110606494] [Medline: 25003177]
2. Lau Y, Htun TP, Tam WS, Klainin-Yobas P. Efficacy of e-technologies in improving breastfeeding outcomes among perinatal women: a meta-analysis. Matern Child Nutr 2016 Jul 21;12(3):381-401 [FREE Full text] [doi: 10.1111/mcn.12202] [Medline: 26194599]
3. Number of smartphone users in Hong Kong from 2015 to 2020 with a forecast until 2026. Statista. URL: https://www.statista.com/statistics/494594/smartphone-users-in-hong-kong/ [accessed 2022-02-11]
4. Rovniak LS, Kong L, Hovell MF, Ding D, Sallis JF, Ray CA, et al. Engineering Online and In-Person Social Networks for Physical Activity: A Randomized Trial. Ann Behav Med 2016 Dec 12;50(6):885-897 [FREE Full text] [doi: 10.1007/s12160-016-9814-8] [Medline: 27405724]
5. Ahmad N, Shariff ZM, Mukhtar F, Lye M. Family-based intervention using face-to-face sessions and social media to improve Malay primary school children's adiposity: a randomized controlled field trial of the Malaysian REDUCE programme. Nutr J 2018 Aug 02;17(1):74 [FREE Full text] [doi: 10.1186/s12937-018-0379-1] [Medline: 30071855]
6. Cheung YTD, Chan CHH, Ho KS, Fok WP, Conway M, Wong CKH, et al. Effectiveness of WhatsApp online group discussion for smoking relapse prevention: protocol for a pragmatic randomized controlled trial. Addiction 2020 Sep 20;115(9):1777-1785 [FREE Full text] [doi: 10.1111/add.15027] [Medline: 32107817]
7. Byrnes HF, Miller BA, Grube JW, Bourdeau B, Buller DB, Wang-Schweig M, et al. Prevention of alcohol use in older teens: A randomized trial of an online family prevention program. Psychol Addict Behav 2019 Feb;33(1):1-14 [FREE Full text] [doi: 10.1037/adb0000442] [Medline: 30640504]
8. Moon H, Woo K. An integrative review on mothers' experiences of online breastfeeding peer support: Motivations, attributes and effects. Matern Child Nutr 2021 Jul 07;17(3):e13200 [FREE Full text] [doi: 10.1111/mcn.13200] [Medline: 33960665]
9. Breastfeeding Survey 2019. Department of Health. URL: https://www.fhs.gov.hk/english/archive/files/reports/BF_survey_2019.pdf [accessed 2022-02-11]
10. Sriraman NK, Kellams A. Breastfeeding: What are the Barriers? Why Women Struggle to Achieve Their Goals. J Womens Health (Larchmt) 2016 Jul;25(7):714-722. [doi: 10.1089/jwh.2014.5059] [Medline: 27111125]
11. Odom EC, Li R, Scanlon KS, Perrine CG, Grummer-Strawn L. Reasons for earlier than desired cessation of breastfeeding. Pediatrics 2013 Mar;131(3):e726-e732. [doi: 10.1542/peds.2012-1295] [Medline: 23420922]
12. Lok KY, Chow CLY, Shing JSY, Smith R, Lam CCO, Bick D, et al. Feasibility, acceptability, and potential efficacy of an innovative postnatal home-based breastfeeding peer support programme in Hong Kong: a feasibility and pilot randomised controlled trial. Int Breastfeed J 2021 Apr 13;16(1):34 [FREE Full text] [doi: 10.1186/s13006-021-00381-5] [Medline: 33849582]
13. Penetration rate of leading social networks in Hong Kong as of 3rd quarter 2020. Statista. URL: https://www.statista.com/statistics/412500/hk-social-network-penetration/ [accessed 2022-02-11]

14. Uzunçakmak T, Gökşin İ, Ayaz-Alkaya S. The effect of social media-based support on breastfeeding self-efficacy: a randomised controlled trial. Eur J Contracept Reprod Health Care 2021 Jul 12:1-7. [doi: 10.1080/13625187.2021.1946500] [Medline: 34251949]
15. Bell ML. New guidance to improve sample size calculations for trials: eliciting the target difference. Trials 2018 Nov 05;19(1):605 [FREE Full text] [doi: 10.1186/s13063-018-2894-y] [Medline: 30396364]
16. World Health Organization. Indicators for Assessing Breast-Feeding Practices. Geneva, Switzerland: World Health Organization; 1991.
17. Ip W, Yeung L, Choi K, Chair S, Dennis C. Translation and validation of the Hong Kong Chinese version of the breastfeeding self-efficacy scale-short form. Res Nurs Health 2012 Oct 12;35(5):450-459. [doi: 10.1002/nur.21493] [Medline: 22693037]
18. Dennis C. The breastfeeding self-efficacy scale: psychometric assessment of the short form. J Obstet Gynecol Neonatal Nurs 2003 Nov;32(6):734-744. [doi: 10.1177/0884217503258459] [Medline: 14649593]
19. Ho Y, McGrath JM. A Chinese version of Iowa Infant Feeding Attitude Scale: reliability and validity assessment. Int J Nurs Stud 2011 Apr;48(4):475-478. [doi: 10.1016/j.ijnurstu.2010.09.001] [Medline: 20961544]
20. Mora ADL, Russell DW, Dungy CI, Losch M, Dusdieker L. The Iowa Infant Feeding Attitude Scale: Analysis of Reliability and Validity1. J Appl Social Pyschol 1999 Nov;29(11):2362-2380. [doi: 10.1111/j.1559-1816.1999.tb00115.x]
21. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009 Jun 29;338(jun29 1):b2393-b2393 [FREE Full text] [doi: 10.1136/bmj.b2393] [Medline: 19564179]
22. Stata statistical software: release 16. StataCorp. 2020 Sep 01. URL: https://www.stata.com/ [accessed 2022-02-22]

Abbreviations

BSES-SF: Breastfeeding Self-Efficacy Scale-Short Form
IIFAS: Iowa Infant Feeding Attitude Scale
mHealth: mobile health

Edited by G Eysenbach; submitted 25.07.21; peer-reviewed by M Herron, L Amir; comments to author 16.09.21; revised version received 17.01.22; accepted 17.01.22; published 09.03.22.

Please cite as:
Lok KYW, Ko RWT, Fan HSL, Chau PH, Wong JYH, Wang MP, Tsang V
Feasibility and Acceptability of an Online WhatsApp Support Group on Breastfeeding: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2022;11(3):e32338
URL: https://www.researchprotocols.org/2022/3/e32338
doi:10.2196/32338
PMID:35262504

Protocol

Strengthening Social Capital to Address Isolation and Loneliness in Long-term Care Facilities During the COVID-19 Pandemic: Protocol for a Systematic Review of Research on Information and Communication Technologies

Idrissa Beogo[1,2], RN, MBA, PhD; Drissa Sia[3,4], MD, PhD; Eric Tchouaket Nguemeleu[3,5], PhD; Junqiang Zhao[1], MSc; Marie-Pierre Gagnon[6], PhD; Josephine Etowa[1], RN RM, PhD

[1]School of Nursing, Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada

[2]College of Nursing, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

[3]Département des sciences infirmières, Université du Québec en Outaouais, Saint-Jerôme, QC, Canada

[4]Département de médecine sociale et préventive, École de santé publique, Université de Montréal, Montréal, QC, Canada

[5]Département de gestion, d'évaluation et de politique de santé, École de santé publique, Université de Montréal, Montréal, QC, Canada

[6]Faculté des sciences infirmières, Université Laval, Québec, QC, Canada

Corresponding Author:
Idrissa Beogo, RN, MBA, PhD
School of Nursing
Faculty of Health Sciences
University of Ottawa
451 Smith road
Roger-Guindon Hall
Ottawa, ON, K1H 8M5
Canada
Phone: 1 6135625800 ext 8148
Email: ibeogo@uottawa.ca

Abstract

Background: The COVID-19 pandemic has had the greatest impact in long-term care facilities (LTCFs) by disproportionately harming older adults and heightening social isolation and loneliness (SIL). Living in close quarters with others and in need of around-the-clock assistance, interactions with older adults, which were previously in person, have been replaced by virtual chatting using information and communication technologies (ICTs). ICT applications such as FaceTime, Zoom, and Microsoft Teams video chatting have been overwhelmingly used by families to maintain residents' social capital and subsequently reduce their SIL.

Objective: Because of the lack of substantive knowledge on this ever-increasing form of social communication, this systematic review intends to synthesize the effects of ICT interventions to address SIL among residents in LTCFs during the COVID-19 period.

Methods: We will include studies published in Chinese, English, and French from December 2019 onwards. Beyond the traditional search strategy approach, 4 of the 12 electronic databases to be queried will be in Chinese. We will include quantitative and intervention studies as well as qualitative and mixed methods designs. Using a 2-person approach, the principal investigator and one author will blindly screen eligible articles, extract data, and assess risk of bias. In order to improve the first round of screening, a pilot-tested algorithm will be used. Disagreements will be resolved through discussion with a third author. Results will be presented as structured summaries of the included studies. We plan to conduct a meta-analysis if sufficient data are available.

Results: A total of 1803 articles have been retrieved to date. Queries of the Chinese databases are ongoing. The systematic review and subsequent manuscript will be completed by the fall of 2022.

Conclusions: ICT applications have become a promising avenue to reduce SIL by providing a way to maintain communication between LTCF residents and their families and will certainly remain in the post–COVID-19 period. This review will investigate

and describe context-pertinent and high-quality programs and initiatives to inform, at the macro level, policy makers and researchers, frontline managers, and families. These methods will remain relevant in the post–COVID-19 era.

International Registered Report Identifier (IRRID): DERR1-10.2196/36269

(JMIR Res Protoc 2022;11(3):e36269) doi:10.2196/36269

KEYWORDS

ICT, long-term care facilities, COVID-19, social isolation, loneliness, pandemic, implementation sciences, protocol; nursing home; long-term care; mental health; aging; older adults; virtual communication; virtual care; information technology; healthcare; healthcare sector; health care

Introduction

Background

The COVID-19 pandemic has disproportionately harmed older adults and subsequently heightened their social isolation and loneliness (SIL) [1,2] and suicidal vulnerability [3-5]. Long-term care facilities (LTCFs) are the most profoundly affected sector with the highest mortality rate of 81% [6,7]. Apart from being the hardest hit during the pandemic, residents have already been experiencing long-lasting SIL before the beginning of the COVID-19 pandemic [8]. To address this SIL, information and communication technologies (ICTs) have become key resources for families to maintain a socio-emotional connection with their loves secluded in LTCFs [9]. Families remain LTCF residents' strongest support system. The Canadian Institutes for Health Information contends that the majority (82%) of older adults need family involvement in their instrumental activities of daily living [10-12].

Systematic [13] and Cochrane [14] reviews undertaken before the COVID-19 pandemic were inconclusive with regard to the effect of ICT in reducing SIL in older adults. However, the pandemic has confirmed the importance of ICT's applications worldwide to provide remote chatting conferencing—a feature that maintains vital interactions between families and their loved ones. Owing to stringent public health restrictions on physical access to LTCFs, ICT has been the only alternative to maintaining social capital and subsequently reducing SIL among seniors in LTCFs [15]. Residents in these settings are more fraught with SIL, varying from 40% [16] to 71.6% [17]. As noted by Abbasi [18], during the current COVID-19 pandemic, SIL resulted in an uptake of antidepressants, antipsychotics, and anxiolytics in LTCF residents. Along with being physically and psychologically separated from the community, many older adults find that relocation to a LTCF itself is a stressful life event. The literature clearly contends that the majority of older adults are reluctant to be relocated [19-21]. Many older adults are widowed, under guardianship or tutorship, and an unknown proportion—4% in the Canadian general population in 2018 [22]—sel-identify as lesbian, gay, bisexual, transgender, questioning, or two-spirit [23]. These vulnerable older adults often become further isolated once admitted into a LTCF and experience greater levels of SIL [24], which underlines the necessity for providing interactive connections with others.

SIL negatively affects quality of life and is associated with an increase in all-cause mortality, an effect that is slightly stronger in men than in women [25-27]. Owing to the public health restrictions imposed during the COVID-19 pandemic, LTCF residents had limited contact with staff and were often secluded in their rooms and no longer able to partake in communal meals, in-person activities, or family visits. The threat of infection and the loss of contact with loved ones has contributed to SIL and has likely exaggerated subsequent negative outcomes in older people. Nevertheless, SIL in LTCFs has just started being considered in the literature. With an increasing incidence of COVID-19, the use of ICT applications has skyrocketed to facilitate social communication [15], providing a necessary support for older people in long-term care [28].

The use of ICT to reduce SIL has been extensively studied, including the effects of internet-based interventions [29-33] and Humanoid Robot approaches [34-36]. Methods currently used in LTCFs to connect older adults with their families and friends range from the conventional telephone to web-based platforms such as Skype, FaceTime, Zoom, or Google Meet (for a review see Banskota et al [37] or Chen and Schulz [38]). Zamir et al [39] employed an intercare home group "Skype" to reduce SIL in 3 care homes and found that video calls reduced feelings of loneliness in residents seemed acceptable and was a feasible, low-cost model, especially during times of public crisis such as during the COVID-19 pandemic.

Research Question and Objectives

This review intends to assess the effect of ICT interventions implemented in LTCFs to address SIL among residents during the COVID-19 pandemic. The following objectives will be considered to address this research question:

1. To synthesize the effects of ICT interventions to address SIL in LTCF residents during the COVID-19 period;
2. To identify studies that use ICT, namely through a varied function of communication such as messaging or chat, video, voice mail, or photo as a strategy for interaction and connection with older family members living in LTCFs;
3. To measure the impact of ICT on the interaction between families and their family members in LTCF facilities.

Rationale for This Study

Older adults have been relying on family members to monitor their "health, well-being, and safety" through virtual visits during the COVID-19 pandemic [40]. SIL in older people has been identified as a risk factor for premature mortality [27] for both poor physical (eg, cardiovascular and obesity) [41-43] and psychiatric health (eg, depression and anxiety) [44]. The significance of communication technology has been featured prominently in local and national news outlets that have

highlighted the stories of residents and families being connected via various ICT applications during the pandemic [45]. Liotta et al [46] explained how increased social connectedness was a powerful tool for nursing home residents that it decreased SIL during the pandemic. Thus, innovative use of digital tools can provide a method to address this urgent public health matter on a long-term basis [47]. As opportunities offered by ICT applications have the potential for long-term solutions for the COVID-19 pandemic and postpandemic period, this systematic review will help inform policy and practice interventions in this area.

The proposed systematic review is also necessary to shed light on the pre–COVID-19 pandemic literature on the impact of ICT on SIL. Systematically examining the evidence on the association between ICT and SIL will help establish up-to-date knowledge to develop best practices and support evidence-based policy decision-making.

Methods

Identification of Data Sources and Studies

This systematic review will be conducted following the Cochrane Collaboration methods [48] and the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) checklist to ensure the completeness of this protocol. We will consult the Synthesis Without Meta-analysis (SWiM) guidelines to guide the use of alternative synthesis methods. The review is registered with the OSF registries [49]. The search will be performed in English, French, and Chinese. A pilot exploratory search on Ovid MEDLINE will be undertaken to create a robust string that is well calibrated in order to improve the likelihood of retrieving articles that are as relevant as possible. This interactive process will include both free vocabulary and descriptors. Beyond the traditional approach that emphasizes on French or English search strings, this review will be enhanced by searching on Chinese databases. Tables 1 and 2 exhibit the search strategies in English for the Ovid MEDLINE database and in Chinese for the China National Knowledge Infrastructure (CNKI) database. The final search strategy involved the following databases: PsycINFO, Ovid MEDLINE, Embase, CINAHL, Cochrane Library, Web of Science, Communication & Mass Media Complete, Association for Computing Machinery (ACM) Digital Library, IEEE Xplore, CNKI, WanFang, Weipu (VIP), and SinoMed. Two review authors, the principal investigator (IB) and a PhD student (JZ), who is a native Mandarin speaker and writer, will canvas all the titles and abstracts; a third review author (DS) will resolve the conflicts. A pilot test will be implemented using a pilot-tested algorithm, which is shown in Figure S1 in Multimedia Appendix 1.

Table 1. Ovid MEDLINE search strategy (to be modified as needed for other databases).

Number	Query
#1	(lonel* or 'social connect*' or connectedness or 'social distanc*' or aloneness or solitude or Seclu* or confin* or separat* or quarantine* or remote* or 'emotional isolation' OR Quarantine). .ab,kw,ti.
#2	exp Loneliness/ or exp Quarantine/
#3	#1 OR #2
#4	(isolat* or deprivation or network or support).ab,kw,ti. . AND social.ab,kw,ti.
#5	exp Social Isolation/
#6	#4 OR #5
#7	#3 OR #6
#8	('Long-Term Care' or 'Assisted-Living Facilit*' or 'Homes for the Aged' or 'Nursing Home*' or Geriatrics or 'homes for the aged' or 'Housing for the Elderly').ab,ti.
#9	(long-term-care OR Geriatrics OR 'Older Adult*' OR elde* OR senior OR aged OR retirement).ab,kw,ti. AND (Home, OR facilit* OR residen*).ab,kw,ti.
#10	exp Nursing Homes/ OR exp Housing for the Elderly/
#11	#8 OR #9 OR #10
#12	(Pandemi* or epidemi* or andemi* OR Outbreak) or (coronavirus or COVID-19 or SARS-COV2).ab,ti.+kw
#13	exp pandemic/ or exp pandemic/ OR COVID-19/ OR SARS-CoV-2/
#14	#12 OR #13
#15	('Digital technology' OR Zoom OR facebook OR 'information technology' OR Skype OR 'FaceTime' OR 'cell phone*' or smartphone* or on-line or online or web-based or 'webbased' or 'web based' or 'world wide web' OR 'Cellular Phone*' or 'mobile phone*' or internet).ab,kw,ti.
#16	information technology/ OR information technology/ OR Cell Phone/ OR smartphone/ OR Online Social Networking/ OR internet/
#17	#15 OR #16
#18	#7 AND #11 AND #14 AND #17

Table 2. Planned search strategy in CNKI.

Number	Query
#1	TKA=社会资本 OR 社会隔绝 OR 社会隔离 OR 社会疏远 OR 社会疏离 OR 社会距离 OR 社群隔离 OR 隔离 OR 空间隔离 OR 情感隔离 OR 孤独 OR 孤独感 OR 生活质量
#2	TKA=老年
#3	TKA=传染病 OR 疫情 OR 冠状病毒 OR COVID-19 OR 新型冠状病毒肺炎 OR 新冠肺炎 OR 冠状病毒肺炎
#4	TKA=计算机通信网络 OR 信息通信技术 OR 信息通讯技术 OR 信息技术 OR 网络 OR 互联网 OR 智能手机 OR 便携式电话 OR 手机 OR 电话 OR 短信 OR 视频 OR 微信
#5	#1 AND #2 AND #3 AND #4

Selection and Data Extraction

All retrieved articles will be uploaded into Rayyan Intelligent Systematic Review [50]. After the removal of duplicates, 2 reviewers will independently screen the titles and abstracts identified by the literature search for inclusion. By April 30, 2022, the full text of potentially relevant articles will be screened to determine final inclusion, followed by a data extraction phase. To increase the reliability of screening by the two independent reviewers, a random sample of articles will be screened on the basis of the eligibility criteria in a pilot test phase. For the screening, a structured algorithm that was previously validated by IB, SD, and ETN will be used.

A standardized data extraction grid that has been developed and piloted will be employed to extract data from all full texts. This includes the following: authors, year of publication, language, design, objectives, participants' characteristics, ICT intervention, outcomes, population, setting, and the fields of the AMSTAR checklist. The *K* statistic will then be calculated to determine the intrarater agreement for study inclusion [51]. Studies excluded during the screening phase will be recorded along with the reason for exclusion by each reviewer. As we intend to complete the review by October 2022, an updated search will be run shortly before this time in order to capture any recent peer-reviewed publications.

Quality Assessment

Given the potential heterogeneity in study designs, IB and JZ will independently assess the methodological quality of studies using the design-specific appraisal tool the Cochrane risk-of-bias (RoB 2) for randomized clinical trials, or the Newcastle-Ottawa Scale also used in nonobservational cohort and case-control studies.

Inclusion and Exclusion Criteria

Inclusion and exclusion criteria will be based on the Population, Interventions, Comparators and designs, Outcomes (PICO) framework, summarized below.

Population (P)

This systematic review will consider studies that target SIL reduction as an outcome in older adults aged 65 years and over living in LTCFs (eg, nursing homes or assisted living arrangements). We will exclude studies of the following persons: (1) with a terminal illness, or (2) who are hospitalized, or (3) with severe neurocognitive disorders, or (4) with severely impaired cognition (measured by specific tools such as the Mini-Mental State Examination [52]), or (5) targeting community dwellers.

Intervention (I)

This systematic review will focus on the use of information technology, namely through communication modes such as chat, video, voice mail, or photo, which maintain or improve the connection between the older adults in LTCFs and their families. This sector of the Canadian health system—and possibly worldwide—appears to be the poor counterpart of all the segments, above all in the new technologies of communication [53].

In addition to regular communication technology such as the telephone, the main ICT intervention component must be based on the use of the internet to fulfill social networking. This is one of 7 elements that can help older adults maintain their independence, proposed in a white paper by the Center for Technology and Aging [54]. Targeted interventions can be delivered individually or in groups and can take place over one or more sessions of various time frames. Any type of digital tool will be considered, including computers, smartphones, or tablets, with the ultimate goal of addressing SIL using commercial applications including Facebook or Zoom for conversation. Any form of connection involving an important face-to-face component in the conversation, or for the purpose of medical treatments, will be excluded.

Comparator (C) or Designs

We will include quantitative studies, specifically randomized controlled trials (RCTs) and quasi‐RCTs (including cluster designs), quasi-experimental, cohort, cross-sectional, and pre-post intervention studies. Qualitative and mixed methods studies will be also included. We will exclude all ICT-based therapeutic interventions, although they also have an interactive component and are capable of reducing SIL. Studies that compare ICT interventions to alternative ones such as visits through widows or contactless control groups during the pandemic will be included. Telehealth or telemedicine, as defined by the World Health Organization [55], although delivered by video, will be excluded as their main intention is not SIL reduction. Further comparison is foreseen through between-group comparisons involving; for instance, phone calls versus calls with a visual component.

Outcome (O)

This systematic review will target the following outcomes, irrespective of whether a psychometric measure is used or not.

The primary outcomes were as follows:

- Measures of SIL (ie, scores on any appropriate and validated tool);
- Measures of SIL through proxy outcomes including the following: companionship, friendship, feeling of being forgotten and not belonging, and connection with family.

We will exclude interventions that include an important face-to-face component or technologies that do not support an interactive component.

The secondary outcomes were as follows:

- Self-report measures of symptoms of depression (ie, scores on any self-report questionnaire that is designed to quantify the severity of symptoms of depression);
- Self-report measures of quality of life (ie, scores on any self-report questionnaire that is designed to allow people to rate their quality of life, either overall or within specific domains).

Timeline (T)

This study was carried out over the COVID-19 pandemic from December 2019 onward.

Data Synthesis and Analysis

Extracted data from all included studies will be summarized in tabular format. Data will be categorized and aggregated by type of intervention and type of setting (eg, nursing homes or assisted living arrangements). A narrative synthesis will be completed. Significant as well as nonsignificant results will be collected, analyzed, and discussed within the relevant outcome category. We plan to run a subgroup narrative analysis. A meta-analysis is planned on the basis of the quality and quantity of data, and heterogeneity will be measured through I^2 statistics. All the limitations will be discussed.

Availability of Data and Material

The data sets generated and analyzed during this study, which would be necessary to interpret, replicate, and build on the findings reported in the review article will be made publicly available as requested by the funding institution. All requests should be addressed to the corresponding author as these data will be stored on a secured server of the Université de Saint-Boniface.

Ethical Considerations

As this systematic review is part of a "Social isolation and loneliness project," we received ethical approval from the Ethics Committees for Research of the University of Ottawa (H-08-21-7314) the University of Moncton (dossier 2021-073) and the Research Ethics Board of the Primary Care and Population Health Research Sector of the CIUSSS of the Capitale-Nationale (2021-2303, _SPPL).

Patient involvement

No patients will be involved. Patients will not be invited to comment on the study protocol design and were not consulted as to how this work may inform patient-relevant outcomes or how a patient might interpret results. However, findings will be disseminated to the public and health care professional networks via conferences, publications, and presentations.

Results

In this review, no patient will be involved. Data extraction and analysis, as well as writing of the manuscript, are expected to be completed by the end of summer 2022.

Discussion

According to the International Federation on Ageing, the number one emerging issue faced by seniors in Canada is keeping older people socially connected and active [56]. SIL can be a chronic issue that burdens older adults. This has sharply augmented with the arrival of COVID-19. Active and vigorous programs exist throughout the world to address SIL such as the following: "End Loneliness" in the United Kingdom [57], Danmark spiser sammen [58], "ALONE" of Ireland [59], or even "Better Together," the Canadian family and caregiver presence initiative [60,61] that has resulted in a significant involvement of families (ie, 20% of families spending over 10 hours per week in supportive activities) [62]. Despite this, more work is needed, and ICT interventions are a good fit.

Virtual communication and technologies that have come to the forefront as the primary mode for LTCF residents during the COVID-19 pandemic appear to be a promising new avenue to maintain social connections and capitalize on the ties among families, their loved ones, and the "outside world." Furthermore, many older adults have higher levels of eHealth literacy with the baby boomer generation becoming older seniors. Nonetheless, many LTCFs do not have the technological capabilities to support modern-day technologies. This is one area of health care system that the COVID-19 pandemic has shed light on. Indeed, in 2020, Canada Infoway reported that LTCFs health technology is the least funded component of the entire health system; a reality that may also be applicable to the rest of the world [53].

At the same time, the current number of studies on ITC does not reflect the mortality inflicted by COVID-19 in long-term care. When and if visiting condition return to "normal," greater efforts will be required to further develop and promote a secure way of virtually connecting LTCF residents with their loves in the community. Furthermore, initiatives should be tailored to address individual social isolation needs. This review is a step toward highlighting the need for more high-quality program evaluation of interventions and other initiatives implemented in the course of the COVID-19 pandemic. The findings of this systematic review will help draw attention from relevant stakeholders of health systems, or specifically LTCF-oriented ones, to address this urgent issue. Although SIL is a socially complex issue that requires a multi-sectorial approach, knowledge gathered and synthesized from this exercise may inform the actions of governments, researchers, and frontline LTCF managers.

Acknowledgments

This paper stemmed from the research grant protocol of the principal investigator's research program on aging and health technology, granted by the Canadian Institutes for Health Research (CIHR) for the program "Implementation Science Teams – Strengthening Pandemic Preparedness in Long-Term Care." We would also like to thank our information specialist Frederic Bergeron for his help in developing the search strategy and selecting databases. This is project is funded by the CIHR for the program "Implementation Science Teams – Strengthening Pandemic Preparedness in Long-Term Care" (operating grant fund FRN 174865). The CIHR had no role in the process (study design, execution, data analyses, or published findings).

Authors' Contributions

IB conceived the initial idea for the study and is the guarantor of the review. IB, SD, JE, MPG, JZ, and ETN were involved in writing the protocol, undertaking the preliminary literature review, as well as providing their expert input in LTCF and digital technology. They were also involved in the editing of the protocol and search strategy. JZ designed the Chinese search strategy. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
First round screening algorithm.
[PNG File , 381 KB - resprot_v11i3e36269_app1.png]

References

1. Hwang T, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. Int Psychogeriatr 2020 Oct;32(10):1217-1220 [FREE Full text] [doi: 10.1017/S1041610220000988] [Medline: 32450943]
2. He W, Zhang Z, Li W. Information technology solutions, challenges, and suggestions for tackling the COVID-19 pandemic. Int J Inf Manage 2021 Apr;57:102287 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102287] [Medline: 33318721]
3. Conejero I, Berrouiguet S, Ducasse D, Leboyer M, Jardon V, Olié E, et al. [Suicidal behavior in light of COVID-19 outbreak: Clinical challenges and treatment perspectives]. Encephale 2020 Jun;46(3S):S66-S72 [FREE Full text] [doi: 10.1016/j.encep.2020.05.001] [Medline: 32471707]
4. Sheffler JL, Joiner TE, Sachs-Ericsson NJ. The Interpersonal and Psychological Impacts of COVID-19 on Risk for Late-Life Suicide. Gerontologist 2021 Jan 21;61(1):23-29 [FREE Full text] [doi: 10.1093/geront/gnaa103] [Medline: 32959869]
5. Wand APF, Zhong B, Chiu HFK, Draper B, De Leo D. COVID-19: the implications for suicide in older adults. Int Psychogeriatr 2020 Oct;32(10):1225-1230 [FREE Full text] [doi: 10.1017/S1041610220000770] [Medline: 32349837]
6. Grant K. 81% of COVID-19 deaths in Canada were in long-term care – nearly double OECD average. The Globe and Mail. 2020. URL: https://www.theglobeandmail.com/canada/article-new-data-show-canada-ranks-among-worlds-worst-for-ltc-deaths/ [accessed 2022-03-16]
7. Pandemic Experience in the Long-Term Care Sector: How Does Canada Compare With Other Countries? Canadian Institute for Health Information. 2020. URL: https://www.cihi.ca/sites/default/files/document/covid-19-rapid-response-long-term-care-snapshot-en.pdf [accessed 2022-03-16]
8. El Haj M, Altintas E, Chapelet G, Kapogiannis D, Gallouj K. High depression and anxiety in people with Alzheimer's disease living in retirement homes during the covid-19 crisis. Psychiatry Res 2020 Sep;291:113294 [FREE Full text] [doi: 10.1016/j.psychres.2020.113294] [Medline: 32763552]
9. Kanekar A, Sharma M. COVID-19 and Mental Well-Being: Guidance on the Application of Behavioral and Positive Well-Being Strategies. Healthcare (Basel) 2020 Sep 12;8(3):336 [FREE Full text] [doi: 10.3390/healthcare8030336] [Medline: 32932613]
10. Dementia in long-term care. Canadian Institute for Health Information. URL: https://www.cihi.ca/en/dementia-in-canada/dementia-care-across-the-health-system/dementia-in-long-term-care#:~:text=Within%20long%2Dterm%20care%20homes,or%20trauma)%20was%2087%25 [accessed 2022-03-16]
11. Recherche quantitative originale – Santé mentale positive et changement perçu de la santé mentale chez les adultes au Canada pendant la deuxième vague de la pandémie de COVID-19. Government of Canada. 2021 Sep 27. URL: https://www.canada.ca/fr/sante-publique/services/rapports-publications/promotion-sante-prevention-maladies-chroniques-canada-recherche-politiques-pratiques/vol-41-no-11-2021/sante-mentale-positive-changement-percu-adultes-canada-deuxieme-vague-pandemie-covid-19.html [accessed 2022-03-16]
12. Bangerter LR, Van Haitsma K, Heid AR, Abbott K. "Make Me Feel at Ease and at Home": Differential Care Preferences of Nursing Home Residents. Gerontologist 2016 Aug;56(4):702-713 [FREE Full text] [doi: 10.1093/geront/gnv026] [Medline: 26035895]

13. 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]
14. Noone C, McSharry J, Smalle M, Burns A, Dwan K, Devane D, et al. Video calls for reducing social isolation and loneliness in older people: a rapid review. Cochrane Database Syst Rev 2020 May 21;5:CD013632 [FREE Full text] [doi: 10.1002/14651858.CD013632] [Medline: 32441330]
15. Sacco G, Lléonart S, Simon R, Noublanche F, Annweiler C, TOVID Study Group. Communication Technology Preferences of Hospitalized and Institutionalized Frail Older Adults During COVID-19 Confinement: Cross-Sectional Survey Study. JMIR Mhealth Uhealth 2020 Sep 18;8(9):e21845 [FREE Full text] [doi: 10.2196/21845] [Medline: 32896832]
16. Trybusińska D, Saracen A. Loneliness in the Context of Quality of Life of Nursing Home Residents. Open Med (Wars) 2019;14:354-361 [FREE Full text] [doi: 10.1515/med-2019-0035] [Medline: 31157300]
17. Prieto-Flores M, Forjaz MJ, Fernandez-Mayoralas G, Rojo-Perez F, Martinez-Martin P. Factors associated with loneliness of noninstitutionalized and institutionalized older adults. J Aging Health 2011 Feb;23(1):177-194. [doi: 10.1177/0898264310382658] [Medline: 20881107]
18. Abbasi J. Social Isolation-the Other COVID-19 Threat in Nursing Homes. JAMA 2020 Aug 18;324(7):619-620. [doi: 10.1001/jama.2020.13484] [Medline: 32692848]
19. Banerjee D, D'Cruz M, Sathyanarayana Rao T. Coronavirus disease 2019 and the elderly: Focus on psychosocial well-being, agism, and abuse prevention – An advocacy review. J Geriatr Ment Health 2020;7(1):4-10. [doi: 10.4103/jgmh.jgmh_16_20]
20. Laing W. Strategic commissioning of long term care for older people: Can we get more for less? LaingBuisson. 2014. URL: https://www.patientlibrary.net/tempgen/762.pdf [accessed 2022-03-16]
21. Brownie S, Horstmanshof L, Garbutt R. Factors that impact residents' transition and psychological adjustment to long-term aged care: a systematic literature review. Int J Nurs Stud 2014 Dec;51(12):1654-1666. [doi: 10.1016/j.ijnurstu.2014.04.011] [Medline: 24813582]
22. A statistical portrait of Canada's diverse LGBTQ2+ communities. Statistics Canada. 2021. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/210615/dq210615a-eng.htm [accessed 2022-03-16]
23. Rétablir la confiance : La COVID-19 et l'avenir des soins de longue durée. La Société royale du Canada. URL: https://rsc-src.ca/fr/r%C3%A9tablir_la_confiance_la_covid-19_et_l_avenir_des_soins_de_longue_dur%C3%A9e [accessed 2022-03-16]
24. Kneale D, Henley J, Thomas J, French R. Inequalities in older LGBT people's health and care needs in the United Kingdom: a systematic scoping review. Ageing Soc 2021 Mar;41(3):493-515 [FREE Full text] [doi: 10.1017/S0144686X19001326] [Medline: 34531622]
25. Musich S, Wang SS, Hawkins K, Yeh CS. The Impact of Loneliness on Quality of Life and Patient Satisfaction Among Older, Sicker Adults. Gerontol Geriatr Med 2015;1:2333721415582119 [FREE Full text] [doi: 10.1177/2333721415582119] [Medline: 28138454]
26. Singh K, Srivastava S. Loneliness and quality of life among elderly people. J Psychosoc Res 2014;9(1):11-18. [doi: 10.1007/978-94-007-0753-5_101124]
27. Rico-Uribe LA, Caballero FF, Martín-María N, Cabello M, Ayuso-Mateos JL, Miret M. Association of loneliness with all-cause mortality: A meta-analysis. PLoS One 2018;13(1):e0190033 [FREE Full text] [doi: 10.1371/journal.pone.0190033] [Medline: 29300743]
28. De R, Pandey N, Pal A. Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. Int J Inf Manage 2020 Dec;55:102171 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102171] [Medline: 32836633]
29. Dickens AP, Richards SH, Greaves CJ, Campbell JL. Interventions targeting social isolation in older people: a systematic review. BMC Public Health 2011 Aug 15;11:647 [FREE Full text] [doi: 10.1186/1471-2458-11-647] [Medline: 21843337]
30. Ballantyne A, Trenwith L, Zubrinich S, Corlis M. 'I feel less lonely': what older people say about participating in a social networking website. Quality Ageing Older Adults 2010 Sep 13;11(3):25-35. [doi: 10.5042/qiaoa.2010.0526]
31. Blažun H, Saranto K, Kokol P, Vošner J. Information and communication technology as a tool for improving physical and social activity of the elderly. NI 2012 (2012) 2012;2012:26 [FREE Full text] [Medline: 24199042]
32. Bradley N, Poppen W. Assistive technology, computers and Internet may decrease sense of isolation for homebound elderly and disabled persons. TAD 2003 May 21;15(1):19-25. [doi: 10.3233/tad-2003-15104]
33. Blažun H, Saranto K, Rissanen S. Impact of computer training courses on reduction of loneliness of older people in Finland and Slovenia. Comput Hum Behav 2012 Jul;28(4):1202-1212 [FREE Full text] [doi: 10.1016/j.chb.2012.02.004]
34. Beer J, Takayama L. Mobile remote presence systems for older adults: acceptance, benefits, and concerns. 2011 Presented at: HRI'11: International Conference on Human-Robot Interaction; March 6-9, 2011; Lausanne. [doi: 10.1145/1957656.1957665]
35. Wada K, Shibata T. Social Effects of Robot Therapy in a Care House - Change of Social Network of the Residents for Two Months. 2007 Presented at: 2007 IEEE International Conference on Robotics and Automation; April 10-14, 2007; Rome. [doi: 10.1109/robot.2007.363156]
36. Bickmore T, Caruso L, Clough-Gorr K, Heeren T. 'It's just like you talk to a friend' relational agents for older adults. Interacting with Computers 2005 Dec;17(6):711-735 [FREE Full text] [doi: 10.1016/j.intcom.2005.09.002]

37. Banskota S, Healy M, Goldberg E. 15 Smartphone Apps for Older Adults to Use While in Isolation During the COVID-19 Pandemic. West J Emerg Med 2020 Apr 14;21(3):514-525 [FREE Full text] [doi: 10.5811/westjem.2020.4.47372] [Medline: 32302279]
38. Chen YRR, 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]
39. Zamir S, Hennessy C, Taylor A, Jones R. Intergroup 'Skype' Quiz Sessions in Care Homes to Reduce Loneliness and Social Isolation in Older People. Geriatrics (Basel) 2020 Nov 11;5(4):90 [FREE Full text] [doi: 10.3390/geriatrics5040090] [Medline: 33187242]
40. Hado E, Friss Feinberg L. Amid the COVID-19 Pandemic, Meaningful Communication between Family Caregivers and Residents of Long-Term Care Facilities is Imperative. J Aging Soc Policy 2020;32(4-5):410-415. [doi: 10.1080/08959420.2020.1765684] [Medline: 32441209]
41. National Academies of Sciences, Engineering, and Medicine. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System. Washington, DC: The National Academies Press; 2020.
42. Shankar A, McMunn A, Banks J, Steptoe A. Loneliness, social isolation, and behavioral and biological health indicators in older adults. Health Psychol 2011 Jul;30(4):377-385. [doi: 10.1037/a0022826] [Medline: 21534675]
43. 2030 vision: The best – and worst – futures for older people in the UK. ILC-UK. 2014 Mar 13. URL: https://tinyurl.com/2p8pnane [accessed 2022-03-16]
44. Evans IEM, Martyr A, Collins R, Brayne C, Clare L. Social Isolation and Cognitive Function in Later Life: A Systematic Review and Meta-Analysis. J Alzheimers Dis 2019;70(s1):S119-S144 [FREE Full text] [doi: 10.3233/JAD-180501] [Medline: 30372678]
45. Ramesar V. Absence of volunteers creates staffing pressures at N.S. nursing homes. CBC News. 2020 Aug 20. URL: https://www.cbc.ca/news/canada/nova-scotia/nursing-homes-pandemic-recreation-covid-19-1.5692866 [accessed 2022-03-16]
46. Liotta G, Marazzi M, Orlando S, Palombi L. Is social connectedness a risk factor for the spreading of COVID-19 among older adults? The Italian paradox. PLoS One 2020;15(5):e0233329 [FREE Full text] [doi: 10.1371/journal.pone.0233329] [Medline: 32437377]
47. Eghtesadi M. Breaking Social Isolation Amidst COVID-19: A Viewpoint on Improving Access to Technology in Long-Term Care Facilities. J Am Geriatr Soc 2020 May;68(5):949-950 [FREE Full text] [doi: 10.1111/jgs.16478] [Medline: 32277470]
48. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. Hoboken, NJ: John Wiley & Sons; 2008.
49. Social Capitalization Through Information and Communication Technologies to Reduce Social Isolation and Loneliness in Long-Term Care Homes residents during the COVID-19 epidemic: a Systematic Review of Published Research. OSF. 2021. URL: https://osf.io/sjhpb/ [accessed 2022-03-17]
50. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
51. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
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. Sondage national des infirmières et infirmiers du Canada 2020 : Utilisation des technologies de santé numériques au travail. Inforoute Santé du Canada. 2020 May 11. URL: https://tinyurl.com/yckpmay2 [accessed 2022-03-16]
54. Technologies to Help Older Adults Maintain Independence: Advancing Technology Adoption. Center for Technology and Aging. 2009. URL: http://www.thescanfoundation.org/sites/default/files/CTA%20White%20Paper%20Final%20070809.pdf [accessed 2022-03-16]
55. Telemedicine: opportunities and developments in Member States: report on the second global survey on eHealth. World Health Organization. URL: https://apps.who.int/iris/handle/10665/44497 [accessed 2022-03-16]
56. Report on the Social Isolation of Seniors. Government of Canada. URL: https://tinyurl.com/2p8f2xv3 [accessed 2022-03-16]
57. Safeguarding the Convoy: A call to action from the Campaign to End Loneliness. Age UK Oxfordshire. URL: https://www.campaigntoendloneliness.org/wp-content/uploads/Safeguarding-the-Convoy.-A-call-to-action-from-the-Campaign-to-End-Loneliness.pdf [accessed 2022-03-16]
58. Eat together and create new communities. Sagen A. 2018. URL: https://wwwaeldresagendk/om-aeldresagen/lige-nu/sammenmed [accessed 2022-03-16]
59. Have a Laugh for Loneliness with ALONE this Christmas. ALONE. 2018. URL: http://aloneie/have-a-laugh-for-loneliness-with-alone-thischristmas/ [accessed 2022-03-16]
60. Dokken D, Barden A, Tuomey M, Giammarinaro N, Johnson B. Families as Care Partners: Implementing the Better Together Initiative Across a Large Health System. J Clin Outcomes Manag 2020;27(1).
61. Dokken D, Kaufman J, Johnson B, Perkins S, Benepal J, Roth A. Changing Hospital Visiting Policies: From Families as “Visitors” to Families as Partners. J. Clin. Outcomes Manag 2015;22(1).

62. Turcotte M, Sawaya C. Senior care: Differences by type of housing. Statistics Canada. URL: https://www150.statcan.gc.ca/n1/pub/75-006-x/2015001/article/14142-eng.htm [accessed 2022-03-16]

Abbreviations

CIHR: Canadian Institutes for Health Research
CNKI: China National Knowledge Infrastructure
ICT: information and communication technologies
LTCF: long-term care facility
PICO: Population, Interventions, Comparators and designs, Outcomes
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
RCT: randomized controlled trial
SIL: social isolation and loneliness
SWiM: Synthesis Without Meta-analysis

Edited by T Leung; submitted 08.01.22; this is a non–peer-reviewed article;accepted 22.02.22; published 24.03.22.

Please cite as:
Beogo I, Sia D, Tchouaket Nguemeleu E, Zhao J, Gagnon MP, Etowa J
Strengthening Social Capital to Address Isolation and Loneliness in Long-term Care Facilities During the COVID-19 Pandemic: Protocol for a Systematic Review of Research on Information and Communication Technologies
JMIR Res Protoc 2022;11(3):e36269
URL: https://www.researchprotocols.org/2022/3/e36269
doi:10.2196/36269
PMID:35275841

Protocol

Reporting of Model Performance and Statistical Methods in Studies That Use Machine Learning to Develop Clinical Prediction Models: Protocol for a Systematic Review

Colin George Wyllie Weaver[1], BSc; Robert B Basmadjian[1], MSc; Tyler Williamson[1], PhD; Kerry McBrien[1,2], MPH, MD; Tolu Sajobi[1], PhD; Devon Boyne[3], PhD; Mohamed Yusuf[4], MSc; Paul Everett Ronksley[1], PhD

[1]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[2]Department of Family Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[3]Department of Oncology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[4]Faculty of Science & Engineering, Manchester Metropolitan University, Manchester, United Kingdom

Corresponding Author:
Paul Everett Ronksley, PhD
Department of Community Health Sciences
Cumming School of Medicine
University of Calgary
Teaching, Research, and Wellness Building 3E18B
3280 Hospital Drive NW
Calgary, AB, T2N 4Z6
Canada
Phone: 1 403 220 8820
Email: peronksl@ucalgary.ca

Abstract

Background: With the growing excitement of the potential benefits of using machine learning and artificial intelligence in medicine, the number of published clinical prediction models that use these approaches has increased. However, there is evidence (albeit limited) that suggests that the reporting of machine learning–specific aspects in these studies is poor. Further, there are no reviews assessing the reporting quality or broadly accepted reporting guidelines for these aspects.

Objective: This paper presents the protocol for a systematic review that will assess the reporting quality of machine learning–specific aspects in studies that use machine learning to develop clinical prediction models.

Methods: We will include studies that use a supervised machine learning algorithm to develop a prediction model for use in clinical practice (ie, for diagnosis or prognosis of a condition or identification of candidates for health care interventions). We will search MEDLINE for studies published in 2019, pseudorandomly sort the records, and screen until we obtain 100 studies that meet our inclusion criteria. We will assess reporting quality with a novel checklist developed in parallel with this review, which includes content derived from existing reporting guidelines, textbooks, and consultations with experts. The checklist will cover 4 key areas where the reporting of machine learning studies is unique: modelling steps (order and data used for each step), model performance (eg, reporting the performance of each model compared), statistical methods (eg, describing the tuning approach), and presentation of models (eg, specifying the predictors that contributed to the final model).

Results: We completed data analysis in August 2021 and are writing the manuscript. We expect to submit the results to a peer-reviewed journal in early 2022.

Conclusions: This review will contribute to more standardized and complete reporting in the field by identifying areas where reporting is poor and can be improved.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42020206167; https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=206167

International Registered Report Identifier (IRRID): RR1-10.2196/30956

(JMIR Res Protoc 2022;11(3):e30956) doi:10.2196/30956

KEYWORDS

machine learning; clinical prediction; research reporting; statistics; research methods; clinical prediction models; artificial intelligence; modeling; eHealth; digital medicine; prediction

Introduction

Machine learning is commonly defined as computers learning from data [1], especially when they are not explicitly programmed [2]. There is considerable optimism around the potential benefits of using machine learning approaches in medicine, including for prediction [3-9]. This is in part because of the ever-increasing volume and variety of health care data collected and readily available for clinical and research purposes. Although few prediction models developed using machine learning are currently used in clinical care [10], some researchers believe that machine learning will greatly increase predictive performance relative to more traditional regression techniques and will replace many regression-based prediction models and tasks previously performed by clinicians, such as diagnostic image interpretation [4,10,11]. These increases in performance purport to result in more accurate diagnoses and prognoses for patients and may ultimately improve patient outcomes.

Excitement about these potential benefits has led to a recent increase in the number of publications reporting the use of machine learning approaches for clinical prediction. While several reviews have evaluated the reporting quality of studies that use machine learning for clinical prediction [12-17], they all assessed the studies' general reporting using items similar to those found on the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) checklist instead of using machine learning–specific items (eg, related to tuning). We therefore do not know much about the quality of machine learning–specific reporting in studies that use machine learning to develop clinical prediction models. However, there is evidence (albeit limited) suggesting that the reporting of machine learning–specific aspects in these studies is poor [16]. A review by Christodoulou et al [16] evaluated the performance of clinical prediction models developed using machine learning versus regression and found poor reporting of the machine learning tuning methods.

Another review of the reporting of studies that use machine learning for clinical prediction by O'Shea et al [17] assessed the reporting of studies that used convolutional neural networks for the radiological diagnosis of cancer; this review used the CLAIM (Checklist for Artificial Intelligence in Medical Imaging) [18], which was developed for machine learning–based studies. However, most of the CLAIM items are similar to items on non–machine learning checklists and are not unique to machine learning (eg, validation, handling of missing data). Among the unique items, most were specific to artificial neural networks. Two additional reviews evaluated the reporting of studies that used machine learning for clinical prediction [15,16]. However, both these reviews used the TRIPOD checklist to assess reporting quality and therefore the items used were not specific to machine learning. Lastly, a review by Yusuf et al [14] assessed the reporting quality of studies that used machine learning to develop clinical prediction models, but the review focused on the reporting of information about study participants and did not assess reporting of unique machine learning items.

There are no broadly accepted guidelines for the reporting of clinical prediction studies that use machine learning [19-22]. While the TRIPOD statement [23], published in 2015, outlines items that should be reported in all sections of a paper on the development or validation of a clinical prediction model, the specificity of these items as they relate to machine learning methods is limited. Consequently, there is a need for additional guidance on how to report the methods and results of these studies, which differ from traditional regression studies [20,24]. Six reporting guidelines or checklists have been developed for studies that use machine learning approaches [21]: CLAIM [18], MI-CLAIM (Minimum Information About Clinical Artificial Intelligence Modeling) [25], MINIMAR (Minimum Information for Medical Artificial Intelligence Reporting) [26], The Machine Learning Reproducibility Checklist [27], and guidelines by Luo et al [28] and Stevens et al [29]. However, these checklists have substantial limitations, including considerably overlapping with TRIPOD instead of being an extension to TRIPOD. The checklist by Luo et al [28] was the only one to use a Delphi consensus process, a recommended approach [30]. To our knowledge, none of these checklists have become widely used in machine learning clinical prediction research [19-22]. Therefore, the TRIPOD group is developing an extension to the TRIPOD statement for studies that use machine learning—TRIPOD-Artificial Intelligence [31].

The absence of broadly accepted reporting guidelines may contribute to poor reporting. Additionally, studies that use machine learning often have more complex study designs than regression-based clinical prediction model development studies (eg, more complex resampling procedures to accommodate unbiased tuning or model comparison). Complete reporting in clinical prediction research, including research using machine learning, allows for assessment of model validity, assessment of risk of bias in predictive performance, and enables readers to trust and be able to use and externally validate models. Given the lack of reviews assessing machine learning–specific reporting and the evidence of poor reporting of tuning methods, we aim to conduct a systematic review to assess reporting quality using a sample of 100 recently published studies that use machine learning to develop clinical prediction models. The review will consider reporting within 4 key areas where the reporting of machine learning studies is unique: modelling steps, model performance, statistical methods, and presentation of models. The specific objectives of this review are to (1) assess the current reporting quality of machine learning–specific aspects (modelling steps, model performance, statistical methods, and presentation of models) in studies that use machine learning to develop clinical prediction models; (2) evaluate whether reporting quality differs by journal discipline; and (3) identify the most common machine learning algorithms, tuning methods, and internal validation procedures currently used in this field.

Methods

Registration and Reporting of Results

This protocol was registered with PROSPERO (registration number CRD42020206167) in September 2020 [32]. The reporting of this protocol follows the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols statement (Multimedia Appendix 1) [33]. The reporting of the review will follow the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [34]. Ethical considerations do not apply to this review as it will exclusively use data from published studies.

Search Methods

MEDLINE will be searched without language restrictions via Ovid using a search strategy developed in consultation with a research librarian (Table 1).

Table 1. MEDLINE (Ovid) search strategy using Ovid MEDLINE(R) and Ovid MEDLINE Epub Ahead of Print, In-Process and Other Non-Indexed Citations and Daily.

Concept	MEDLINE (Ovid) search terms	Explanation
Supervised machine learning algorithm	exp artificial intelligence/[a] OR (artificial intelligence OR machine learning OR supervised learning OR statistical learning OR deep learning OR ensemble learning OR regression tree*[b] OR classification tree* OR "C4.5" OR probability estimation tree* OR random forest* OR support vector machine* OR relevance vector machine* OR artificial neural net* OR deep neural net* OR deep artificial neural net* OR recurrent neural net* OR feedforward neural net* OR feed-forward neural net* OR convolution* neural net* OR perceptron* OR gradient boost* OR adaboost OR k-nearest neighbo* OR nearest neighbo* OR k-NN* OR bayesian network* OR bayesian additive regression tree* OR model tree* OR naive bayes tree* OR bootstrap aggregat* OR hyperparamet* OR hyper-paramet* OR scikit OR tensorflow OR pytorch OR ((neural net* OR decision tree* OR bagging OR boosting OR ensemble* OR tuning OR torch OR CART) AND (cross-valid* OR crossvalid* OR bootstrap* OR predict* OR classifier* OR train OR validat* OR discrimination OR calibration OR ROC curve* OR c-statistic* OR c statistic* or area under the curve OR AUC))).mp[c]	• Names of common supervised machine learning algorithms (and their implementations) as well as terms related to the use of supervised learning algorithms • Terms with non–machine learning meanings as well (eg, bagging) are AND'ed with terms commonly used when discussing prediction models
Animal study	exp animals/[d] NOT humans/	• Animal studies are excluded
Reviews and other non–original research	(review OR meta-analysis OR editorial OR comment OR letter).pt[e] OR (review OR meta-analysis).ti[f]	• Reviews and other publication types not representing original research are excluded
Combining concepts	1 NOT 2 NOT 3	• Combining concepts
Limit to 2019	4 AND "2019*".dp[g]	• We are using a random sample of these studies

[a]/: Medical Subject Heading.

[b]*: search term truncation.

[c].mp: multipurpose (searches titles, abstracts, and author-specified keywords).

[d]exp …/: exploded Medical Subject Heading.

[e].pt: publication type.

[f].ti: title.

[g]dp: date of publication.

The machine learning search terms used by Christodoulou et al [16] and the list of machine learning algorithms they identified were used to develop the search terms for this review. An initial search without time limits yielded 141,401 citations. The purpose of this review is to identify common reporting deficiencies in recently published articles and not to conduct an exhaustive assessment of reporting quality of machine learning clinical prediction models over time. We will therefore review citations from only 2019 in a pseudorandom order until we obtain a fixed sample of 100 included articles. This time-restricted sampling approach aims to limit the number of studies in our review because of the impracticality of including all studies that would meet our inclusion criteria (preliminary pilot screening indicates this would be approximately 30,000). Further, this sampling approach is similar to the approach used in other reviews that have aimed to evaluate reporting quality [35]. The number 100 was chosen to obtain a balance between CI width and the time it will take to review articles and extract data. The year 2019 was chosen because it is recent and was not dominated by articles on COVID-19 or influenced by

COVID-19–related changes to manuscript writing, reviewing, and publishing practices that may impact reporting quality.

Only one database (MEDLINE) will be used because there are differences in how databases record dates of publication and entry and our review is not meant to be exhaustive. We verified that several key journals that publish machine learning clinical prediction studies are indexed in MEDLINE (eg, *Journal of the American Medical Informatics Association*, *Journal of Medical Internet Research*, *Bioinformatics*, *BMJ Open*, *PLOS One*). Grey literature will not be searched as the objective is to assess reporting quality in peer-reviewed literature.

Study Selection

Titles and abstracts will be screened independently by 2 reviewers, and abstracts included by either reviewer will proceed to full-text review. Abstracts will be included if they represent peer-reviewed original research where a supervised machine learning algorithm is used to develop a prediction model. The research must have a human health application and the prediction model must be intended for use in clinical practice (specifically, the prediction model must provide a diagnosis, provide a prognosis, or identify candidates for health care interventions). Further details are provided in the complete exclusion criteria list in Textbox 1.

All study designs, medical disciplines, predictor data types (including images, video, and audio), and outcome data types (binary, categorical, ordinal, measured) will be included. Supervised machine learning algorithms create prediction models by learning the relationship between predictors and outcome and are thus most appropriate for this review. In contrast, unsupervised machine learning algorithms identify patterns in data with no specified outcomes; while they can be used as a step in the development of a prediction model, they cannot create prediction models themselves. The reasons for article exclusion will be recorded at both review stages, with criteria applied hierarchically (Textbox 1). After the first 12 abstracts, the reviewers will discuss discrepancies and clarify exclusion criteria.

The same 2 reviewers will independently review full-text articles using the same exclusion criteria, and disagreements will be resolved by discussion and consensus or by consulting a third individual. After the first 7 full-texts, the reviewers will discuss discrepancies and clarify exclusion criteria.

To obtain a random sample of 100 included articles, we will screen citations in the same pseudorandom order at both the title and abstract screening and the full-text review stages. After removing duplicate PubMed identifiers, citations will be sorted by PubMed unique identifier (smallest to largest). In this order, citations will then be assigned a pseudorandomly generated number in R using the rnorm() function with set.seed(3486) and will be sorted by this number (smallest to largest). Title and abstract screening will be conducted using this order until there are 120 studies that will proceed to full-text review. If fewer than 100 of these are included after full-text review, more titles and abstracts will be screened. No full-text articles will be assessed for inclusion after 100 have been included.

A PRISMA flow diagram will be used to describe the stages and reasons for exclusion. Only studies assigned a pseudorandom number less than or equal to the 100th included study will be described in the flow diagram (ie, studies that could not have been included in the review because 100 were obtained will not be described in the flow diagram). The number of records identified by the search strategy without time limits and in 2019 will be reported. Agreement between the reviewers for both the title and abstract screening and the full-text review stages will be assessed by calculating Cohen κ coefficient (including the pilot of 12 abstracts and 7 full-text articles). Google Sheets and Google Forms will be used for study selection, data extraction, and data storage.

Textbox 1. Exclusion criteria (applied in order, with the first-listed applicable criterion recorded as the reason for exclusion).

1. The study is not original research.

- Description
  - Includes reviews, meta-analyses, editorials, comments, or letters

2. No prediction model is developed.

- Description
  - Examples are a study that validates a previously developed machine learning prediction model and a robotics study (using machine learning to develop a robot)
  - Studies are also excluded if the primary aim is to identify important predictors or estimate associations between predictors and outcomes rather than to provide individual predictions (develop a prediction model). Studies focused on assessing the incremental value of a new predictor are included
  - Studies are also excluded if their primary objective is to develop a new machine learning methodology or type of machine learning and they place less emphasis on the prediction model(s) developed as a test of the methodology. This exclusion reason is only used in full-text review
- Rationale
  - Most of the reporting items of interest are only applicable to model development studies

3. No supervised machine learning algorithm is used.

- Description
  - The study must use at least one supervised machine learning algorithm to develop the prediction model. Studies that only use machine learning to select predictors or process data and not to develop the prediction model are excluded
  - The following are considered machine learning algorithms:
    - Decision tree (classification or regression)
    - Random forest
    - Support vector machine
    - Relevance vector machine
    - Artificial neural network (including single layer perceptron)
    - Boosted tree algorithms (including gradient boosting machines)
    - k-nearest neighbors
    - Bayesian network (only if they learn the graph structure from the data)
    - Bayesian additive regression tree
    - Model tree, including naïve Bayes tree
    - Algorithms employing boosting or bagging
    - Ensemble methods if they include at least one machine learning algorithm listed above
  - The following (by themselves) are not considered machine learning algorithms:
    - Generalized linear model (regression)
    - Regularized (penalized) regression, including lasso, ridge, and elastic net
    - Partial least squares
    - Principal component regression
    - Generalized additive models
    - Regression with splines
    - Multivariate adaptive regression splines
    - Generalized estimating equations
    - Bayesian regression
    - Naïve Bayes

        - Discriminant analysis
        - Genetic algorithms
        - Latent class analysis
        - Fuzzy logic
        - Autoencoder (because not supervised)
    - Studies that only use unsupervised or reinforcement learning are not included because the reporting requirements should be different (no prediction model is developed)
    - Studies that use machine learning to select predictors or process data are excluded because the algorithms are not used to develop the prediction model itself and hence many of the reporting items do not apply (eg, tuning)

- Rationale
    - These algorithms not considered machine learning are mostly regression-based or regression-like and do not learn the structure of the relationship between predictors and outcomes using the data. The implementation, reporting, and risks of bias in these algorithms should be similar to those in regression; the purpose of this review is to assess the reporting quality of machine learning algorithms that may be different or more challenging to report than regression or have different risks of bias

4. The research does not have human health application.

- Description
    - Examples are animal study, in vitro study, and research to increase the efficiency of a laboratory procedure

5. The application is not to provide a diagnosis, provide a prognosis, or identify candidates for health care interventions.

- Description
    - The prediction model must provide a diagnosis (predict whether a health condition is currently present, screen for a condition, or determine the subtype of a condition), provide a prognosis (predict future health status, outcome, or event), or identify patients who would be better candidates for a health care intervention
    - The prediction model developed must be intended for use in clinical practice either on publishing of the results or on further validation. Similarly, if the aims of the study are mostly etiological and not to develop an accurate prediction model, the study will be excluded
    - Additionally, the prediction model must directly provide a diagnosis or prognosis or identify candidates for health care interventions; models simply aiding humans to do so are excluded (eg, image contouring or segmentation). Models diagnosing image regions individually are included (eg, diagnosing lesions as benign or malignant)
    - The following types of studies will be excluded:
        - Studies that develop case definitions to identify individuals or events in health databases, which would be used for surveillance, quality improvement, or research and not for routine clinical care
        - Studies that develop prediction models to quantify a patient aspect that is not a health condition (eg, height)
        - Studies that develop prediction models to optimize the operation of health care technology or a procedure (eg, a prediction model to determine for which patients setting A should be set to X)

6. The paper is for a conference.

- Description
    - Papers published as part of a conference are excluded

7. The paper is not peer-reviewed.

8. The research is reported only in abstract, poster, or presentation form.

9. Full text cannot be found.

10. Full text is not available in English.

Data Extraction

Reporting Quality

Reporting quality will be assessed independently by 2 reviewers using a checklist developed in parallel with this review. This checklist includes content derived from existing reporting guidelines, textbooks, and consultations with machine learning experts; a Delphi procedure will not be used. We will then pilot the checklist for 5 studies and revise as needed. Checklist items will be specific to machine learning studies and will not overlap with non–machine learning reporting guidelines (ie, the checklist will be an extension to the TRIPOD checklist). We will not use the TRIPOD adherence assessment form [36] because we are focusing on the machine learning–specific reporting aspects; our review will evaluate these unique reporting items rather than provide a comprehensive reporting assessment of the 100 machine learning studies. This is because we expect the reporting completeness of the aspects not specific to machine learning to be similar to the reporting completeness found in prior reviews of the reporting quality of clinical prediction model studies [24,37,38].

In a draft of the checklist, we identified 4 key areas of study reporting that the checklist will focus on: modelling steps (order and data used for each step), model performance (eg, reporting the performance of each model compared), statistical methods (eg, describing the tuning approach), and presentation of models (eg, specifying the predictors that contributed to the final model). For the assessment of reporting quality, each item can receive 1 of 3 assessments: "reported," "not reported," (including incomplete reporting) and "not applicable." Disagreements will be resolved by discussion and consensus or by consulting a third individual. Agreement on reporting quality assessment between the 2 reviewers will be measured by Cohen κ coefficient for each item and will exclude the pilot of 5 articles.

Reviewers will make their assessments based on what is reported in the text, any supplementary appendices, and the text of any related studies (eg, by the same authors) that are referenced as providing more information on methods used. Reviewers will not contact study authors or look at statistical code to assess reporting quality. Well-reported studies should report these items clearly in words to make it easy for others to understand and critically assess the methods and results. While providing statistical code may be necessary to facilitate understanding and exactly replicate the approach used, in isolation it will not be regarded as sufficient to consider an item "reported." If there is more than one outcome, analysis, or machine learning algorithm used in the study, reporting quality will be assessed for the primary analysis only. If the primary analysis is not clearly specified, the outcome, analysis, or algorithm mentioned first in the discussion section will be considered primary.

Quality Assessment (Risk of Bias)

While the focus of our review will be on reporting quality, we recognize that study quality (ie, risk of bias) may be associated with reporting quality. For this reason, risk of bias within our identified studies will be assessed using the Prediction Model Risk of Bias Assessment Tool (PROBAST) [39]. PROBAST was chosen because it is the only risk of bias tool designed for studies that develop clinical prediction models. Given that we will not have the clinical expertise and context to effectively evaluate prediction models that span multiple medical specialties, we will not assess the first 3 PROBAST domains (participants, predictors, and outcome). Our quality assessment will therefore be limited to the analysis domain within PROBAST. We will not use PROBAST signaling question 4.3 because it requires clinical context to assess or the signaling question 4.9 because it lacks relevance for machine learning studies. We will use a modified version of signaling question 4.8 to improve clarity for machine learning studies (adding the term "data leakage"). We will assign each study a low, high, or unclear risk of bias based on the analysis signaling questions. One reviewer will assess risk of bias using PROBAST's published form [39], and a second reviewer will verify this information. Disagreements will be resolved by discussion and consensus or by consulting a third individual.

Other Data Extraction

In addition to reporting quality and risk of bias, we will extract general study characteristics, prediction model characteristics, tuning method, and internal validation procedures (Textbox 2). These will be extracted by 1 reviewer using a standardized data extraction form and verified by a second reviewer. Disagreements will be resolved by discussion and consensus or by consulting a third individual. We will contact study authors for further information where needed.

Textbox 2. Study characteristics to extract in addition to those for reporting quality and risk of bias.

1. General characteristics

- First author's last name
- Year of publication
- Title
- Journal
- Journal discipline (clinical, radiology, computer science or engineering, other)
- Country of first author

2. Prediction model characteristics (if applicable, consider primary outcome or analysis only)

- Type of outcome predicted (diagnosis of a condition, prognosis of a condition, identification of candidates for a health care intervention)
- Type of data emphasized the most in the title and abstract (coded or structured data, imaging data or video, language [text, audio], genomic or other 'omic data, signal [eg, electrocardiogram])
- Sample size used for model development

3. Tuning methods (in primary analysis)

- Were the data used to tune or were fixed or default tuning parameters used? (yes the data were used to tune, no, not mentioned, unclear—if no or not mentioned, the remaining tuning questions are not applicable)
- Search method (grid, random, ad hoc, Bayesian optimization, gradient-based optimization, evolutionary optimization, other, unclear)

4. Internal validation procedures (in primary analysis)

- Resampling method, if applicable (data split one or more times, k-fold crossvalidation [including repeated k-fold crossvalidation, leave-one-out crossvalidation], bootstrap, other, unclear)
- Was there a holdout test set not used at any point in the training process? (yes, no, unclear)
- Was nested crossvalidation (or crossvalidation) used in a manner as good as in a holdout test set? (yes, no, unclear)

5. Other methods

- Machine learning algorithms used

Data Synthesis

The primary outcomes are the proportions of included studies that report each of the checklist items across the 4 domains: modelling steps, model performance, statistical methods, and presentation of models. The proportion of applicable items reported per study will also be described using median, first quartile, and third quartile. Items deemed not applicable will be excluded from the proportion denominators. All included studies will be analyzed. As this is a systematic review assessing reporting quality and not synthesizing the findings of individual studies, a meta-analysis will not be performed [35].

We will complete 1 subgroup analysis (objective 2); because of the small sample size, this will be considered exploratory. The proportion of applicable items reported per study will be described by journal discipline (4 categories: clinical, radiology, computer science or engineering, and other). We hypothesize that articles in computer science and engineering journals will report machine learning methods and results more completely, especially the tuning approach.

Ninety-five percent CIs will be reported, including binomial exact CIs for proportions. With 100 included studies, the binomial exact CI widths will be a maximum of 0.20 (at a proportion of 0.5) and a width of 0.18 at a proportion of 0.25, assuming the item is applicable to all studies. One hypothesis test, a Kruskal-Wallis rank sum test, will be performed to determine whether there are differences between the journal disciplines in terms of the proportion of applicable items reported per study. R (version 3.6.2; R Foundation for Statistical Computing) will be used for all analyses [40]. All data collected and code used for this review will be made public via a data repository.

Results

We completed data analysis in August 2021 and are writing the manuscript. We expect to submit the results to a peer-reviewed journal in early 2022.

Discussion

Despite the growing use of machine learning for clinical prediction, there is little understanding of the completeness of reporting in these studies, especially reporting aspects unique to machine learning. This review will identify common reporting deficiencies in these areas in 100 recently published studies. Dissemination of these findings to researchers developing machine learning models will increase awareness and the importance of these deficiencies and consequently improve

reporting completeness. The novel checklist that is an extension to TRIPOD will also provide an easy way for researchers to improve reporting and allow peer reviewers and editors to assess reporting completeness prior to publication.

We will examine whether reporting quality differs by journal discipline in a subgroup analysis. Results from this secondary objective may highlight key differences in reporting across disciplines and facilitate targeted dissemination or educational activities for disciplines where reporting quality is poor. We will also document the machine learning algorithms currently used in the literature and the accompanying tuning methods and internal validation procedures. Future studies may be able to use this information to identify and improve areas where less preferred approaches (eg, those introducing biases or reducing performance) are often used.

Based on the results of previous reviews on reporting quality [12-14,16,17], we expect to find overall poor reporting quality in our review and possibly a high percentage of studies with high risk of bias. However, we will identify areas where reporting is generally complete and areas where it is lacking in order to improve reporting in this field. Until TRIPOD-Artificial Intelligence [20] is published, the checklist we developed for this review will be the most comprehensive tool available to assess reporting quality of methodological aspects unique to clinical prediction studies that use machine learning. We hope that the checklist we developed and the understanding we gain of areas where reporting is poor will aid the development of TRIPOD-Artificial Intelligence [30].

This review has several strengths. It is the first review to assess the reporting quality of machine learning–specific aspects in published clinical prediction studies that use machine learning. The reporting checklist developed for this review is also novel and focuses on items particularly relevant to models developed using machine learning and not simply on items close to the well-known TRIPOD items [23]. However, this study also has some limitations. First, this review will focus on reporting and will not assess whether the machine learning techniques employed are in line with preferred methodological practices or might introduce bias or reduce performance. The review will only assess risk of bias using the PROBAST analysis domain, which is not specific to machine learning approaches. Second, we chose not to use the more robust Delphi procedure to develop the checklist but chose a smaller and expedited expert consultation. Third, the fixed sample size of 100 studies is relatively small and some estimates of reporting quality, especially in the subgroup analysis, will be imprecise. Fourth, the included studies are likely to be very heterogeneous, especially in terms of the types of machine learning (‘omics to computer vision) and disciplines, which will result in differences in conventions and terminology. This will make it more difficult to assess reporting completeness and may challenge the idea that a single reporting checklist can apply across heterogeneous uses of machine learning for clinical prediction. Findings from this review will help determine if certain types of machine learning may require separate or additional reporting checklists. Fifth, assigning journals to discipline categories requires arbitrary determination, but we believe that our comparison of reporting by discipline remains useful. Finally, it is possible that the search strategy is biased toward better reported studies (ie, studies that use method terms in their abstracts). We have tried to mitigate such bias by keeping our search terms broad.

The use of machine learning in the setting of clinical prediction is growing rapidly, but there is evidence (albeit limited) that suggests that the reporting of machine learning–specific aspects within these studies is poor. This is the first review to assess the reporting of these aspects and enable measurement of current reporting completeness and identification of areas where reporting is lacking. Both the identification of these areas and the novel checklist developed for the review will contribute to more standardized and complete reporting in this field.

Acknowledgments

We thank Diane Lorenzetti, PhD (University of Calgary), for helping us develop the search strategy. DB is supported by a postdoctoral fellowship from the Canadian Institutes of Health Research. CGWW was supported by studentships from the Canadian Institutes of Health Research, Alberta Innovates (201610416), and the University of Calgary. Neither these organizations nor the academic institutions we are affiliated with played a role in developing this protocol.

Authors' Contributions

CGWW conceived of the study with TW, led the design of the study, and wrote the first draft of the manuscript with contributions from PER. All authors contributed to the design of the study, critically reviewed the manuscript, and contributed to revising the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-P checklist.
[PDF File (Adobe PDF File), 124 KB - resprot_v11i3e30956_app1.pdf]

References

1. Hastie T, Tibsharani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. New York, NY: Springer; 2009.
2. Koza J, Bennett IF, Andre D, Keane M. Automated design of both the topologysizing of analog electrical circuits using genetic programming. In: Gero JS, Sudweeks F. eds. Artificial Intelligence in Design '96. Dordrechtpringer; 1996:151-170.
3. Darcy AM, Louie AK, Roberts LW. Machine learning and the profession of medicine. JAMA 2016 Feb 09;315(6):551-552. [doi: 10.1001/jama.2015.18421] [Medline: 26864406]
4. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA 2017 Aug 08;318(6):517-518. [doi: 10.1001/jama.2017.7797] [Medline: 28727867]
5. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA 2018 Apr 03;319(13):1317-1318. [doi: 10.1001/jama.2017.18391] [Medline: 29532063]
6. Beam AL, Kohane IS. Translating artificial intelligence into clinical care. JAMA 2016 Dec 13;316(22):2368-2369. [doi: 10.1001/jama.2016.17217] [Medline: 27898974]
7. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med 2016 Sep 29;375(13):1216-1219 [FREE Full text] [doi: 10.1056/NEJMp1606181] [Medline: 27682033]
8. Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: humanism and artificial intelligence. JAMA 2018 Jan 02;319(1):19-20. [doi: 10.1001/jama.2017.19198] [Medline: 29261830]
9. Chen JH, Asch SM. Machine learning and prediction in medicine - beyond the peak of inflated expectations. N Engl J Med 2017 Jun 29;376(26):2507-2509 [FREE Full text] [doi: 10.1056/NEJMp1702071] [Medline: 28657867]
10. Deo RC. Machine learning in medicine. Circulation 2015 Nov 17;132(20):1920-1930 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.001593] [Medline: 26572668]
11. Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. Health Aff (Millwood) 2014 Jul;33(7):1123-1131. [doi: 10.1377/hlthaff.2014.0041] [Medline: 25006137]
12. Andaur Navarro CL, Damen JAA, Takada T, Nijman SWJ, Dhiman P, Ma J, et al. Completeness of reporting of clinical prediction models developed using supervised machine learning: a systematic review. medRxiv. 2021 Jul 01. URL: https://www.medrxiv.org/content/10.1101/2021.06.28.21259089v1 [accessed 2021-07-31]
13. Dhiman P, Ma J, Navarro CA, Speich B, Bullock G, Damen JA, et al. Reporting of prognostic clinical prediction models based on machine learning methods in oncology needs to be improved. J Clin Epidemiol 2021 Oct;138:60-72 [FREE Full text] [doi: 10.1016/j.jclinepi.2021.06.024] [Medline: 34214626]
14. Yusuf M, Atal I, Li J, Smith P, Ravaud P, Fergie M, et al. Reporting quality of studies using machine learning models for medical diagnosis: a systematic review. BMJ Open 2020 Mar 23;10(3):e034568. [doi: 10.1136/bmjopen-2019-034568] [Medline: 32205374]
15. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020 Mar 25;368:m689 [FREE Full text] [doi: 10.1136/bmj.m689] [Medline: 32213531]
16. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol 2019 Jun;110:12-22. [doi: 10.1016/j.jclinepi.2019.02.004] [Medline: 30763612]
17. O'Shea RJ, Sharkey AR, Cook GJR, Goh V. Systematic review of research design and reporting of imaging studies applying convolutional neural networks for radiological cancer diagnosis. Eur Radiol 2021 Oct 16;31(10):7969-7983 [FREE Full text] [doi: 10.1007/s00330-021-07881-2] [Medline: 33860829]
18. Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. Radiol Artif Intell 2020 Mar 01;2(2):e200029 [FREE Full text] [doi: 10.1148/ryai.2020200029] [Medline: 33937821]
19. El Naqa I, Ruan D, Valdes G, Dekker A, McNutt T, Ge Y, et al. Machine learning and modeling: Data, validation, communication challenges. Med Phys 2018 Oct 24;45(10):e834-e840 [FREE Full text] [doi: 10.1002/mp.12811] [Medline: 30144098]
20. Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. Lancet 2019 Apr 20;393(10181):1577-1579. [doi: 10.1016/S0140-6736(19)30037-6] [Medline: 31007185]
21. Ibrahim H, Liu X, Denniston AK. Reporting guidelines for artificial intelligence in healthcare research. Clin Exp Ophthalmol 2021 Jul 25;49(5):470-476. [doi: 10.1111/ceo.13943] [Medline: 33956386]
22. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health 2019 Oct;1(6):e271-e297 [FREE Full text] [doi: 10.1016/S2589-7500(19)30123-2] [Medline: 33323251]
23. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015 Jan 07;350:g7594. [doi: 10.1136/bmj.g7594] [Medline: 25569120]

24. Heus P, Damen JAAG, Pajouheshnia R, Scholten RJPM, Reitsma JB, Collins GS, et al. Poor reporting of multivariable prediction model studies: towards a targeted implementation strategy of the TRIPOD statement. BMC Med 2018 Jul 19;16(1):120-112 [FREE Full text] [doi: 10.1186/s12916-018-1099-2] [Medline: 30021577]
25. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. Nat Med 2020 Sep 09;26(9):1320-1324 [FREE Full text] [doi: 10.1038/s41591-020-1041-y] [Medline: 32908275]
26. Hernandez-Boussard T, Bozkurt S, Ioannidis J, Shah N. MINIMAR (MINimum Information for Medical AI Reporting): Developing reporting standards for artificial intelligence in health care. J Am Med Inform Assoc 2020 Dec 09;27(12):2011-2015 [FREE Full text] [doi: 10.1093/jamia/ocaa088] [Medline: 32594179]
27. Pineau J, Vincent-Lamarre P, Sinha K. Improving reproducibility in machine learning research (a report from the NeurIPS 2019 Reproducibility Program). J Mach Learn Res 2021:22-20 [FREE Full text]
28. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. J Med Internet Res 2016 Dec 16;18(12):e323 [FREE Full text] [doi: 10.2196/jmir.5870] [Medline: 27986644]
29. Stevens LM, Mortazavi BJ, Deo RC, Curtis L, Kao DP. Recommendations for reporting machine learning analyses in clinical research. Circ Cardiovasc Qual Outcomes 2020 Oct;13(10):e006556 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.120.006556] [Medline: 33079589]
30. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med 2010 Feb 16;7(2):e1000217 [FREE Full text] [doi: 10.1371/journal.pmed.1000217] [Medline: 20169112]
31. Collins GS, Dhiman P, Andaur Navarro CL, Ma J, Hooft L, Reitsma JB, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ Open 2021 Jul 09;11(7):e048008 [FREE Full text] [doi: 10.1136/bmjopen-2020-048008] [Medline: 34244270]
32. University of York Centre for Reviews and Dissemination. PROSPERO. PROSPERO: international prospective register of systematic reviews. 2022. URL: https://www.crd.york.ac.uk/prospero/ [accessed 2022-01-26]
33. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 01;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
34. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009 Jul 21;339(jul21 1):b2535-b2535 [FREE Full text] [doi: 10.1136/bmj.b2535] [Medline: 19622551]
35. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. PLoS Med 2007 Mar 27;4(3):e78 [FREE Full text] [doi: 10.1371/journal.pmed.0040078] [Medline: 17388659]
36. Heus P, Damen JAAG, Pajouheshnia R, Scholten RJPM, Reitsma JB, Collins GS, et al. Uniformity in measuring adherence to reporting guidelines: the example of TRIPOD for assessing completeness of reporting of prediction model studies. BMJ Open 2019 Apr 24;9(4):e025611 [FREE Full text] [doi: 10.1136/bmjopen-2018-025611] [Medline: 31023756]
37. Takemura T, Kataoka Y, Uneno Y, Otoshi T, Matsumoto H, Tsutsumi Y, et al. The reporting quality of prediction models in oncology journals: a systematic review. In: Annals of Oncology.: Annals of Oncology. Vol 29. Elsevier BV; 2018 Nov Presented at: European Society for Medical Oncology Asia Congress; Nov 23-25, 2018; Singapore p. ix171. [doi: 10.1093/annonc/mdy433.004]
38. Wynants L, Kent DM, Timmerman D, Lundquist CM, Van Calster B. Untapped potential of multicenter studies: a review of cardiovascular risk prediction models revealed inappropriate analyses and wide variation in reporting. Diagn Progn Res 2019 Feb 22;3(1):6-17 [FREE Full text] [doi: 10.1186/s41512-019-0046-9] [Medline: 31093576]
39. Wolff RF, Moons KG, Riley RD, Whiting PF, Westwood M, Collins GS, PROBAST Group†. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019 Jan 01;170(1):51-58 [FREE Full text] [doi: 10.7326/M18-1376] [Medline: 30596875]
40. R Core Team. R: The R Project for Statistical Computing. R: A Language and Environment for Statistical Computing. Vienna, Austria URL: https://www.r-project.org [accessed 2022-01-25]

Abbreviations

CLAIM: Checklist for Artificial Intelligence in Medical Imaging
MI-CLAIM: Minimum Information About Clinical Artificial Intelligence Modeling
MINIMAR: Minimum Information for Medical Artificial Intelligence Reporting
PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analysis
PROBAST: Prediction Model Risk of Bias Assessment Tool
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by G Eysenbach; submitted 13.06.21; peer-reviewed by R Bajpai, B Kaas-Hansen, V Guardabasso; comments to author 15.09.21; revised version received 09.12.21; accepted 31.12.21; published 03.03.22.

Please cite as:
Weaver CGW, Basmadjian RB, Williamson T, McBrien K, Sajobi T, Boyne D, Yusuf M, Ronksley PE
Reporting of Model Performance and Statistical Methods in Studies That Use Machine Learning to Develop Clinical Prediction Models: Protocol for a Systematic Review
JMIR Res Protoc 2022;11(3):e30956
URL: https://www.researchprotocols.org/2022/3/e30956
doi:10.2196/30956
PMID:35238322

Protocol

Device- and Nondevice-Guided Slow Breathing to Reduce Blood Pressure in Patients with Hypertension: Protocol for a Systematic Review and Meta-analysis

Kamila Shelry de Freitas Gonçalves[1*], PhD[‡]; Ana Carolina Queiroz Godoy Daniel[1*], PhD; José Luiz Tatagiba Lamas[2*], PhD; Henrique Ceretta Oliveira[2*], MSc; Renata C C P Silveira[1*], PhD; Lyne Cloutier[3*], PhD; Eugenia Velludo Veiga[1*], PhD

[1]WHO Collaborating Centre for Nursing Research Development, University of São Paulo, Ribeirão Preto, Brazil

[2]School of Nursing, University of Campinas, Campinas, Brazil

[3]Département des sciences infirmières, Université du Québec à Trois-Rivières, Quebec, QC, Canada

[‡]Interdisciplinary Research Group for Research in Arterial Hypertension

[*]all authors contributed equally

Corresponding Author:
Kamila Shelry de Freitas Gonçalves, PhD
WHO Collaborating Centre for Nursing Research Development
University of São Paulo
R Prof Hélio Lourenço 3900 Vila Monte Alegre
Ribeirão Preto, 14040-902
Brazil
Phone: 55 16 33154321
Email: kamilashelry@icloud.com

Abstract

Background: Physiotherapy can include both device-guided slow breathing (DGSB) and nondevice-guided slow breathing (NDGSB) in the treatment of systemic arterial hypertension.

Objective: The aim of this study is to summarize the effects of DGSB on blood pressure levels of patients with hypertension based on the published literature to date.

Methods: A systematic search of all published randomized controlled trials (RCTs) on the effects of device-guided and nondevice-guided slow breathing in patients with hypertension, without language restriction, was carried out up to a publication date of January 2020 in nine databases: PubMed/MEDLINE, Latin American and Caribbean Health Sciences Literature (LILACS), EMBASE, CENTRAL (Cochrane Central Register of Controlled Trials), Physiotherapy Evidence Database (PEDro), CINAHL (Cumulative Index to Nursing and Allied Health Literature), Scopus, Web of Science, and Livivo. Clinical trial records databases (ClinicalTrials.gov), and bases for the open gray literature, including Gray Literature Report and ProQuest Central (Citation, Abstract or Indexing, and Dissertations and Theses), were also searched for potentially eligible RCTs. The quality assessment of the included studies will be performed using the Cochrane Risk of Bias Tool for Randomized Trials. The overall quality of the evidence for each outcome will be assessed using the GRADE (Grading of Recommendations, Development and Evaluation) system.

Results: As of December 2021, the review was completed and all data from continuous variables referring to blood pressure values (mmHg) were synthesized.

Conclusions: This systematic review will provide a summary of the current evidence on the effects of both DGSB and NDGSB on blood pressure levels. This information can contribute to decision-making by health professionals related to the use of these interventions in patients with hypertension.

Trial Registration: PROSPERO (Prospective International Register of Systematic Reviews) CRD42020147554; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=147554

International Registered Report Identifier (IRRID): RR1-10.2196/33579

(JMIR Res Protoc 2022;11(3):e33579) doi:10.2196/33579

KEYWORDS

hypertension; breathing exercises; device-guided breathing; respirate; systematic review; physical therapy; blood pressure; clinical decision making; health care professional; physiotherapy

Introduction

Hypertension is a multifactorial chronic disease, and is the main risk factor for the development of cardiovascular diseases and chronic kidney disease. Hypertension affects 32% of adults and more than 60% of the elderly, being responsible for half of the deaths from cardiovascular disease (CVD) in Brazil [1]. In addition, the complications of hypertension can lead to decreased work productivity and family income [1,2]. In high-income countries such as Canada, the hypertension prevalence has declined; in middle-income countries such as those of Latin America, Asia, the Middle East, and North Africa, detection and treatment hypertension have enhanced, whereas low detection and treatment rates persist in the poorest nations such as those of sub-Saharan Africa and Oceania [1,3,4].

Considering the high prevalence rates, the treatment of hypertension is necessary, not only to reduce blood pressure levels but also to prevent the development of CVD, cerebrovascular diseases, and kidney diseases. This treatment can be medication, which will be determined according to the blood pressure values obtained either in medical consultations or at home; cardiovascular risk factors; and the presence of target organ damage identified during anamnesis. Nonpharmacological treatment has also been shown to be effective in reducing blood pressure levels in patients with hypertension [1-4], including body weight control, the establishment of healthy eating habits (specifically reducing salt consumption), alcohol consumption control, smoking cessation, stress control, aerobic and isometric physical exercises, and slow breathing with or without device guide [1-4].

The physiotherapy prescription for the treatment of hypertension may include both exercise and device-guided slow breathing (DGSB) or nondevice-guided slow breathing (NDGSB). These breathing exercises consist of slow and deep breathing with 6 to 10 breaths per minute, and can be performed with or without a guiding device. Isometric exercises have been shown to be effective in reducing blood pressure levels, along with aerobic and dynamic exercises [4-6]. However, the application of DGSB remains controversial. Since DGSB activates cardiac and pulmonary stretching receptors, decreases sympathetic activity, and increases parasympathetic activity and vagal tone, thus changing the heart rate and blood pressure, it would be clinically sound to consider that it would reduce blood pressure levels. With blood pressure reduction, there is an increase in baroreflex sensitivity, which promotes improvements in the autonomic balance of patients with hypertension [7].

The American Heart Association reported that there is no strong evidence on the effectiveness of DGSB, whereas the 8th Brazilian Hypertension Guidelines reported a IIa degree of recommendation, level of evidence A [1,4]. A review [8] indicated that there is currently insufficient evidence of data to recommend the routine use of DGSB in patients with hypertension, even though this device has been cleared by the US Food and Drug Administration and the UK National Health Service. In their review, Cernes et al [9] stated that DGSB, as long as it is monitored by a health professional, can be recommended for patients with hypertension who cannot obtain full control of their blood pressure with drug treatment or cannot tolerate the side effects of treatment. Barros et al [10] performed a controlled clinical study with 15 individuals in the control group and 17 in the experimental group, who practiced DGSB for 15 to 20 minutes a day, with 6 to 10 breaths per minute, and concluded that DGSB, in the long term, did not reduce blood pressure values, catecholamine levels, or muscle sympathetic nerve activity in patients with hypertension. However, the use of DGSB was indicated in the 7th Brazilian Hypertension Guidelines [11].

Recommendations for the use of DGSB or NDGSB in clinical practice should be guided by a systematic, high-quality literature review. Recently, Chaddha et al [12] published a review that fulfills this requirement, which compared DGSB to NDGSB (*pranayama*) for 4 weeks in patients with prehypertension and hypertension. The review included 17 studies, with systolic blood pressure reported in 1017 subjects and diastolic blood pressure reported in 964 subjects. Although interesting, this review did not specifically include patients with hypertension and exclusively compared DGSB to *pranayama*. Therefore, a systematic review of the antihypertensive effects of DGSB or NDGSB applied by physical therapists is necessary to provide the best evidence available to clinical physical therapists and patients with hypertension. In addition, it is also important to summarize the evidence on the effectiveness of DGSB or NDGSB compared to usual care.

Accordingly, the aim of this systematic review and meta-analysis is to summarize the effects of DGSB on blood pressure levels of patients with hypertension compared with control conditions (such as minimal intervention, usual care, placebo, and no treatment), other interventions (NDGSB), and when used as an adjunct to other treatments (medication). Thus, the research question for this systematic review of randomized controlled trials (RCTs) is: What are the effects of the prolonged use of device-guided or nondevice-guided slow breathing compared to usual care on the blood pressure values of patients with hypertension?

Methods

Study Design

The design of this systematic review is inspired by the recommendations of the Cochrane Handbook of Systematic Reviews [13] and the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) guidelines [14].

The articles were selected based on inclusion criteria according to the type of study, participants, and intervention.

Inclusion Criteria

Types of Studies

RCTs published up to January 2020 were included in this systematic review, without language or year of publication restrictions.

Types of Participants

Studies including patients with hypertension, with or without comorbidity, who were over 18 years old, of both sexes, and with or without antihypertensive medication treatment were included for review.

Types of Interventions

Interventions considered had to involve DGSB and NDGSB compared to control conditions (such as minimal intervention, usual care, placebo, and no treatment) and interventions as an adjunct to other treatments (medication). Any dosage of device-guided breathing treatment was accepted. Regarding the follow-up time, studies with a duration of a minimum of 8 weeks were considered.

Exclusion Criteria

RCTs that also used other interventions along with DGSB/NDGSB, such as physical activity (aerobic exercises, Tai Chi, resistance training, and isometric exercises); salt reduction and salt substitution; stress control (meditation, Qigong, yoga, progressive muscle and reduction programs for attention-based stress disorders); dietary interventions, including a dietary approach to stop hypertension, low-carbohydrate diet, Mediterranean diet, high-protein diet, low-fat diet, vegetarian diet, paleolithic diet, and low index glycemic/load; and lifestyle interventions (comprehensive lifestyle modification, smoking cessation, alcohol restriction, sleep, home heating, and weight loss) were excluded since it was not possible to identify the specific effect of DGSB/NDGSB in such studies.

Outcome Measures

The primary outcome was the systolic blood pressure and diastolic blood pressure values (measured at home, in the office, or by ambulatory blood pressure monitoring), expressed in mmHg, measured after the interventions, as well as their variations. The secondary outcome was a reduction in the quantity/dosage of drugs administered to control hypertension, if relevant.

Identification and Selection of Studies

A systematic search of all published RCTs on the effects of device and nondevice-guided slow breathing in patients with hypertension, without language restriction, was carried out until January 2020 in nine databases: Pubmed/MEDLINE, Latin American and Caribbean Health Sciences Literature (LILACS), EMBASE, CENTRAL (Cochrane Central Register of Controlled Trials), Physiotherapy Evidence Database (PEDro), CINAHL (Cumulative Index to Nursing and Allied Health Literature), Scopus, Web of Science, and Livivo. Clinical trial record databases (ClinicalTrials.gov), and sites for the open grey literature, such as Gray Literature Report and ProQuest Central (Citation, Abstract or Indexing, and Dissertations and Theses), were also searched. Completed and ongoing RCTs were searched up to January 2020 and, when possible, only peer-review papers were included, because the grey literature was also searched.

Search Strategy

An example of the search strategy used in PubMed/MEDLINE is shown in Textbox 1.

Two reviewers independently analyzed all titles and abstracts retrieved with the search. When there was agreement on a particular record, the full text was analyzed by both reviewers, according to the eligibility criteria. In the presence of disagreement between the reviewers, a third reviewer was convened. When additional information was needed, the authors of the potentially eligible studies were contacted.

Textbox 1. Search strategy in the PubMed/MEDLINE database.

Hypertension [Mesh: NoExp], hypertensi*, high blood pressure, high blood pressures, blood pressure [Mesh], blood pressure, arterial pressures, arterial tension, arterial tensions, Arterial Pressure [Mesh], Arterial pressure, arterial blood pressure, arterial blood pressures, elevated blood pressure, acute hypertension, arterial hypertension, cardiovascular hypertension, controlled hypertension, hypertensive disease, systemic hypertension, increased blood pressure, artery blood pressure, artery pressure, systemic arterial pressure, systemic artery pressure, Hypertension [Mesh: NoExp] OR hypertensi* OR high blood pressure OR high blood pressures OR blood pressure [Mesh] OR blood pressure OR arterial pressures OR arterial tension OR arterial tensions OR Arterial Pressure [Mesh] OR Arterial pressure OR arterial blood pressure OR arterial blood pressures OR elevated blood pressure OR acute hypertension OR arterial hypertension OR cardiovascular hypertension OR controlled hypertension OR hypertensive disease OR systemic hypertension OR increased blood pressure OR artery blood pressure OR artery pressure OR systemic arterial pressure OR systemic artery pressure, Breathing Exercises [Mesh: NoExp], breathing exercises, respiratory muscle training, device-guided breathing, loaded breathing, slow breathing exercises, paced breathing, controlled breathing, breathing exercise, breathing therapy, chest physical therapy, chest physiotherapy, respiration exercise, respiration therapy, respiratory exercise, respiratory physiotherapy, deep inspiration, deep respiration, deep breathing, Breathing Exercises [Mesh: NoExp] OR breathing exercises OR respiratory muscle training OR device-guided breathing OR loaded breathing OR slow breathing exercises OR paced breathing OR controlled breathing OR breathing exercise OR breathing therapy OR chest physical therapy OR chest physiotherapy OR respiration exercise OR respiration therapy OR respiratory exercise OR respiratory physiotherapy OR deep inspiration OR deep respiration OR deep breathing OR deep respiration OR deep breathing, Hypertension [Mesh: NoExp] OR hypertensi* OR high blood pressure OR high blood pressures OR blood pressure [Mesh] OR blood pressure OR arterial pressures OR arterial tension OR arterial tensions OR Arterial Pressure [Mesh] OR Arterial pressure OR arterial blood pressure OR arterial blood pressures OR elevated blood pressure OR acute hypertension OR arterial hypertension OR cardiovascular hypertension OR controlled hypertension OR hypertensive disease OR systemic hypertension OR increased blood pressure OR artery blood pressure OR artery pressure OR systemic arterial pressure OR systemic artery pressure AND Breathing Exercises [Mesh: NoExp] OR breathing exercises OR respiratory muscle training OR device-guided breathing OR loaded breathing OR slow breathing exercises OR paced breathing OR controlled breathing OR breathing exercise OR breathing therapy OR chest physical therapy OR chest physiotherapy OR respiration exercise OR respiration therapy OR respiratory exercise OR respiratory physiotherapy OR deep inspiration OR deep respiration OR deep breathing OR deep respiration OR deep breathing

Two reviewers independently extracted the following data from the included trials: author, publication date, country of publication, study type, sample size, participant characteristics (age, gender), use or not of antihypertensive medications, presence of comorbidities, categories of blood pressure, details of intervention (type of device used in the DGSB, whether DGSB was performed with or without load, how the NDGSB was performed, breaths per minute for DGSB and NDGSB, time of use of the device per day and for how many months), details for blood pressure measurement (device used, type of measurement [home or office], and protocol used for measurement, including preparation), and outcome measures (systolic and diastolic blood pressure). A third reviewer was called in case of disagreement. When necessary, the authors of RCTs included were contacted to provide additional information.

Assessment of Study Characteristics

The quality assessment of the included studies was conducted using the Cochrane Risk of Bias Tool for Randomized Trials (RoB2) [15], which includes a randomization process, deviations from the intended interventions, conflicting result data, result measurement, selection of the reported result, and general biases. The same two reviewers performed an independent assessment. Disagreements between reviewers were resolved by discussion and, if necessary, the opinion of a third reviewer was requested. The same two reviewers performed data extraction, using standardized forms regarding the methodological characteristics of the studies, interventions, and results. Disagreements were again resolved by discussion and, if necessary, the opinion of a third reviewer was requested.

Data Analysis

All data from continuous variables referring to blood pressure values (mmHg) will be synthesized according to the mean difference and respective 95% CIs. Standard deviations were also extracted from the studies for analysis.

The effects of interventions on blood pressure values will be analyzed separately. The data will be evaluated according to the type of intervention (DGSB or NDGSB); however, only studies lasting at least 8 weeks will be considered for meta-analysis (results evaluated after 8 weeks of randomization). Results where there was an intention-to-treat analysis will be used whenever possible.

The presence of statistical heterogeneity between RCTs will be assessed using the I^2 statistic. The quality of the evidence will be considered inconsistent if considerable heterogeneity between the groups (I^2>50%) is observed. When sufficient evidence is available, a funnel graph will be used to investigate possible publication bias**.**

Data Synthesis

The overall quality of the evidence for each outcome will be assessed using the GRADE (Grading of Recommendations, Development and Evaluation) [16] system, regardless of whether or not the information is sufficient to summarize the data in a quantitative analysis. The following five factors will be considered when classifying the quality of the evidence: risk of bias (>25% of the RCTs included in the comparison are classified as having a high risk of bias), inconsistency (I^2>50%), indirectness (>50% of participants were not related to the target audience trial), imprecision (<400 participants in the comparison for continuous outcomes), and publication bias (assessed using a funnel plot when >10 trials are in the same comparison). For each factor not met, the quality of the evidence is reduced by one level (from high to moderate, low, or very low). Single trial comparisons (<400 participants for continuous results) were found to be inconsistent and inaccurate, providing “low quality evidence,” which could be downgraded to “very low quality evidence” if limitations are identified in relation to the risk of bias [16].

The quality of the evidence will be categorized as follows: the evidence is of high quality if the results are consistent in ≥75% of the participants, with a low risk of bias, without publication bias, and with consistent direct and accurate data; further research is unlikely to alter the estimate or confidence in such results. The evidence will be considered to be of moderate quality when one of the five classification factors above is met; further research can alter the estimated effect and impact confidence in the effect in this case. The evidence will be considered to be of poor quality when two of the five classification factors are not met. In this situation, future research is likely to alter the estimated effect and have a significant impact on confidence in the effect. The evidence will be considered to be of very low quality when three of the five classification factors are not met, as any estimate of effect is uncertain in this case [16].

Results

As of December 2021, the review is complete and all data from continuous variables referring to blood pressure values (in mmHg) have been synthesized.

Discussion

This systematic review aims to provide the best available evidence on the effectiveness of DGSB or NDGSB in patients with hypertension, as well as whether DGSB/NDGSB allows for a reduction in the amount/dosage of antihypertensive medication administration. All recommendations in the Cochrane Manual of Systematic Reviews will be followed to ensure that the review is of high quality. It is believed that the results of this systematic review will be important because both DGSB and NDGSB are accessible and can be performed even at home, as long as the use is prescribed and guided by a physical therapist.

In addition, this review is the first to assess NDGSB unrelated to the effect of *pranayama* on the blood pressure of patients with hypertension, since this type of intervention is routinely used in a physiotherapist’s clinical practice, following a completely different method from that adopted in yoga. Therefore, this evidence will inform health care professionals and patients about the potential benefits of this intervention. This review may also identify gaps in the literature that can be addressed in future studies.

Authors' Contributions

KSFG, EVV, and RCCPS developed the research question and methods. KSFG wrote the first draft of the manuscript. KSFG, ACQGD, EVV, JLTL, RCCPS, and LC contributed to the development of the methods and search strategies. All authors contributed to the drafting of the review protocol and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Barroso W, Rodrigues C, Bortolotto L, Mota-Gomes MA, Brandão AA, de Magalhães FAD, et al. Brazilian Guidelines of Hypertension - 2020. Arq Bras Cardiol 2021 Mar;116(3):516-658 [FREE Full text] [doi: 10.36660/abc.20201238] [Medline: 33909761]
2. Rabi D, McBrien K, Sapir-Pichhadze R, Nakhla M, Ahmed SB, Dumanski SM, et al. Hypertension Canada's 2020 Comprehensive Guidelines for the Prevention, Diagnosis, Risk Assessment, and Treatment of Hypertension in Adults and Children. Can J Cardiol 2020 May;36(5):596-624. [doi: 10.1016/j.cjca.2020.02.086] [Medline: 32389335]
3. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021 Sep 11;398(10304):957-980 [FREE Full text] [doi: 10.1016/S0140-6736(21)01330-1] [Medline: 34450083]
4. Whelton P, Carey R, Aronow W, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension 2018 Jun;71(6):e13-e115 [FREE Full text] [doi: 10.1161/HYP.0000000000000065] [Medline: 29133356]
5. Hansen D, Dendale P, Coninx K, Vanhees L, Piepoli MF, Niebauer J, et al. The European Association of Preventive Cardiology Exercise Prescription in Everyday Practice and Rehabilitative Training (EXPERT) tool: a digital training and decision support system for optimized exercise prescription in cardiovascular disease. Concept, definitions and construction methodology. Eur J Prev Cardiol 2017 Jul;24(10):1017-1031. [doi: 10.1177/2047487317702042] [Medline: 28420250]
6. Pescatello L, MacDonald H, Ash G, Lamberti LM, Farquhar WB, Arena R, et al. Assessing the existing professional exercise recommendations for hypertension: a review and recommendations for future research priorities. Mayo Clin Proc 2015 Jun;90(6):801-812. [doi: 10.1016/j.mayocp.2015.04.008] [Medline: 26046413]
7. Schein M, Gavish B, Herz M, Rosner-Kahana D, Naveh P, Knishkowy B, et al. Treating hypertension with a device that slows and regularises breathing: a randomised, double-blind controlled study. J Hum Hypertens 2001 Apr;15(4):271-278. [doi: 10.1038/sj.jhh.1001148] [Medline: 11319676]
8. Mahtani K, Beinortas T, Bauza K, Nunan D. Device-guided breathing for hypertension: a summary evidence review. Curr Hypertens Rep 2016 Apr;18(4):33. [doi: 10.1007/s11906-016-0631-z] [Medline: 27059041]
9. Cernes R, Zimlichman R. Role of paced breathing for treatment of hypertension. Curr Hypertens Rep 2017 Jun;19(6):45. [doi: 10.1007/s11906-017-0742-1] [Medline: 28470470]
10. de Barros S, da Silva GV, de Gusmão JL, de Araújo TG, de Souza DR, Cardoso CG, et al. Effects of long term device-guided slow breathing on sympathetic nervous activity in hypertensive patients: a randomized open-label clinical trial. Blood Press 2017 Dec;26(6):359-365. [doi: 10.1080/08037051.2017.1357109] [Medline: 28724309]
11. Malachias MVB. 7th Brazilian Guideline of Arterial Hypertension: Presentation. Arq Bras Cardiol 2016 Sep;107(Suppl 3) [FREE Full text] [doi: 10.5935/abc.20160140] [Medline: 27819379]
12. Chaddha A, Modaff D, Hooper-Lane C, Feldstein D. Device and non-device-guided slow breathing to reduce blood pressure: a systematic review and meta-analysis. Complement Ther Med 2019 Aug;45:179-184. [doi: 10.1016/j.ctim.2019.03.005] [Medline: 31331557]
13. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions. 2nd edition. Chichester, UK: John Wiley & Sons; 2019.
14. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009 Jul 21;339:b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
15. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
16. Schünemann H, Brozek J, Guyatt G, Oxman A. GRADE Handbook. 2013. URL: https://gdt.gradepro.org/app/handbook/handbook.html [accessed 2022-02-25]

Abbreviations

CENTRAL: Cochrane Central Register of Controlled Trials
CINAHL: Cumulative Index to Nursing and Allied Health Literature
CVD: cardiovascular disease
DGSB: device-guided slow breathing
GRADE: Grading of Recommendations, Development and Evaluation
LILACS: Literatura Latino-Americana e do Caribe em Ciências da Saúde
NDGSB: nondevice-guided slow breathing
PEDro: Physiotherapy Evidence Database
PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analysis
RCT: randomized controlled trial
RoB2: Cochrane Risk of Bias Tool for Randomized Trials

Edited by G Eysenbach; submitted 14.09.21; peer-reviewed by S Omboni, S Hiremath; comments to author 01.11.21; revised version received 23.11.21; accepted 22.12.21; published 04.03.22.

Please cite as:
de Freitas Gonçalves KS, Queiroz Godoy Daniel AC, Tatagiba Lamas JL, Ceretta Oliveira H, Silveira RCCP, Cloutier L, Velludo Veiga E
Device- and Nondevice-Guided Slow Breathing to Reduce Blood Pressure in Patients with Hypertension: Protocol for a Systematic Review and Meta-analysis
JMIR Res Protoc 2022;11(3):e33579
URL: https://www.researchprotocols.org/2022/3/e33579
doi:10.2196/33579
PMID:35254284

Protocol

Technology-Based Interventions to Promote the HIV Preexposure Prophylaxis (PrEP) Care Continuum: Protocol for a Systematic Review

Chenglin Hong[1,2], MSW, MPH; Laura S Abrams[1], PhD; Ian W Holloway[1,2], MSW, MPH, PhD

[1]Department of Social Welfare, Luskin School of Public Affairs, University of California, Los Angeles, Los Angeles, CA, United States

[2]Gay Sexuality and Social Policy Initiative, Luskin School of Public Affairs, University of California, Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Chenglin Hong, MSW, MPH
Department of Social Welfare
Luskin School of Public Affairs
University of California, Los Angeles
3250 Public Affairs Building
Los Angeles, CA, 90095
United States
Phone: 1 4087969472
Email: chenglinhong@ucla.edu

Abstract

Background: Preexposure prophylaxis (PrEP) is a promising biomedical intervention for HIV prevention. Researchers have proposed the PrEP care continuum to guide and evaluate PrEP implementation programs. Technology-based interventions (TBIs) have been widely used in HIV prevention and treatment programs, including for the promotion of the PrEP care continuum. The rapid development of new interventions using technology and electronic health methods emphasizes the need for a review of the effectiveness of these TBIs.

Objective: The aim of this systematic review is to summarize the effectiveness and acceptability of TBIs used to promote the HIV PrEP care continuum.

Methods: We will conduct a systematic literature search in PubMed, Embase, MEDLINE, PsycINFO, Web of Science, CINAHL, and the Cochrane Central Register of Controlled Trials following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Only intervention studies (ie, studies meeting the criteria of randomized controlled trials or quasi-experimental studies) evaluating the effectiveness of TBIs will be included. We will search the National Institutes of Health Research Portfolio Online Reporting Tools (NIH RePORT) for interventions involving PrEP. At least 2 reviewers will independently screen and select the studies, extract the data, and evaluate the quality of the studies, and discrepancies will be resolved by a senior author. We will provide a narrative synthesis of the included studies and present details about the study populations, interventions, and PrEP-related outcomes of significance.

Results: The protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO; CRD42021249562). As of August 2021, we have completed the initial search and identified 1213 records. Study screening and data extracting are in progress. We expect the results to be ready by summer 2022.

Conclusions: The findings of this review will summarize successful experiences and lessons learned from the existing literature and therefore inform the design and implementation of intervention studies for PrEP care promotion.

Trial Registration: PROSPERO CRD42021249562; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=249562

International Registered Report Identifier (IRRID): DERR1-10.2196/33045

(JMIR Res Protoc 2022;11(3):e33045) doi:10.2196/33045

KEYWORDS

PrEP; pre-exposure prophylaxis; HIV; technology-based intervention; eHealth; mHealth; systematic review; HIV; HIV prevention

Introduction

The efficacy and safety of preexposure prophylaxis (PrEP) for HIV prevention has been demonstrated through clinical trials among a range of at-risk populations, including men who have sex with men (MSM), people who inject drugs, and other high-risk populations [1-3]. The biomedical approach of daily oral antiretroviral medication is highly effective for preventing HIV when taken as prescribed [4,5]. To better facilitate the scale-up of PrEP and measure implementation progress, researchers proposed the PrEP care continuum (or PrEP care cascade, hereafter referred to as the PrEP continuum). Kelley et al [6] proposed a PrEP continuum framework to achieve protection for MSM that includes the following components: (1) at-risk MSM, (2) an awareness of and willingness to take PrEP, (3) access to health care, (4) a prescription for PrEP, and (5) adherence to effective PrEP. Nunn et al [7] expanded this framework as follows and included the additional step of retaining patients in PrEP care: (1) identifying individuals at highest risk for contracting HIV, (2) increasing HIV risk awareness among those individuals, (3) enhancing PrEP awareness, (4) facilitating PrEP access, (5) linking to PrEP care, (6) prescribing PrEP, (7) initiating PrEP, (8) adhering to PrEP, and (9) retaining individuals in PrEP care. Each step of the continuum represents a critical point for interventions, which can substantially increase the overall PrEP effectiveness at a population level. These frameworks guide researchers and practitioners to measure PrEP outcomes, and there are increasing numbers of published studies using the PrEP continuum to design and evaluate interventions to promote PrEP implementation worldwide [8-10].

In the past decade, electronic health and mobile health technologies, collectively referred to as technology-based interventions (TBIs), have seen exponential growth in the field of health care and health promotion. Examples of TBIs include disseminating health information on popular social media platforms, monitoring sleep duration and quality on wearable devices, sending push notifications to patients' devices to support medication adherence, and using web-based chatrooms and meeting platforms for remote counseling [11]. These TBIs have several key features and functions that provide them with advantages over traditional health promotion interventions and programs, including their ability to reach a broader range of communities and individuals across diverse contexts, provide consistent services without the limitation of physical distance, and potentially be cost-effective and widely scalable once developed and adapted. Previous evaluation studies and meta-analyses have demonstrated and verified the effectiveness of such TBIs for supporting health behavior changes and disease management in several cases, including cardiovascular disease [12], diabetes [13], medication adherence [14], mental health, physical health [15], and sexual health including HIV and sexually transmitted infection prevention and care [16].

Compared to traditional HIV prevention interventions, TBIs can facilitate large-scale information dissemination and can effectively deliver just-in-time services and feedback to at-risk populations including MSM and people who inject drugs [17-19]. A recently published systematic review of TBIs related to HIV prevention and care cataloged published studies and other funded projects between 2014 and 2018 and found that there was a growing trend of TBIs and funded studies targeting various points on the HIV prevention and care continua with a substantial emphasis on education and behavior changes; however, there were no TBIs aiming to promote PrEP adherence and persistence [20]. From a nonsystematic review of various technologies used in HIV testing research, Romero et al [21] also noted that TBIs have the ability to increase outreach and HIV self-testing. To date, there is only one published review of TBIs involving PrEP; this review summarizes several telehealth programs aiming to improve PrEP availability, uptake, and adherence [22]. However, this review had some limitations. First, it is not clear whether it followed a systematic review guideline, and the authors only focused on telehealth approaches. Therefore, the results may not apply to all TBIs involving PrEP. In addition, similar to Maloney's review [20], their analyses did not summarize the efficacy of each intervention and specifically evaluate which features of the interventions influenced each step of the PrEP continuum [22]. Finally, there was a lack of summary and evaluation on whether the TBIs were designed with guidance from a theoretical framework. In addition, little is known about the acceptability of and user experience associated with these TBIs and whether the interventions were scaled up and used in the real world after the completion of the study.

These considerations suggest the need for a comprehensive and rigorous review of the state of the science regarding TBIs and PrEP implementation. Therefore, the objectives of this systematic review are to (1) describe the use of TBIs in promoting steps along the PrEP continuum by population studied, (2) summarize the theoretical frameworks that have been used to inform and design the interventions, (3) summarize the evidence on and effectiveness of promoting steps along the PrEP continuum using the interventions, and (4) describe the acceptability of the TBIs among users, user experiences, and the sustainability of the intervention after study completion.

Methods

Protocol Registration and Design

This review was designed and reported in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [23,24]. A protocol outlining the review methods was registered in the International Prospective Register of Systematic Reviews (PROSPERO; CRD42021249562).

Search Strategy

A systematic literature search was conducted using the following databases: PubMed, Embase, MEDLINE, PsycINFO, Web of Science, CINAHL, and the Cochrane Central Register of Controlled Trials. The search strategies were developed in consultation with a health science librarian and in accordance with previous reviews on similar topics [20,25]. The search terms were created using a combination of Medical Subject Headings, keywords, and phrases, including "PrEP" or "pre-exposure prophylaxis" in combination with, but not limited to, any of the following: "eHealth", "mHealth", "smartphone",

"mobile phone", "mobile application", "app", "internet", "text messages", "social media", "dating app", "wearable device", "website", and "technology". A manual search will be conducted using the reference lists of included publications and using Google Scholar to find other relevant studies. In addition, we will search the National Institutes of Health Research Portfolio Online Reporting Tools (NIH RePORT) for interventions still underway and contact the principal investigators for a basic summary of the interventions. Multimedia Appendix 1 provides details of the search strategy and terms used.

Eligibility Criteria

We will include published records of intervention studies (eg, randomized trials and quasi-experimental studies) that compare outcomes for individuals who participated in the TBI to any control group. As PrEP was first approved by the US Food and Drug Administration for HIV prevention in 2012, we used this year for the start of the search. We will exclude articles published in languages other than English and those that did not have a PrEP-related outcome as one of the primary outcomes. Study protocols and reports of ongoing trials will also be excluded. We will only include studies with original data; therefore, nonempirical studies such as letters, perspectives, and editorials will not be included.

Study Population

There are no restrictions on the sociodemographic characteristics of the study participants in the intervention studies.

Interventions

To date, there is no consensus on the definition of TBIs and what interventions fall under this category. We first created a list of TBIs and their features based on the extant literature and iteratively refined this list based on a review of the included studies. Given the broad definition, we included studies in which an intervention used information technology for service delivery to facilitate one or multiple steps of the PrEP continuum. These interventions include, but were not limited to, mobile apps, social media, web-based products, SMS text messaging, web-based chatrooms and quizzes, push notifications, and interactive geolocation maps.

Outcomes

The included studies must report at least one PrEP-related outcome. Although the previous PrEP continuum consists of multiple steps, this review will primarily focus on PrEP awareness, willingness to take PrEP, PrEP uptake, and PrEP adherence [8,10]. Although these broad definitions were summarized from the existing literature, we will extract how each study defined and measured these outcomes. Generally, PrEP awareness will be defined as having heard of PrEP and having accurate knowledge of PrEP efficacy for HIV prevention. Willingness to use PrEP is the hypothetical willingness and perceived acceptability of using PrEP in the future. PrEP uptake is the initiation and actual use of PrEP among the study participants. Although alternative PrEP options, such as long-acting injectable PrEP, are under development, this review will only focus on oral PrEP (Truvada, Descovy, and generics [26]). PrEP adherence is the level of medication adherence sufficient for the prevention of HIV acquisition. We will consider both daily and "on-demand" dosing (taking PrEP only when individuals are at risk for getting HIV is known as on-demand PrEP, or the "2-1-1 schedule," which means taking 2 pills between 2 and 24 hours before sexual intercourse, 1 pill 24 hours after the first dose, and 1 pill 24 hours after the second dose [27]) if the articles clearly demonstrated how adherence was defined and measured in their studies. This could include self-reported days of missing the pills in the months prior to the study or lab results of tenofovir-diphosphate concentrations in dried blood spot samples [28]. All PrEP-related outcomes must be assessed and reported quantitatively (eg, using rates, percentages, or odds ratios).

Secondary outcomes include users' experiences, acceptability of the TBI, and satisfaction rates. User acceptability and satisfaction can be assessed using multiple methods, including exit interviews or standardized questionnaires or scales, including the Computer Usability Satisfaction Questionnaire [29] and the Health Information Technology Usability Evaluation Scale [30].

Study Selection and Screening

We will import all the search records to Covidence (Veritas Health Innovation Ltd), a software that automatically removes duplicate studies. The remaining studies will be screened under blinded conditions by at least 2 independent reviewers. Article titles and abstracts will first be screened for eligibility, and any studies that meet the eligibility criteria or those that the reviewers cannot decide to include or exclude will progress to the full-text review. Next, 2 independent reviewers (CH and another reviewer) will screen the full texts of the articles to determine whether they will be included in the review and data extraction. The results of each step will be compared, and any inconsistency or conflict will be resolved through discussion or arbitration from a senior author (IWH) until a consensus is reached.

Data Extraction and Analysis

A standardized data extraction form will be created to enter the data from the included articles. A total of 2 reviewers will extract the data independently (CH and another reviewer), and the results will be compared and discussed until a consensus is reached. For each selected article, we will record the general study characteristics and intervention-specific data. Study characteristics include study authors, year of publication, study location, participants' sociodemographic characteristics, study design, recruitment methods, and sample size. Intervention-specific data include the name of the study and a brief description of the intervention, study design, type of technology used for the intervention (eg, internet or social media), theoretical framework used to guide the intervention design (if applicable), comparator used, PrEP continuum domain that the study aimed to improve, and quantitative data on the comparisons between the intervention and the control group (eg, changes in percentages, rates, and odds ratios). Information on whether the intervention continued to be implemented after the study completion will be extracted from the included studies. The reviewer (CH) will contact the corresponding authors for clarification and further information where study details are missing or unclear. The studies will be categorized and grouped

according to the type of intervention and the PrEP continuum domain that was targeted.

Risk of Bias Assessment

The quality of the studies will be assessed using the Cochrane Risk of Bias Tool for randomized controlled trials [31]. The domains being assessed include random sequence generation, allocation concealment, blinding of participants, blinding of personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other sources of bias. All included studies will be scored as having a low, high, or unclear risk of bias.

Results

As of August 2021, we have completed the preliminary searches on the databases previously listed. A total of 1213 records were identified. Eligibility screening and data extraction are ongoing. It is anticipated that the final review will be completed and submitted for publication in summer 2022.

Discussion

The emergence of TBIs in public health research has created opportunities for health promotion and disease prevention. A growing number of intervention studies have investigated the effect of TBIs in supporting the PrEP continuum among various at-risk populations. This systematic review will provide an overview of these TBIs and describe the evidence supporting their use and their efficacy. The review will highlight which kind of interventions have been implemented and which PrEP continuum step(s) are primarily targeted. Furthermore, this review will be the first to summarize the theoretical frameworks used for informing TBI design. The findings of this review will summarize successful experiences and lessons learned from published literature, conference abstracts, NIH RePORT, and unpublished data. The findings will also inform the design and implementation of future studies on PrEP care promotion.

Acknowledgments

CH is supported by the University of California, Los Angeles (UCLA) Graduate Division Summer Mentored Research Fellowship. We would like to thank the health science librarian at UCLA for their work on the literature search strategy for this review.

Authors' Contributions

CH conceptualized, designed, and initiated this study, conducted the initial search and screening of studies, and drafted the manuscript. LSA and IWH reviewed the manuscript and provided critical feedback. All authors approved the final manuscript for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The search strategy used to find articles in PubMed.
[DOCX File , 14 KB - resprot_v11i3e33045_app1.docx]

References

1. 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]
2. Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, Partners PrEP Study Team. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N Engl J Med 2012 Aug 02;367(5):399-410 [FREE Full text] [doi: 10.1056/NEJMoa1108524] [Medline: 22784037]
3. 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]
4. Haberer J. Current concepts for PrEP adherence in the PrEP revolution: from clinical trials to routine practice. Curr Opin HIV AIDS 2016 Jan;11(1):10-17 [FREE Full text] [doi: 10.1097/COH.0000000000000220] [Medline: 26633638]
5. Amico KR, Stirratt MJ. Adherence to preexposure prophylaxis: current, emerging, and anticipated bases of evidence. Clin Infect Dis 2014 Jul;59 Suppl 1:S55-S60 [FREE Full text] [doi: 10.1093/cid/ciu266] [Medline: 24926036]
6. Kelley CF, Kahle E, Siegler A, Sanchez T, Del Rio C, Sullivan PS, et al. Applying a PrEP continuum of care for men who have sex with men in Atlanta, Georgia. Clin Infect Dis 2015 Nov 15;61(10):1590-1597 [FREE Full text] [doi: 10.1093/cid/civ664] [Medline: 26270691]

7. Nunn A, Brinkley-Rubinstein L, Oldenburg C, Mayer KH, Mimiaga M, Patel R, et al. Defining the HIV pre-exposure prophylaxis care continuum. AIDS 2017 Mar 13;31(5):731-734 [FREE Full text] [doi: 10.1097/QAD.0000000000001385] [Medline: 28060019]
8. Ezennia O, Geter A, Smith DK. The PrEP care continuum and Black men who have sex with men: a scoping review of published data on awareness, uptake, adherence, and retention in PrEP care. AIDS Behav 2019 Oct 28;23(10):2654-2673. [doi: 10.1007/s10461-019-02641-2] [Medline: 31463711]
9. Mistler CB, Copenhaver MM, Shrestha R. The Pre-exposure prophylaxis (PrEP) care cascade in people who inject drugs: a systematic review. AIDS Behav 2021 May 04;25(5):1490-1506. [doi: 10.1007/s10461-020-02988-x] [Medline: 32749627]
10. Glick JL, Russo R, Jivapong B, Rosman L, Pelaez D, Footer KHA, et al. The PrEP care continuum among cisgender women who sell sex and/or use drugs globally: a systematic review. AIDS Behav 2020 May 05;24(5):1312-1333 [FREE Full text] [doi: 10.1007/s10461-019-02733-z] [Medline: 31807973]
11. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of mobile apps to promote health and manage disease: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2021 Jan 11;9(1):e21563 [FREE Full text] [doi: 10.2196/21563] [Medline: 33427672]
12. Gandapur Y, Kianoush S, Kelli HM, Misra S, Urrea B, Blaha MJ, et al. The role of mHealth for improving medication adherence in patients with cardiovascular disease: a systematic review. Eur Heart J Qual Care Clin Outcomes 2016 Oct 01;2(4):237-244 [FREE Full text] [doi: 10.1093/ehjqcco/qcw018] [Medline: 29474713]
13. Bonoto BC, de Araújo VE, Godói IP, de Lemos LLP, Godman B, Bennie M, et al. Efficacy of mobile apps to support the care of patients with diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2017 Mar 01;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]
14. Linn AJ, Vervloet M, van Dijk L, Smit EG, Van Weert JC. Effects of eHealth interventions on medication adherence: a systematic review of the literature. J Med Internet Res 2011 Dec 05;13(4):e103 [FREE Full text] [doi: 10.2196/jmir.1738] [Medline: 22138112]
15. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
16. Muessig KE, Bien CH, Wei C, Lo EJ, Yang M, Tucker JD, et al. A mixed-methods study on the acceptability of using eHealth for HIV prevention and sexual health care among men who have sex with men in China. J Med Internet Res 2015 Apr 21;17(4):e100 [FREE Full text] [doi: 10.2196/jmir.3370] [Medline: 25900881]
17. Ovalle A, Goldstein O, Kachuee M, Wu ESC, Hong C, Holloway IW, et al. Leveraging social media activity and machine learning for HIV and substance abuse risk assessment: development and validation study. J Med Internet Res 2021 Apr 26;23(4):e22042 [FREE Full text] [doi: 10.2196/22042] [Medline: 33900200]
18. Platteau T, Herrijgers C, de Wit J. Digital chemsex support and care: The potential of just-in-time adaptive interventions. Int J Drug Policy 2020 Nov;85:102927 [FREE Full text] [doi: 10.1016/j.drugpo.2020.102927] [Medline: 32932125]
19. Meiksin R, Melendez-Torres GJ, Falconer J, Witzel TC, Weatherburn P, Bonell C. eHealth interventions to address sexual health, substance use, and mental health among men who have sex with men: systematic review and synthesis of process evaluations. J Med Internet Res 2021 Apr 23;23(4):e22477 [FREE Full text] [doi: 10.2196/22477] [Medline: 33890855]
20. Maloney KM, Bratcher A, Wilkerson R, Sullivan PS. Electronic and other new media technology interventions for HIV care and prevention: a systematic review. J Int AIDS Soc 2020 Jan;23(1):e25439 [FREE Full text] [doi: 10.1002/jia2.25439] [Medline: 31909896]
21. Romero RA, Klausner JD, Marsch LA, Young SD. Technology-delivered intervention strategies to bolster HIV testing. Curr HIV/AIDS Rep 2021 Aug 09;18(4):391-405 [FREE Full text] [doi: 10.1007/s11904-021-00565-y] [Medline: 34109549]
22. Touger R, Wood BR. A Review of telehealth innovations for HIV pre-exposure prophylaxis (PrEP). Curr HIV/AIDS Rep 2019 Feb 30;16(1):113-119. [doi: 10.1007/s11904-019-00430-z] [Medline: 30701404]
23. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
24. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160 [FREE Full text] [doi: 10.1136/bmj.n160] [Medline: 33781993]
25. Muessig KE, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman LB. A systematic review of recent smartphone, Internet and Web 2.0 interventions to address the HIV continuum of care. Curr HIV/AIDS Rep 2015 Mar 28;12(1):173-190 [FREE Full text] [doi: 10.1007/s11904-014-0239-3] [Medline: 25626718]
26. About PrEP. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/hiv/basics/prep/about-prep.html [accessed 2021-08-03]
27. On-demand PrEP. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/hiv/basics/prep/on-demand-prep.html [accessed 2021-08-03]
28. Anderson P, Glidden D, Liu A, Buchbinder S, Lama JR, Guanira JV, iPrEx Study Team. 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]

29. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int J Hum Comput Interact 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]
30. Yen P, Sousa KH, Bakken S. Examining construct and predictive validity of the Health-IT Usability Evaluation Scale: confirmatory factor analysis and structural equation modeling results. J Am Med Inform Assoc 2014 Oct 01;21(e2):e241-e248 [FREE Full text] [doi: 10.1136/amiajnl-2013-001811] [Medline: 24567081]
31. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343(oct18 2):d5928-d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]

Abbreviations

MSM: men who have sex with men
NIH RePORT: National Institutes of Health Research Portfolio Online Reporting Tools
PrEP: preexposure prophylaxis
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
TBI: technology-based interventions
UCLA: University of California, Los Angeles

Edited by G Eysenbach; submitted 22.08.21; peer-reviewed by B Castonguay, J Mitchell; comments to author 20.10.21; revised version received 07.12.21; accepted 08.12.21; published 08.03.22.

Please cite as:
Hong C, Abrams LS, Holloway IW
Technology-Based Interventions to Promote the HIV Preexposure Prophylaxis (PrEP) Care Continuum: Protocol for a Systematic Review
JMIR Res Protoc 2022;11(3):e33045
URL: https://www.researchprotocols.org/2022/3/e33045
doi:10.2196/33045
PMID:35258470

Protocol

Digital Health Intervention Design and Deployment for Engaging Demographic Groups Likely to Be Affected by the Digital Divide: Protocol for a Systematic Scoping Review

Catherine L Jenkins[1], BA; Sumayyah Imran[2]; Aamina Mahmood[2]; Katherine Bradbury[3], BSc, MSc, PhD; Elizabeth Murray[2], BA, MBBS, MSc, PhD; Fiona Stevenson[2], MA, PhD; Fiona L Hamilton[2], MBBS, MSc, MRCGP, PhD

[1]North East London Foundation Trust, Essex, United Kingdom

[2]eHealth Unit, Research Department of Primary Care & Population Health, University College London, London, United Kingdom

[3]School of Psychology, University of Southampton, Southampton, United Kingdom

Corresponding Author:
Fiona L Hamilton, MBBS, MSc, MRCGP, PhD
eHealth Unit
Research Department of Primary Care & Population Health
University College London
Upper Third Floor, UCL Medical School (Royal Free Campus)
Rowland Hill Street
London, NW3 2PF
United Kingdom
Phone: 44 2037940272
Email: flodocs@hotmail.com

Abstract

Background: Digital health interventions refer to interventions designed to support health-related knowledge transfer and are delivered via digital technologies, such as mobile apps. Digital health interventions are a double-edged sword: they have the potential to reduce health inequalities, for example, by making treatments available remotely to rural populations underserved by health care facilities or by helping to overcome language barriers via in-app translation services; however, if not designed and deployed with care, digital health interventions also have the potential to increase health inequalities and exacerbate the effects of the digital divide.

Objective: The aim of this study is to review ways to mitigate the digital divide through digital health intervention design, deployment, and engagement mechanisms sensitive to the needs of digitally excluded populations.

Methods: This protocol outlines the procedure for a systematic scoping review that follows the methodology recommended by the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidance. The following databases will be searched for primary research studies published in English from October 1, 2011, to October 1, 2021: Cochrane Library, Epistemonikos, NICE Evidence, PROSPERO, PubMed (with MEDLINE and Europe PMC), and Trip. In addition, the following sources of gray literature will be searched: Conference Proceedings Citation Index, Health Management Information Consortium, International HTA Database, OpenGrey, The Grey Literature Report, Google Scholar Basic Search UK, MedNar Deep Web Search Engine, and Carrot2. We will select publications that meet the following inclusion criteria: primary research papers that evaluated digital health interventions that describe features of digital health intervention design and deployment that enable or hinder access to and engagement with digital health interventions by adults from demographic groups likely to be affected by the digital divide (eg, older age, minority ethnic groups, lower income, and lower education level). A random selection of 25 publications identified from the search will be double screened by four reviewers. If there is >75% agreement for included/excluded publications, the team will continue to screen all the identified publications. For all included publications, study characteristics will be extracted by one author and checked for agreement by a second author, with any disagreements resolved by consensus among the study team. Consultation digital health intervention design and deployment, and digital health intervention users will also be conducted in parallel.

Results: The review is underway and is anticipated to be completed by September 2022.

Conclusions: The results will have implications for researchers and policy makers using digital health interventions for health improvement peripandemic and post pandemic, and will inform best practices in the design and delivery of digital health interventions.

International Registered Report Identifier (IRRID): PRR1-10.2196/32538

(JMIR Res Protoc 2022;11(3):e32538) doi:10.2196/32538

KEYWORDS

digital divide; digital health interventions; DHIs; eHealth; digital health literacy; health inequalities; health inequities; mHealth; mobile health

Introduction

Background

The COVID-19 pandemic has accelerated widespread adoption of digital health interventions internationally. This rapid shift to digital delivery has laid bare the impact of pre-existing and emerging systemic health inequalities on communities most in need of accessible health care, specifically people from ethnic minority backgrounds, people from lower socioeconomic backgrounds, older people (aged ≥65), and people living with disabilities [1,2]. Belonging to one or more of these groups (intersectionality) is a risk factor for experiencing more severe illness and mortality [3]. These same groups are likely to make up a large proportion of people vulnerable to health inequalities regardless of initial illness severity [4].

Health inequalities have been defined as "the systematic, avoidable and unfair differences in health outcomes that can be observed between populations, between social groups within the same population or as a gradient across a population ranked by social position" [5], but the term is also used for differences in access to health care, quality of care received, wider determinants of health such as housing and education, and opportunities to lead healthy lives, including differences in risky behaviors such as smoking [6]. The related term "health inequity" implies a normative judgement about the fairness or otherwise of these differences, "expressing a moral commitment to social justice," as described by Kawachi et al [7]. For this review, we use the term health inequalities.

The digital divide—the gap between populations able to benefit from access to and use of health information and services online and populations unable to take up such opportunities—is a clear example of health inequality and exacerbates the inverse care law [8] such that digitally delivered health care runs the risk of excluding the people who could most benefit. This review seeks to address issues of social justice in the digital delivery of health care by providing a comprehensive overview of the literature on strategies to reduce the digital divide through the design and deployment aspects of digital health interventions.

The Digital Divide and Digital Health Literacy

Health literacy and digital literacy are both key determinants of health [9]. The term digital health literacy (also referred to as eHealth literacy) brings both literacies together to describe the degree to which individuals can access, understand, and apply digitally delivered health information and services to make informed decisions about their health. Importantly, digital health literacy extends beyond personal responsibility to encompass the responsibility of digital health systems and services to support and dynamically respond to the digital health literacy skills, and the confidence and motivation to develop such skills, in the populations they serve [10].

Digital health literacy competence fluctuates depending on context, but a clear link has been demonstrated between low digital health literacy and poor health outcomes [11]. The accelerated shift to digital health care to comply with social distancing measures arising from COVID-19 heightens the risk of poor health outcomes in digitally excluded groups and further complicates an already complex interplay between intersectionality and digital health literacy as a distributed, or *outsourced*, concept. For example, many older people may have the financial capital to afford digitally enabled devices and data plans but not the skills, motivation, or confidence to engage directly with such technology; instead, they may be reliant on their networks—family members, friends—to navigate digital health care on their behalf [12,13].

Digital health intervention design and deployment therefore need to be proportionate and prioritize those groups most affected by low digital health literacy on both individual and systemic levels via a two-pronged approach: first, to develop patients' skills in accessing, understanding, and using digital health information and services, and second, to develop the digital health literacy responsiveness of the systems and health care professionals (HCPs) supporting digital health intervention deployment in practice. This review will have wider implications for research into digital health interventions and will address:

- The digital divide: How should digital health interventions be designed and deployed to mitigate the digital divide (eg, patient access to WiFi-enabled digital devices or subsidized data plans, the inclusiveness of recruitment methods for studies evaluating digital health interventions, strategies to increase uptake and use of a digital health intervention app by people from demographic groups likely to be affected by the digital divide, or the content of onboarding scripts used by HCPs when introducing patients to a digitally delivered health service)?
- Digital health literacy: How can app design and digital skills training be optimized to reduce digital health inequalities arising from low digital health literacy (eg, integration of user testing into the design process, readability of app content, gamification/social features to engage users, or support for HCPs to train as "digital health champions" as part of improving their own skills)?

Rationale for Conducting a Systematic Scoping Review

The decision to conduct a systematic scoping review rather than a systematic review is informed by the purpose of scoping reviews, which is to provide a comprehensive picture of the available evidence to guide further research. This purpose is of relevance to the study of digital health interventions because digital health intervention design and deployment is frequently captured in gray literature (eg, charity reports).

This review will be conducted using methods outlined by the Joanna Briggs Institute (JBI) Reviewers' Manual [14] and will conform to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [15]. The protocol for this review is based on the JBI framework by Peters et al [16] with further process updates from Levac et al [17] and is structured as follows: identification of the review aims and search questions, selection of evidence sources, charting of data extracted from the evidence sources, reporting of results, and consultation with stakeholders (including HCPs, digital health experts, and digital health intervention users). Results from this review will support decision-making when designing and deploying digital health interventions.

Review Question

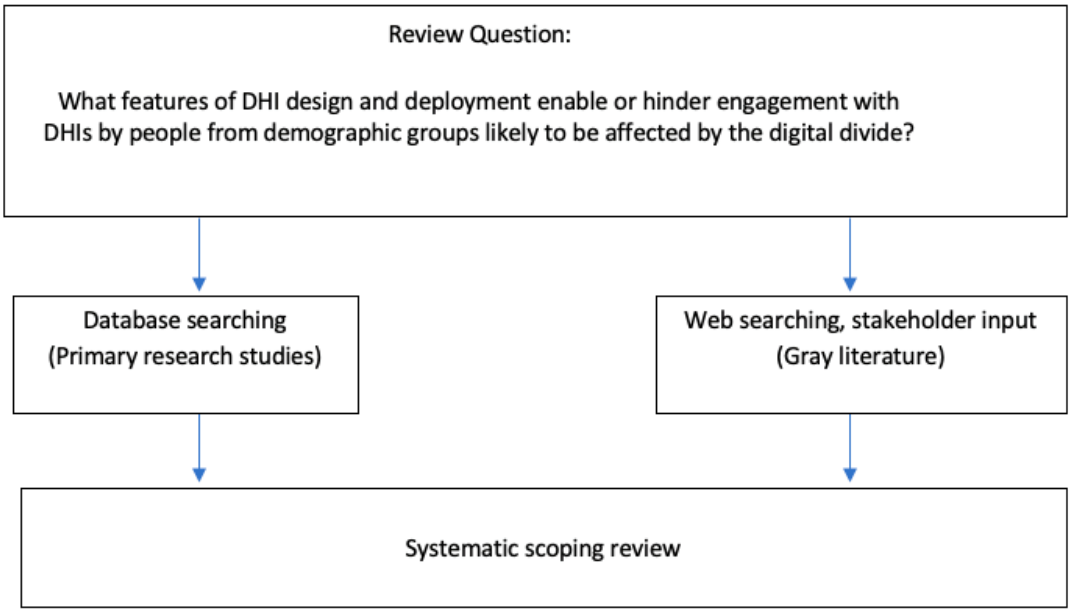
The purpose of this systematic scoping review is to identify research that reports on the design and deployment of digital health interventions to reduce the digital divide and increase digital health literacy at macro- (national policy), meso- (national program), and micro- (localized or individual) levels [18]. It aims to do this by discovering peer-reviewed qualitative, quantitative, and mixed methods research, and gray literature relevant to mapping the contextual and process factors that enable or hinder engagement with digital health services by people vulnerable to the digital divide and individual or systemic low digital health literacy.

This review will seek to answer the review question, "What features of digital health intervention design and deployment enable or hinder engagement with digital health interventions by people from demographic groups likely to be affected by the digital divide?"

Health interventions are increasingly delivered through digital platforms, and it is important that they do not exacerbate or create health inequalities. Our hypothesis is that existing knowledge in the literature can inform our objective to find ways to bridge the digital divide and improve digital health literacy in underserved groups, for example, through digital health intervention design, development, and deployment.

Figure 1 outlines how this review will be conducted.

Figure 1. Systematic scoping review flow diagram. DHI: digital health intervention.



Methods

We will follow the methods recommended by the PRISMA-ScR framework [15].

Protocol Registration

At the time of writing, scoping reviews are ineligible for registration in PROSPERO. For transparency and to enable the systematic scoping review method outlined here to be repurposed for onward research, a summary of the protocol and any supplementary material will be registered with the Open Science Framework [19] and assigned a Digital Object Identifier (DOI) for long-term retrieval.

Eligibility Criteria

We will include primary research studies reporting on the design or deployment of digital health interventions that meet the following inclusion criteria.

Population

Adults from different demographic groups (defined based on gender, age, ethnicity, income, and education) in high-income and low-income countries.

Intervention

Digital health interventions are defined as any service intended to improve physical or mental health, or to promote health improvement through, for example, lifestyle change delivered digitally (formally or informally), such as via smartphone apps, social media, email, SMS text message, using wearable technologies, video games (eg, for motor or cognitive training), websites, or telehealth (eg, remote consultations) but excluding telemedicine if this consists solely of remote monitoring without any input from the patient [15].

Comparator

There may or may not be a comparator, depending on the study design.

Outcome

The primary outcome of interest is the capacity to mitigate the digital divide and increase digital health literacy either through the digital health intervention itself or its deployment. Health impact is not considered an outcome of interest for this review.

- "Mitigate the digital divide" refers to removing barriers to digital inclusion, for example, an intervention that provides free smartphones or tablets, or low-cost cellular data or WiFi for people with low incomes [10].
- "Increase digital health literacy" refers to digital skills development [9], either through the design of the digital health intervention or through programs to deploy the digital health intervention, resulting in higher levels or use or confidence in using the digital health interventions, measured either by interview or self-report (eg, the eHealth Literacy Scale questionnaire [20] or use data). Changes in these outcome measures suggest that both users and the health professionals administering the digital health intervention are confident in the use of digital resources for safe and sustainable self-management and to promote safe and sustainable self-management.

For the purposes of this review, gray literature covers conference abstracts and proceedings, white papers, and stakeholder reports (eg, by charities with an interest in digital literacy or who promote the health and well-being of people from different demographic groups) that enhance the contextual understanding of the field.

There will be no limitation by geography because lessons from low-income countries could inform strategies to reduce digital health inequities in high-income countries. We will report and interpret data within the context of the country where the study was based.

Limitations will be placed on user population and on study date. The user population will prioritize the marginalized demographics previously outlined; the study date will be limited to post-2011 (for preprint or publication) in recognition of the rapid change of pace in digital health intervention development [21]. No limitations will be placed on the study design or publication status. However, we will limit included studies to those published in English.

Decisions on whether data should be included or excluded from further analysis will be guided by criteria that, in line with a systematic application of scoping review method [16], will be iteratively refined based on increasing familiarity with the literature under review.

The use of the term digital health intervention is a possible challenge as it covers a range of digital health technologies [22] and may contribute to an imbalance in this review that favors comprehensiveness over precision. This will be addressed by working closely with information specialists to peer review the preliminary search strategy in line with the PRESS (Peer Review of Electronic Search Strategies) procedure [23], ensuring a feasible and focused approach that is also flexible enough to be iteratively reworked in response to the results retrieved.

Evidence Sources

Reviews

A preliminary search for registered, preprint, or published systematic, scoping, and other review types will be conducted to pilot the search strategy, covering the following databases (via the Ovid interface, where applicable):

- Cochrane Library
- Epistemonikos
- NICE Evidence
- PROSPERO
- PubMed (with MEDLINE and Europe PMC)
- Trip

The preliminary search will provide keywords that can be incorporated into searches across evidence tiers below that of the gold standard represented by reviews. It will also serve to identify Medical Subject Headings (MeSH) and PubMed IDs, which can be inserted into the Yale MeSH Analyzer [24] to extract further MeSH and to identify keywords to capture preprints not yet indexed with MeSH. A cap of 100 review articles will be applied in cases where unfeasible amounts of hits are retrieved.

Model papers for citation "*pearl-growing*"—papers already identified as highly relevant to this review—will be mined for their keywords and applied as test cases for honing the search strategy. Vernacular search terms will be identified from hand searching, including reference list scanning, forward/backward citation snowballing, and table of contents scanning of relevant journals online, such as the *Journal of Medical Internet Research* and its sister journals.

Further sources for other tiers of evidence include:

- Databases of peer-reviewed primary research (qualitative, quantitative, mixed methods studies):
  - CINAHL
  - EMBASE
  - PsycINFO
  - OTseeker
  - PubMed

- Gray literature
  - Conference Proceedings Citation Index
  - Health Management Information Consortium
  - International HTA Database
  - OpenGrey
  - GreyNet
  - The Grey Literature Report

The Grey Literature Report [25], available up to 2016, is a key source of gray literature that, unusually, is indexed using MeSH. It may be helpful in providing alternative MeSH to feed back into the search strand for peer-reviewed literature.

Alongside database searching, web searching will be conducted systematically, to the extent that this is feasible given that web searching is vulnerable to changing web content and algorithms [26]. Search terms will be used consistently between database and web searching and screen captures of content, and where available, DOIs will be saved for transparency. Web searching will also support the discovery of human-computer interaction (HCI) literature, which is not reliably indexed in MEDLINE. Searching the HCI literature will be helpful for the provision of further perspectives on design challenges (eg, designing for people with low vision or motor impairments).

Web Searching

The following databases will be used for the web searching: Google Scholar Basic Search UK, with use of the "Cited by" function (see Figure 2), for the first 100 results retrieved; MedNar Deep Web Search Engine, for the first 100 results retrieved; and Carrot2.

Targeted consultation with experts in digital health intervention design and deployment and digital health intervention users will also be conducted in parallel. This is part of the added value of scoping reviews. Ethics approval is not required for such consultation.

Figure 2. "Cited by" feature for forward citation snowballing in Google Scholar.

Search Strategy

The following search strategy was used for PubMed (November 2021): ((“digital health” OR “digital health intervention*” OR “DHI” OR “digital rehabilitation”) OR “eHealth” OR “mHealth” AND (app OR intervention OR technology) AND (barrier* OR design OR disparit* OR divi* OR engage* OR exclu* OR inequ*)) AND literacy.

This search strategy will be modified by incorporating MeSH and keywords used to index relevant papers and will inform the search strings for parallel searching in Google Scholar Advanced Search UK and MedNar.

Screening of identified literature will follow the recommendations of the JBI [16]. Of the studies and gray literature initially identified from the searches, 25 publications will be selected at random by dividing the total by four (the number of screeners), then selecting every nth paper (where n=the total divided by four) to assign for double-screening of titles, abstracts, or full-text by four reviewers; if there is greater than 75% agreement for included/excluded publications, then the team will continue to screen all of the identified publications. If agreement is 75% or less, we will hold a further training session to resolve differences in interpretation of the criteria and, if necessary, tighten the wording. Further rounds of double-screening and calculation of interrater agreement will continue until we meet concordance. For included publications, study characteristics will be extracted by one author and checked for agreement by a second author, and any uncertainty over data retrieved will be discussed and resolved via consensus with all contributing authors. Retrieval of a model paper [27] served as the test case for the search strategy.

PRISMA-ScR

The PRISMA-ScR framework will be used to transparently record the selection, deduplication, and screening decision process, supplemented by updated recommendations from Tricco et al [28]. Gray literature may not be searchable by abstract (due to a lack of abstract fields in the gray literature discovery interfaces), so the PRISMA-ScR flow diagram will be adapted to accommodate this. Critical appraisal, for example, of the underlying evidence informing the studies retrieved, is not generally operationalized in scoping reviews [16]. Instead, the focus of this review type is on providing as complete a picture as possible of what is currently known about the topic of interest.

Data Extraction and Charting

Data will be extracted from full-text sources relevant to the review question using the JBI data extraction guidance for scoping reviews [14]. If required, missing data will be requested from authors. Google Sheets [29] will be used to record the initial data extraction of all included sources. In accordance

with the iterative approach taken by scoping reviews, the following headers will be trialed and refined as needed [30]:

- Reference
- Region and setting of digital health intervention design/deployment
- Study type (qualitative, quantitative, mixed methods, review)
- Funding source
- Stakeholder involvement
- Digital health intervention type
- Digital health intervention purpose
- User population
- Summary of results relevant to this review question

Collating, Summarizing, and Reporting Results

Unlike a systematic review—where evidence sources that do not meet the quality criteria applied are excluded—scoping reviews seek to present an overview of all material reviewed (including gray literature) [24]. We will summarize the results visually [30] to assist validation of the results by stakeholders (described in the following section) and later dissemination.

Consultation Exercise

In parallel with the aforementioned stages, consultation with stakeholders (eg, representatives from charities, experts in researching digital health interventions, and patients or members of the public) will be conducted to validate emerging results. Stakeholders will be identified both a priori (known to the authors) and iteratively as this review progresses (through their representation in relevant sources).

Results

This systematic scoping review is in progress. The final draft of this review will be submitted by September 2022. We expect to report a narrative summary of the findings of included peer-reviewed primary studies and gray literature that describe differences in the ability of people from different demographic groups or people with lower health literacy or digital health literacy to access and use digital health interventions, and any change in this following specific measures taken by the researchers to make digital health interventions more accessible and usable by people from these groups; and the results of qualitative studies that discuss barriers and facilitators to access and use digital health interventions by people from different demographic groups.

Discussion

This review will present a summary of evidence regarding strategies for optimizing the design and deployment of digital health interventions to mitigate the effects of the digital divide and low digital health literacy on populations disproportionately affected by health inequalities and the digital divide. Its results will inform the design and deployment of digital health interventions at a time when they are more used than ever before and will have wider implications for researchers and policy makers using digital health interventions for health improvement. For example, HCI strategies for developing content for websites and apps, such as use of personas with a range of demographic characteristics (including those of digitally excluded populations) [31,32] and use of video and audio to help people of low health literacy use digital health interventions [33].

In conclusion, this review is among the first to examine the design and deployment of digital health interventions specifically in the context of the challenges presented by the digital divide and low digital health literacy at individual and systemic levels. The results of this review will support researchers, digital health intervention developers, and HCPs in identifying what works to optimize digital health intervention design and deployment, with the aim of promoting social justice.

Acknowledgments

We would like to acknowledge Sarah Rudd, Clinical Librarian, Library and Knowledge Service, North Bristol NHS Trust (National Health Service); NELFT Library & Knowledge Service (North East London Foundation Trust); the London Searching and Training Forum; Dr William Ricketts, Consultant Respiratory Physician, Barts Health NHS Trust; Dr Paul Pfeffer, Consultant Respiratory Physician, Barts Health NHS Trust and Queen Mary University of London; Deb Smith (PPIE [Patient and Public Involvement and Engagement]); Selena Cox (PPIE); and the Living With Covid Recovery study team.

This study is funded by the National Institute for Health Research (NIHR) Crossprogramme (HS&DR [Health & Social Care Delivery Research]) COVID-19 (project reference NIHR132243). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

This report is independent research supported by the National Institute for Health Research ARC North Thames.

Authors' Contributions

FLH, KB, and EM conceived the initial idea for the review. CLJ developed the protocol, which was then revised and approved by all authors. CLJ, SI, AM, and FLH will conduct the search strategy, screening, extraction, and analysis. CLJ will write up the systematic scoping review paper with support and contribution from all authors.

Conflicts of Interest

None declared.

References

1. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health 2020 Nov;74(11):964-968 [FREE Full text] [doi: 10.1136/jech-2020-214401] [Medline: 32535550]
2. Mueller AL, McNamara MS, Sinclair DA. Why does COVID-19 disproportionately affect older people? Aging (Albany NY) 2020 May 29;12(10):9959-9981 [FREE Full text] [doi: 10.18632/aging.103344] [Medline: 32470948]
3. Hankivsky O, Kapilashrami A. Intersectionality offers a radical rethinking of covid-19. BMJ Blogs. 2020 May 15. URL: https://blogs.bmj.com/bmj/2020/05/15/intersectionality-offers-a-radical-rethinking-of-covid-19/ [accessed 2021-11-22]
4. Taylor RR, Trivedi B, Patel N, Singh R, Ricketts WM, Elliott K, et al. Post-COVID symptoms reported at asynchronous virtual review and stratified follow-up after COVID-19 pneumonia. Clin Med (Lond) 2021 Jul;21(4):e384-e391 [FREE Full text] [doi: 10.7861/clinmed.2021-0037] [Medline: 34103378]
5. McCartney G, Popham F, McMaster R, Cumbers A. Defining health and health inequalities. Public Health 2019 Jul;172:22-30 [FREE Full text] [doi: 10.1016/j.puhe.2019.03.023] [Medline: 31154234]
6. Williams E, Buck D, Babalola G. What are health inequalities? The King's Fund. 2020 Feb. URL: https://www.kingsfund.org.uk/publications/what-are-health-inequalities [accessed 2021-11-24]
7. Kawachi I, Subramanian SV, Almeida-Filho N. A glossary for health inequalities. J Epidemiol Community Health 2002 Sep;56(9):647-652 [FREE Full text] [doi: 10.1136/jech.56.9.647] [Medline: 12177079]
8. Hart JT. The inverse care law. Lancet 1971 Feb 27;1(7696):405-412 [FREE Full text] [doi: 10.1016/s0140-6736(71)92410-x] [Medline: 4100731]
9. Nutbeam D, Lloyd JE. Understanding and responding to health literacy as a social determinant of health. Annu Rev Public Health 2021 Apr 01;42:159-173. [doi: 10.1146/annurev-publhealth-090419-102529] [Medline: 33035427]
10. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
11. Kim H, Xie B. Health literacy in the eHealth era: a systematic review of the literature. Patient Educ Couns 2017 Jun;100(6):1073-1082. [doi: 10.1016/j.pec.2017.01.015] [Medline: 28174067]
12. Ageing: science, technology and healthy living. United Kingdom Parliament. URL: https://publications.parliament.uk/pa/ld5801/ldselect/ldsctech/183/18308.htm#_idTextAnchor124 [accessed 2022-11-24]
13. The digital age: new approaches to supporting people in later life get online. Centre for Ageing Better. 2018. URL: https://ageing-better.org.uk/publications/digital-age [accessed 2021-11-24]
14. Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis. Australia: JBI; 2020.
15. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, PRISMA-S Group. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev 2021 Jan 26;10(1):39 [FREE Full text] [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
16. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Chapter 11: scoping reviews. JBI Global Wiki. 2020. URL: https://jbi-global-wiki.refined.site/space/MANUAL/3283910770/Chapter+11%3A+Scoping+reviews [accessed 2022-11-24]
17. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
18. Caldwell SE, Mays N. Health Res Policy Syst 2012 Oct 15;10:32 [FREE Full text] [doi: 10.1186/1478-4505-10-32] [Medline: 23067208]
19. Foster ED, Deardorff A. Open Science Framework (OSF). J Med Libr Assoc 2017 Apr 04;105(2):1. [doi: 10.5195/jmla.2017.88]
20. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
21. Patrick K, Hekler EB, Estrin D, Mohr DC, Riper H, Crane D, et al. The pace of technologic change: implications for digital health behavior intervention research. Am J Prev Med 2016 Nov;51(5):816-824. [doi: 10.1016/j.amepre.2016.05.001] [Medline: 27745681]
22. Soobiah C, Cooper M, Kishimoto V, Bhatia RS, Scott T, Maloney S, et al. Identifying optimal frameworks to implement or evaluate digital health interventions: a scoping review protocol. BMJ Open 2020 Aug 13;10(8):e037643 [FREE Full text] [doi: 10.1136/bmjopen-2020-037643] [Medline: 32792444]
23. 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 Jul;75:40-46 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
24. Grossetta Nardini HK, Wang L. Yale MeSH Analyzer. URL: https://mesh.med.yale.edu/ [accessed 2021-11-22]
25. Grey Literature Report. 2021. URL: http://greylit.org/ [accessed 2021-11-22]
26. Bethel A, Briscoe S, Rogers M. Searching for grey literature. University of Exeter. 2019 Mar 20. URL: https://medicine.exeter.ac.uk/media/universityofexeter/medicalschool/research/pentag/documents/Grey_Lit.pdf [accessed 2021-11-25]

27. Rowsell A, Muller I, Murray E, Little P, Byrne CD, Ganahl K, et al. Views of people with high and low levels of health literacy about a digital intervention to promote physical activity for diabetes: a qualitative study in five countries. J Med Internet Res 2015 Oct 12;17(10):e230 [FREE Full text] [doi: 10.2196/jmir.4999] [Medline: 26459743]
28. 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 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
29. Google Docs. Facebook. URL: https://www.facebook.com/GoogleDocs/ [accessed 2021-11-22]
30. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
31. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
32. Holden RJ, Kulanthaivel A, Purkayastha S, Goggins KM, Kripalani S. Know thy eHealth user: development of biopsychosocial personas from a study of older adults with heart failure. Int J Med Inform 2017 Dec;108:158-167 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.006] [Medline: 29132622]
33. Zelezny-Green R, Vosloo S, Conole G. Digital inclusion for low-skilled and low-literate people: a landscape review. UNESCO Digital Library. URL: https://unesdoc.unesco.org/ark:/48223/pf0000261791 [accessed 2022-11-25]

Abbreviations

DOI: Digital Object Identifier
HCI: human-computer interaction
HCP: health care professional
JBI: Joanna Briggs Institute
MeSH: Medical Subject Headings
NIHR: National Institute for Health Research
PRESS: Peer Review of Electronic Search Strategies
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by G Eysenbach; submitted 31.07.21; peer-reviewed by T Ntalindwa, C Cheng; comments to author 14.09.21; revised version received 25.10.21; accepted 27.11.21; published 18.03.22.

Please cite as:
Jenkins CL, Imran S, Mahmood A, Bradbury K, Murray E, Stevenson F, Hamilton FL
Digital Health Intervention Design and Deployment for Engaging Demographic Groups Likely to Be Affected by the Digital Divide: Protocol for a Systematic Scoping Review
JMIR Res Protoc 2022;11(3):e32538
URL: https://www.researchprotocols.org/2022/3/e32538
doi:10.2196/32538
PMID:35302946

Protocol

Nonalcoholic Fatty Liver Disease in Children and Adolescents Taking Atypical Antipsychotic Medications: Protocol for a Systematic Review and Meta-analysis

Reem Hatem[1]; Faisal A Nawaz[1]; Ghadah A Al-Sharif[1]; Mohammad Almoosa[1]; Wid Kattan[2], MBBS; Christos Tzivinikos[3], MD; E Lila Amirali[4], IMHL, MD; Ammar Albanna[1,3], MD

[1]College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates

[2]Division of Psychiatry, Department of Medicine, College of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

[3]Al Jalila Children's Specialty Hospital, Dubai, United Arab Emirates

[4]Department of Psychiatry and Addiction, Université de Montréal, Montréal, QC, Canada

Corresponding Author:
Ammar Albanna, MD
Al Jalila Children's Specialty Hospital
Oud Metha Road, Al Jaddaf
Dubai, 007662
United Arab Emirates
Phone: 971 502030000
Email: aalbanna@ajch.ae

Abstract

Background: Atypical antipsychotics (AAP) are commonly prescribed to children and adolescents and are associated with important adverse effects including weight gain and metabolic syndrome. Nonalcoholic fatty liver disease (NAFLD) is not only the most common pediatric liver disease but is also associated with serious complications including liver cirrhosis.

Objective: Given that NAFLD and AAP are associated with metabolic syndrome, we aim to comprehensively examine the association between AAP and NAFLD in children and adolescents.

Methods: We will conduct a systematic review of studies exploring NAFLD in subjects younger than 18 years on AAP published in English between 1950 and 2020 following the PRISMA (Preferred Reporting items for Systematic Reviews and Meta-Analysis) guidelines.

Results: A PRISMA flowchart will be used present the study results after comprehensively reviewing studies on NAFLD in children and adolescents taking AAP. The first and second systematic searches will be conducted during December 2021. The results are expected to be published in June 2022.

Conclusions: This research project will serve as a foundation for future studies and assist in devising interventions and reforming clinical guidelines for using AAP to ensure improved patient safety.

International Registered Report Identifier (IRRID): PRR1-10.2196/20168

(JMIR Res Protoc 2022;11(3):e20168) doi:10.2196/20168

KEYWORDS

nonalcoholic fatty liver disease; psychopharmacology; antipsychotics; children; adolescents; overprescribing; pharmaceuticals; antipsychotic medications; medication; pediatric psychopharmacology; pharmacology; child and adolescent psychiatry

Introduction

There is an increasing trend in prescribing atypical antipsychotics (AAP) to children and adolescents [1], and this merits a better understanding of their long-term adverse effects. Metabolic complications such as weight gain, hyperlipidemia, and insulin resistance are known side effects of AAP [2-4], and studies have suggested that youngsters are at an increased risk of these side effects compared to older adults [5]. With the decreased incidence of typical psychotic use in children and adolescents it is reasonable to speculate that children on AAP may be at a greater risk of disorders associated with metabolic syndrome, including nonalcoholic fatty liver disease (NAFLD) [6-9].

NAFLD in children and adolescents younger than 18 years is the most common cause of pediatric liver disease, with a prevalence of approximately 9.6% [10]. The prevalence significantly increases to approximately 77% among obese children [11], suggesting its association with metabolic syndrome [12]. NAFLD comprises a spectrum of liver diseases that range from the accumulation of fat in hepatocytes (steatosis) and inflammatory liver changes including nonalcoholic steatohepatitis (NASH) to more serious complications such as liver cirrhosis [13,14]. The "two-hit" theory attempts to explain the pathogenesis of NAFLD [15]; the first hit is due to the accumulation of triglycerides in hepatocytes causing liver steatosis, and the second hit is attributable to many factors including oxidative stress and increased cytokines leading to inflammatory changes.

NAFLD in children and adolescents is associated with serious long-term adverse outcomes, including increased risk of mortality and liver cirrhosis requiring transplantation [16]. Multiple risk factors and medical conditions including eating disorders have been linked to the etiology of NAFLD [17,18]. Obesity has been consistently identified as an important risk factor for pediatric NAFLD [11]. Insulin resistance is another risk factor for NAFLD; Nobili et al prospectively followed 84 children with NAFLD and reported that they were almost always insulin-resistant regardless of their BMI [19]. Moreover, a sedentary lifestyle and high fructose consumption have been described as risk factors for NAFLD [20,21]. Furthermore, epidemiological studies have identified being male, older age, and Hispanic ethnicity as risk factors for childhood NAFLD [10]. Additionally, genetic, cellular, and hormonal factors [22] have been found to influence the transition to inflammatory hepatic changes.

Liver biopsy is the gold standard diagnostic test for NAFLD [23,24], although it is an invasive procedure that can be associated with serious complications. Liver function tests (LFTs) are among the first-line investigations for NAFLD [25,26]. However, the use of LFT to diagnose NAFLD is challenging due to the low sensitivity of this method and the discrepancy around the appropriate cutoff values [13]. Moreover, the normal LFT level does not exclude the presence of advanced NAFLD. In contrast, hepatic ultrasonography (US) is a safe, noninvasive, and widely available imaging tool for the assessment of NAFLD. In adults, it has acceptable sensitivity and specificity (100% and 90%, respectively), especially when the liver fat percentage exceeds 20% [27]. Similar statistical properties were documented in a pediatric study [28]. Other radiological diagnostic tools for NAFLD include computed tomography (CT), magnetic resonance imaging (MRI), and magnetic imaging spectroscopy (MRS) [29].

There is limited information regarding the risk of NAFLD in children and adolescents treated with AAP despite their association with metabolic complications that are considered risk factors for the development of NAFLD [30,31]. The objective of our study is to conduct a comprehensive systematic review of the available literature to examine the association between AAP use and NAFLD in children and adolescents.

Methods

Study Identification

To capture all relevant literature on NAFLD among children and adolescents on AAP, we plan to conduct 2 systematic literature reviews. In our first search, we aim to identify studies that assess NAFLD in children and adolescents taking AAP. Given that we expect a paucity in such studies based on a pilot search performed, we will include studies of different designs including cohort and case-control studies, case reports, and case series. In our second systematic search, we will comprehensively review AAP trials in children and adolescents, attempting to identify any reports of NAFLD; we intend to determine whether this was documented as a primary outcome, a secondary outcome, or as an incidental finding in these studies.

Our systematic review will be performed according to this predefined protocol that describes the objectives, search strategy, eligibility criteria, and evaluation methods according to the PRISMA (Preferred Reporting items for Systematic Reviews and Meta-Analysis) guidelines [32].

Systematic Review Methodology

We will conduct 2 systematic literature reviews. The search is restricted to studies published in English from January 1, 1950, until March 31, 2021. We selected 1950 as the starting year for the literature search, as it coincides with the development of the first antipsychotic.

In our first search, we aim to identify studies that assess NAFLD in children and adolescents taking AAP. We will use the following search terms: "second-generation neuroleptics," antipsychotics," and "neuroleptics," with their generic and brand names given in Table 1. Further, the following variations of pediatric NAFLD, namely "NAFLD," "NASH," "nonalcoholic fatty liver disease," "nonalcoholic steatohepatitis," "hepatic steatosis," "fatty liver disease," "nonalcoholic fatty liver," and "fatty liver." In our second search, we will include the following keywords: "atypical antipsychotics" and "atypical neuroleptics" [33].

Table 1. Generic and brand names of all the atypical antipsychotics in the inclusion criteria of our first and second literature searches.

Medication name	Brand name(s)
Amisulpride	Amazeo, Amipride, Amival, Solian, Soltus, Sulpitac, Sulprix, Midora, Socian
Aripiprazole	Abilify, Abilify Maintena, Abilicare, Abilia, Abelfiz, Abdin, Abizol, Abyraz, Aceprofen, Adexyl, Adwiprazole, Alcartis, Alembic, Pipzol, Amdoal, Anasil, Andepro, Antredamin, Ao Pai, Apalife, Apaloz, Apipral, Apiprax, Apra, Aprizexen, Arena, Arepexane, Aria, Aribit, Aricogan, Arifay, Arileto, Arilan, Arilex, Arimed, Aripa, Aripat, Aripegis, Aripen, Aripi-Hexal, Aripilek, Aripip, Aripipa, Aripipan, Aripiprazol, Aripiprex, Aripizin, Ariple, Ariply, Arip-MT, Aripra, Ariprax, Ariprazole, Ariprexa, Ariprizol, Aripsan, Ariski, Arisppa, Aristab, Arive, Arives, Arixind, Arize, Arizol, Arlemide, Arpilif, Arpit, Arpizol, Arpoya, Arypiprazol, Glenmark, Aryzalera, Arzip, Arzu, Asduter, Asprito, Astoret, Atfren, Azolar, Azymol, Bipodis, Brisking, Centalify, Confilify, Curexol, Egisazol, Epimate, Explemed, Fixment, Gemplex, Ignis, Ilimit, Ipipral, Irazem, Kavium, Lazurex, Lemidal, Lemilvo, Madepzol, Motruxia, Neoaripi, Oryva, Otsuka Albilify, Parokzol, Paxifor, Pipra-A, Piprason, Prazarit, Rapiproz, Real One, Restigulin, Rima-Fix, Ripazol, Sayfren, Schizofy, Schizopra, Sensaz, Siblix, Siznil, Sizopra, Tevaripiprazole, Trefero, Zolerip, Zolprix, Zydus, Zykalor, Zylaxera, Aristada
Asenapine	Saphris, Sycrest
Blonanserin	Lonasen
Cariprazine	Reagila, Reagyla, Vraylar
Clozapine	Clozaril, Clopine, Clozapine Synthon, Versacloz
Iloperidone	Fanapt, Zomaril
Lurasidone	Latuda, Lusiax, Luradon, Lurap, Lurasidone Hydrochloride, Lustona
Melperone	Bunil, Buronil, Melneurin, Eunerpan
Olanzapine	Abilanz, Absolute, ACT Olanzapine, Aedon, Amulsin, Anzap, Anzatric, Anzorin, Apo Olanzapine, Apzet, Arenbil, Arkolamyl, As-Pineks, Auro Olanzapine, Axonium, Aziva, Balerap, Bloonis, Caprilon, Cap Tiva, Chemmart Olanzapine, Crisapina, Deprex, Domus, Dopin, Dozic, Egolanza, Elynza, Enolex, Epilanz-10, Expolid, Exzapine, Ferzapin, Fontanivio, Fredilan, Irropia, Jamp Olanzapine, Jolyon MD, Joyzol, Ketoconazol Sesderma, Kotico, Kozylex, Lanopin, Lanzafen, Lanzapine, Lanzek, Lanzek Zydis, Lapenza, Lapin, Lapozan, Lazap, Lazapix, Lezapin-MD, Lopez, Lupilan, Malanxin, Manza, Marathon, Marcato, Mar-Olanzapine, Medlanz, Meflax, Meltolan, Midax, Mint Olanzapine, MylanOlanzapine, Nervix, Neupine, Newzypra, Nodoff, Norpen OroNykob, Nyzol, Oceanil, Ofans ODT, Oferta, Oferta Sanovel, Olace, Oladay, Olafer, Olafid,Olan, Olanap, Olandix, Olandoz, Olandus, Olanex, Olanexyn, Olanpax, Olansapiin Mylan, Olansek, Olanstad, Olanz, Olanza, Olanzacor, Olanzalet, Olanzalux, Olanzamed, Olanzapin, Olanzapro, Olanzar, Olanzavitae, Olanzep, Olanzin Olanzyl, Olapex, Olapin, Olapine, Olaprexa, Olastazen, Olavex, Olaxinn, Olazap, Olazax, Olazin, Olazine, Olazofren, Oleanz, Olenxa, Olexar, Olfrex, Olivin, Ollafax, Olmed, Olmyzem, Olnegis, Olpax, Olpin, Olpinat, Oltal, Olza, Olzadin, Olzanid, Olzap, Olzapin, Olzic, Olzin, Onezyp, Onotran, Onza, Onzapin, Opin, Opirap, Ou Lan Ning, Ozapex, Ozapin-MD, Ozapram, Ozaprin, Ozin, Parnasan, Parnassan, Pericam, Pinolza, Placet, PMS-Olanzapine, Polar, Pranza, Prexal, Prexolan, Prolanz, Protil, Pryzex, Psychozap, Ranofren, RAN-Olanzapine, Redilanz, Remital, Revertrix, Rexapin, Rexepi, Rolanzax, Sartina, Simina, Sincris, Sizap, Solazin, Stygapon, Synza, Tolaz, Treana, Trexol, Vaincor, Vaira, Villamos, Ximin, Xoltiva, Xytrex, Zalasta, Zalepin, Zanprex, Zap, Zapiluks, Zapilux, Zapin, Zapinex FT, Zappa, Zaprinel, Zapris, Zelta, Zeprex, Zesten, Ziora, Zirmapina, Zofrenix, Zolafren, Zolamelt, Zolapine, Zolaswift, Zolaxa, Zonapin, Zophix, Zopix, Zopridoxin, Zoxil, Zylanza, Zypadhera, Zypeace, Zypine, Zyprexa, Zyzapin, Zypadhera, Zyprexa Relprevv
Paliperidone	Aspire-XR, Inveda, Invega, Palido, Parnido, Trevicta, Xeplion
Quetiapine	Actawell, Adequet, Aebol, Afidat, Aretaeus, Asicot, As-Kalmeks, Atip, Atrolak XL, Biquelle XL, Biquetan, Bonogren, Brancico XL, Catepsin, Cizyapine, Dopaquel, Equelib, Esertia, Etiagen XR, Etipin, Geldoren, Gofyl, Hedonin, Keday XR, Kenantis, Ketap, Ketidose, Ketilept, Ketinel, Ketipina, Ketipine, Ketipinor, Ketrel, Kvelux, Kventiax, Kvineva, Kwetaplex XR, Loquen, Mintreleq XL PR, Mylan Quetiapine, Nantarid, Norsic, Pincalm, Pinexet, Placidin, Psynil, Psyquel, Psyquet, Q-Mind, Q-Pin, Qpine, Quantia, Queapin, Quel, Quentiax, Queopine, Quepigal, Quepimax, Queropax, Quetap, Queteper, Quetia, Quetiapina, Quetiaros, Quetiazic, Quetidin, Quetimax, Quetin, Quetipax, Quetipin, Quetirel, Quetium, Quetkare, Quetoser, Quitapex, Qurax, Quser, Qutace, Qutan, Quticool, Qutipin, Q-Win, Sequa, Serenase, Serex, Seroquel, Setinin, Sizonorm, Sizoquit, Socalm, Sofrel, Symquel XR, Tevaquel, Tiapinan, Tiapine, Tomel, Treksta, Valir, Vesparax, Volqer, Vorta, Alcreno, Alzen, Anaquetan XR, ApoTiapina, Arezil XR, Asicot, Atrolak, Biatrix, Biquetan, Brevenox, Cacepin, Calm-ez, Cedrina, Centroqueen, Delucon, Dendritex, Derin, Derin Prolong, Dominium, Edagan, Etiaben XR, Etiagen, Etiapin, Etiasel, Eufrenin, Geroquel, Gofyl, Gyrex, Hedonin, Hiloca, Ilufren, Inquetia, Kagitz, Kalm, Kaptan, Kefrenex, Kemoter, Kesaquil, Ketian XR, Ketiap, Ketilept, Ketipine, Ketipinor, Ketya, Kvelux, Kventiax, Kvetiapin, Kwetaplex, Kwetax, Limus, Loquen, Loquen XR, Matepil, Megazone, Nantarid, Netiapin, Neuroquel, Neutapin, Psicotric
Risperidone	Risperdal, Risperdal Consta, Risperdal M-Tab, Risperdal Quicklets, Risperlet
Sertindole	Serdolect, Serlect
Sulpiride	Dogmatil, Dolmatil, Eglonyl, Espiride, Modal, Prometar, Sulpor
Ziprasidone	Geodon, Pramaxima, Vikolus, Ypsila, Zeldox, Zipradon, Zipragen, Zipramyl, Ziprasidon, Zipsydon, Zipwell, Zypsila, Zypsilan, Zeldox
Zotepine	Losizopilon, Lodopin, Setous, Zoleptil

Medications in combination with Fluoxetine will not be included in the research (ie, Olanzapine: Co-Depricap, Depten-OZ,Rexepi Combi,Symbyax, Tagram, Target, Olapin Forte, Olapin Plus,Olanex F, Oladay-F).

In our second systematic search, we will comprehensively review AAP trials in children and adolescents, attempting to identify any reports of NAFLD; we aim to determine whether this was documented as a primary outcome, a secondary outcome, or as an incidental finding during these studies. The following databases will be used to conduct the search: Ovid MEDLINE, Cochrane, and Cumulative Index to Nursing and Allied Health Literature (CINAHL). We also searched the following databases up to 2016: Embase, Web of Science, BIOSIS Previews, and PsycINFO.

A systematic, computer-assisted search of the following databases will be performed by 4 medical students RH, FN, GA, and MA at the Mohammed Bin Rashid University (MBRU): Ovid MEDLINE, Embase, Web of Science, BIOSIS Previews, PsycINFO, and CINAHL. The search strategy will be supported by librarian Shakeel Tegginmani at the MBRU Library. Furthermore, the bibliographies of the retrieved and relevant articles will be manually searched for relevant articles. Full articles in English published in peer-reviewed journals will be included in this review.

As mentioned earlier, we will conduct 2 searches that will be limited to children and adolescents younger than 18 years of age. The first search will include the following variations of AAP: “atypical antipsychotics,” “atypical neuroleptics,” “second-generation antipsychotics,” “second-generation neuroleptics,” “antipsychotics,” and “neuroleptics” for AAP [34].

Eligibility Criteria

We determined the following inclusion and exclusion criteria based on a literature review and the objectives of this study. All medication trials published in English pertaining to AAP in subjects younger than 18 years of age that reported NAFLD as an outcome, assessed by radiological methods (including liver ultrasound) or liver biopsy, will be included in this review. Moreover, prospective and retrospective observational studies of children and adolescents on AAP with reported NAFLD indicators will be included. Finally, case series and reports of NAFLD in children and adolescents on AAPs will be included. We will exclude conference abstracts, editorials, letters to editors, treatment guidelines, and studies published as abstracts only. Reviews (systematic and nonsystematic) will be excluded; however, the bibliographies of relevant papers will be reviewed. We will exclude studies in which AAP were used as an add-on or in combination with other medications (such as antiepileptics, first-generation antipsychotics, and selective serotonin uptake inhibitors), and medications that are a combination of AAP and selective serotonin uptake inhibitors including Fluoxetine; however, we will include studies where AAP constitute 1 arm of the study [35]. Due to the overlap of NAFLD risk factors with other comorbidities and varying nutritional backgrounds, we will also exclude studies examining patients with pre-existing medical conditions including eating disorders [17,18]. Our search will be limited to studies published in English, and nonhuman studies will be excluded. The review and comparison of results will be conducted using Endnote.

Outcomes

The primary outcome of the present study is NAFLD in children and adolescents, as assessed by either liver biopsy or using a radiological tool including hepatic US, MRI, MRS, and transient elastography among children and adolescents on AAP. Despite the relationship between NAFLD and metabolic syndrome, we did not include variables of metabolic syndrome in our outcomes because it is not the focus of our review and it has been previously studied, including in a recent meta-analysis [36]. Secondary outcomes will include changes in liver enzymes.

Data Extraction

Investigators RH, FAN, GAA, and MA will independently review the titles and abstracts of the retrieved studies and exclude duplicates and irrelevant studies based on the aforementioned eligibility criteria. All the potentially eligible abstracts will be further assessed for eligibility by thoroughly reviewing their full texts. Results will be compared, and discrepancies will be resolved by consensus and by consulting investigators AA, CT, and ELA when needed. The Cohen κ will be calculated as a measure of interrater agreement. All studies that meet the eligibility criteria will be included in our analyses. Data extraction will be carried out using a standard form.

The quality of the studies will be assessed using the STROBE (Strengthening the Reporting of Observational studies in Epidemiology) guidelines [37] for cohort, case-control, and cross-sectional studies and the Newcastle-Ottawa Scale for nonrandomized studies. The quality of randomized controlled trials will be evaluated using the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) method.

If the results allow for conducting a meta-analysis, we will report the outcomes by calculating the weighted pooled estimate of changes in the outcomes [38]. The risk of bias will be evaluated using three types of homogeneity tests: (1) forest plot, (2) Cochrane Q test (chi-square test), and (3) Higgins I2 statistics. In the forest plot, greater overlap between the CIs indicates greater homogeneity.

Results

The search results will be presented using the PRISMA flowchart. This study will comprehensively review literature pertaining to NAFLD in children and adolescents taking AAP.

The first and the second systematic searches will be conducted during December 2021. The title and abstract review will be performed by RH, FAN, GAA, and MA between December 1 and December 15, 2021. Further, the full-text review will be performed by the same researchers between December 15 and December 25, 2021. The results are expected to be published in June 2022. The results of this study may inform clinical guidelines for AAP use in children and adolescents.

Discussion

Clinical Significance

Studies have suggested that the pediatric population on AAP is more at risk of developing long-term adverse effects due to AAP, including weight gain, hyperlipidemia, and insulin resistance [14]. Therefore, we have grounds to speculate that children on AAP may be at a greater risk of developing disorders associated with metabolic syndrome, including NAFLD. This research project is specifically important due to the current trends in overprescribing AAP in children aged below 18 years [39].

Due to the dearth of systematic reviews on this vital topic, the need for understanding the possible association between AAP and NAFLD in this population across the literature landscape is warranted. Conducting this systematic review would serve as a comprehensive foundation for future studies and help devise interventions for the child and adolescent population.

Conclusions

The PRISMA flowchart was chosen to present the results of this study. There is a global rise in the use of AAP in treating children and adolescents below the age of 18 [39], which can cause long-term adverse effects and metabolic complications. We hope this research project will serve as a foundation for future studies and assist in devising interventions and reforming clinical guidelines of AAP use for improved patient safety.

Conflicts of Interest

None declared.

References

1. Pringsheim T, Lam D, Patten SB. The pharmacoepidemiology of antipsychotic medications for Canadian children and adolescents: 2005–2009. J Child Adolesc Psychopharmacol 2011 Dec;21(6):537-543. [doi: 10.1089/cap.2010.0145] [Medline: 22136092]
2. Fedorowicz VJ, Fombonne E. Metabolic side effects of atypical antipsychotics in children: a literature review. J Psychopharmacol 2005 Sep;19(5):533-550. [doi: 10.1177/0269881105056543] [Medline: 16166191]
3. Perez-Iglesias R, Vazquez-Barquero JL, Amado JA, Berja A, Garcia-Unzueta MT, Pelayo-Terán JM, et al. Effect of antipsychotics on peptides involved in energy balance in drug-naive psychotic patients after 1 year of treatment. J Clin Psychopharmacol 2008 Jun;28(3):289-295. [doi: 10.1097/JCP.0b013e318172b8e6] [Medline: 18480685]
4. Reynolds GP, Kirk SL. Metabolic side effects of antipsychotic drug treatment–pharmacological mechanisms. Pharmacol Ther 2010 Jan;125(1):169-179. [doi: 10.1016/j.pharmthera.2009.10.010] [Medline: 19931306]
5. Hammerman A, Dreiher J, Klang SH, Munitz H, Cohen AD, Goldfracht M. Antipsychotics and diabetes: an age-related association. Ann Pharmacother 2008 Sep;42(9):1316-1322. [doi: 10.1345/aph.1L015] [Medline: 18664607]
6. Erdogan A, Atasoy N, Akkurt H, Ozturk D, Karaahmet E, Yalug I, et al. Risperidone and liver function tests in children and adolescents: a short-term prospective study. Prog Neuropsychopharmacol Biol Psychiatry 2008 Apr;32(3):849-857. [doi: 10.1016/j.pnpbp.2007.12.032] [Medline: 18258348]
7. Erdogan A, Karaman MG, Ozdemir E, Yurteri N, Tufan AE, Kurcer MA. Six months of treatment with risperidone may be associated with nonsignificant abnormalities of liver function tests in children and adolescents: a longitudinal, observational study from Turkey. J Child Adolesc Psychopharmacol 2010 Oct;20(5):407-413. [doi: 10.1089/cap.2009.0113] [Medline: 20973711]
8. Karaman MG, Erdoğan A, Tufan E, Yurteri N, Ozdemir E, Ankarali H. Liver function tests in children and adolescents receiving risperidone treatment for a year: a longitudinal, observational study from Turkey. Int J Psychiatry Clin Pract 2011 Sep;15(3):204-208. [doi: 10.3109/13651501.2011.582537] [Medline: 22121930]
9. Patel NC, Crismon ML, Hoagwood K, Johnsrud MT, Rascati KL, Wilson JP, et al. Trends in the use of typical and atypical antipsychotics in children and adolescents. J Am Acad Child Adolesc Psychiatry 2005 Jun;44(6):548-556. [doi: 10.1097/01.chi.0000157543.74509.c8] [Medline: 15908837]
10. Schwimmer JB, Deutsch R, Kahen T, Lavine JE, Stanley C, Behling C. Prevalence of fatty liver in children and adolescents. Pediatrics 2006 Oct;118(4):1388-1393. [doi: 10.1542/peds.2006-1212] [Medline: 17015527]
11. Chan DFY, Li AM, Chu WCW, Chan MHM, Wong EMC, Liu EKH, et al. Hepatic steatosis in obese Chinese children. Int J Obes Relat Metab Disord 2004 Oct;28(10):1257-1263. [doi: 10.1038/sj.ijo.0802734] [Medline: 15278103]
12. Fu J, Shi H, Liu L, Jiang P, Liang L, Wang C, et al. Non-alcoholic fatty liver disease: an early mediator predicting metabolic syndrome in obese children? World J Gastroenterol 2011 Feb;17(6):735-742 [FREE Full text] [doi: 10.3748/wjg.v17.i6.735] [Medline: 21390143]
13. Vajro P, Lenta S, Socha P, Dhawan A, McKiernan P, Baumann U, et al. Diagnosis of nonalcoholic fatty liver disease in children and adolescents: position paper of the ESPGHAN Hepatology Committee. J Pediatr Gastroenterol Nutr 2012 May;54(5):700-713. [doi: 10.1097/MPG.0b013e318252a13f] [Medline: 22395188]
14. Holtmann M, Kopf D, Mayer M, Bechtinger E, Schmidt MH. Risperidone-associated steatohepatitis and excessive weight-gain. Pharmacopsychiatry 2003 Sep;36(5):206-207. [doi: 10.1055/s-2003-43045] [Medline: 14571356]

15. AlKhater SA. Paediatric non-alcoholic fatty liver disease: an overview. Obes Rev 2015 May;16(5):393-405. [doi: 10.1111/obr.12271] [Medline: 25753407]
16. Bhala N, Angulo P, van der Poorten D, Lee E, Hui JM, Saracco G, et al. The natural history of nonalcoholic fatty liver disease with advanced fibrosis or cirrhosis: an international collaborative study. Hepatology 2011 Oct;54(4):1208-1216 [FREE Full text] [doi: 10.1002/hep.24491] [Medline: 21688282]
17. Yasutake K, Kohjima M, Kotoh K, Nakashima M, Nakamuta M, Enjoji M. Dietary habits and behaviors associated with nonalcoholic fatty liver disease. World J Gastroenterol 2014 Feb;20(7):1756-1767 [FREE Full text] [doi: 10.3748/wjg.v20.i7.1756] [Medline: 24587653]
18. Zhang J, Abbasi O, Malevanchik L, Mohan N, Denicola R, Tarangelo N, et al. Pilot study of the prevalence of binge eating disorder in non-alcoholic fatty liver disease patients. Ann Gastroenterol 2017 Oct;30(6):664-669 [FREE Full text] [doi: 10.20524/aog.2017.0200] [Medline: 29118561]
19. Nobili V, Marcellini M, Devito R, Ciampalini P, Piemonte F, Comparcola D, et al. NAFLD in children: a prospective clinical-pathological study and effect of lifestyle advice. Hepatology 2006 Aug;44(2):458-465. [doi: 10.1002/hep.21262] [Medline: 16871574]
20. Nseir W, Hellou E, Assy N. Role of diet and lifestyle changes in nonalcoholic fatty liver disease. World J Gastroenterol 2014 Jul;20(28):9338-9344 [FREE Full text] [doi: 10.3748/wjg.v20.i28.9338] [Medline: 25071328]
21. Jin R, Welsh JA, Le N, Holzberg J, Sharma P, Martin DR, et al. Dietary fructose reduction improves markers of cardiovascular disease risk in Hispanic-American adolescents with NAFLD. Nutrients 2014 Aug;6(8):3187-3201 [FREE Full text] [doi: 10.3390/nu6083187] [Medline: 25111123]
22. Hassan K, Bhalla V, El Regal ME, A-Kader HH. Nonalcoholic fatty liver disease: a comprehensive review of a growing epidemic. World J Gastroenterol 2014 Sep;20(34):12082-12101 [FREE Full text] [doi: 10.3748/wjg.v20.i34.12082] [Medline: 25232245]
23. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, Nonalcoholic Steatohepatitis Clinical Research Network. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005 Jun;41(6):1313-1321. [doi: 10.1002/hep.20701] [Medline: 15915461]
24. Mann JP, Goonetilleke R, McKiernan P. Paediatric non-alcoholic fatty liver disease: a practical overview for non-specialists. Arch Dis Child 2015 Jul;100(7):673-677. [doi: 10.1136/archdischild-2014-307985] [Medline: 25633064]
25. Kumra S, Herion D, Jacobsen LK, Briguglia C, Grothe D. Case study: risperidone-induced hepatotoxicity in pediatric patients. J Am Acad Child Adolesc Psychiatry 1997 May;36(5):701-705. [doi: 10.1097/00004583-199705000-00022] [Medline: 9136506]
26. Copur M, Erdogan A. Risperidone rechallenge for marked liver function test abnormalities in an autistic child. Recent Pat Endocr Metab Immune Drug Discov 2011 Sep;5(3):237-239. [doi: 10.2174/187221411797265872] [Medline: 21913889]
27. Dasarathy S, Dasarathy J, Khiyami A, Joseph R, Lopez R, McCullough AJ. Validity of real time ultrasound in the diagnosis of hepatic steatosis: a prospective study. J Hepatol 2009 Dec;51(6):1061-1067 [FREE Full text] [doi: 10.1016/j.jhep.2009.09.001] [Medline: 19846234]
28. Shannon A, Alkhouri N, Carter-Kent C, Monti L, Devito R, Lopez R, et al. Ultrasonographic quantitative estimation of hepatic steatosis in children With NAFLD. J Pediatr Gastroenterol Nutr 2011 Aug;53(2):190-195 [FREE Full text] [doi: 10.1097/MPG.0b013e31821b4b61] [Medline: 21788761]
29. Lee SS, Park SH. Radiologic evaluation of nonalcoholic fatty liver disease. World J Gastroenterol 2014 Jun;20(23):7392-7402 [FREE Full text] [doi: 10.3748/wjg.v20.i23.7392] [Medline: 24966609]
30. Manco M, Marcellini M, Devito R, Comparcola D, Sartorelli MR, Nobili V. Metabolic syndrome and liver histology in paediatric non-alcoholic steatohepatitis. Int J Obes (Lond) 2008 Feb;32(2):381-387. [doi: 10.1038/sj.ijo.0803711] [Medline: 18087267]
31. Lui SY, Tso S, Lam M, Cheung EFC. Possible olanzapine-induced hepatotoxicity in a young Chinese patient. Hong Kong Med J 2009 Oct;15(5):394-396 [FREE Full text] [Medline: 19801701]
32. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009 Oct;62(10):e1-e34 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.06.006] [Medline: 19631507]
33. Lauressergues E, Staels B, Valeille K, Majd Z, Hum DW, Duriez P, et al. Antipsychotic drug action on SREBPs-related lipogenesis and cholesterogenesis in primary rat hepatocytes. Naunyn Schmiedebergs Arch Pharmacol 2010 May;381(5):427-439. [doi: 10.1007/s00210-010-0499-4] [Medline: 20333360]
34. Oh K, Park J, Lee SY, Hwang I, Kim JB, Park T, et al. Atypical antipsychotic drugs perturb AMPK-dependent regulation of hepatic lipid metabolism. Am J Physiol Endocrinol Metab 2011 Apr;300(4):E624-E632 [FREE Full text] [doi: 10.1152/ajpendo.00502.2010] [Medline: 21224484]
35. Bonhaus DW, Weinhardt KK, Taylor M, DeSouza A, McNeeley PM, Szczepanski K, et al. RS-102221: a novel high affinity and selective, 5-HT2C receptor antagonist. Neuropharmacology 1997 Apr;36(4-5):621-629. [doi: 10.1016/s0028-3908(97)00049-x] [Medline: 9225287]

36. Almandil NB, Liu Y, Murray ML, Besag FMC, Aitchison KJ, Wong ICK. Weight gain and other metabolic adverse effects associated with atypical antipsychotic treatment of children and adolescents: a systematic review and meta-analysis. Paediatr Drugs 2013 Apr;15(2):139-150. [doi: 10.1007/s40272-013-0016-6] [Medline: 23519708]
37. da Costa BR, Cevallos M, Altman DG, Rutjes AWS, Egger M. Uses and misuses of the STROBE statement: bibliographic study. BMJ Open 2011 Feb;1(1):e000048 [FREE Full text] [doi: 10.1136/bmjopen-2010-000048] [Medline: 22021739]
38. Ahn E, Kang H. Introduction to systematic review and meta-analysis. Korean J Anesthesiol 2018 Apr;71(2):103-112 [FREE Full text] [doi: 10.4097/kjae.2018.71.2.103] [Medline: 29619782]
39. Harrison JN, Cluxton-Keller F, Gross D. Antipsychotic medication prescribing trends in children and adolescents. J Pediatr Health Care 2012 Mar;26(2):139-145 [FREE Full text] [doi: 10.1016/j.pedhc.2011.10.009] [Medline: 22360933]

Abbreviations

AAP: atypical antipsychotics
LFT: liver function test
MRI: magnetic resonance imaging
MRS: magnetic resonance spectroscopy
NAFLD: nonalcoholic fatty liver disease
PRISMA: Preferred Reporting items for Systematic Reviews and Meta-Analysis
US: ultrasonography

Edited by G Eysenbach; submitted 26.03.21; peer-reviewed by T Ewais, B Foroutan; comments to author 14.08.21; revised version received 14.11.21; accepted 30.11.21; published 21.03.22.

Please cite as:
Hatem R, Nawaz FA, Al-Sharif GA, Almoosa M, Kattan W, Tzivinikos C, Amirali EL, Albanna A
Nonalcoholic Fatty Liver Disease in Children and Adolescents Taking Atypical Antipsychotic Medications: Protocol for a Systematic Review and Meta-analysis
JMIR Res Protoc 2022;11(3):e20168
URL: https://www.researchprotocols.org/2022/3/e20168
doi:10.2196/20168
PMID:35311689

Protocol

Associations Between Prenatal Exposure to Serotonergic Medications and Biobehavioral Stress Regulation: Protocol for a Systematic Review and Meta-analysis

Enav Z Zusman[1,2,3], MSc, PharmD; Alekhya Lavu[4], PharmD; Colleen Pawliuk[1], MLIS; Jodi Pawluski[5], PhD; Sarah M Hutchison[1,3], PhD; Robert W Platt[6,7,8], PhD; Tim F Oberlander[1,9], MD

[1]Department of Pediatrics, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

[2]Department of Obstetrics & Gynaecology, University of British Columbia, Vancouver, BC, Canada

[3]British Columbia Children's Hospital Research Institute, Vancouver, BC, Canada

[4]College of Pharmacy, University of Manitoba, Winnipeg, MB, Canada

[5]Univ Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail), UMR_S 1085, F-35000, Rennes, France

[6]Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, Montreal, QC, Canada

[7]Department of Pediatrics, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada

[8]Department of Epidemiology, Biostatistics, and Occupational Health, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada

[9]School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Enav Z Zusman, MSc, PharmD
Department of Pediatrics
Faculty of Medicine
University of British Columbia
4480 Oak St
Vancouver, BC, V6H 3V4
Canada
Phone: 1 (604) 875 3200
Email: enav.zusman@bcchr.ca

Abstract

Background: Up to 20% of mothers experience antenatal depression and approximately 30% of these women are treated with serotonergic psychotropic pharmacological therapy during pregnancy. Serotonergic antidepressants readily cross the placenta and the fetal blood-brain barrier, altering central synaptic serotonin signaling and potentially altering serotonin levels in the developing fetal brain.

Objective: The aim of this study is to assess the impact of prenatal exposure to serotonergic antidepressants, accounting for maternal mood disturbances, on markers of stress regulation during childhood.

Methods: We will follow PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines and will search MEDLINE, Embase, CINAHL, PsycINFO, and ClinicalTrials.gov for full-length studies that assessed physiological (eg, cortisol level, heart rate variability, salivary amylase, pupillary size, C-reactive protein) indices of stress regulation in children of pregnant people who were treated with a serotonergic antidepressant at any point during pregnancy. We will assess the quality of observational studies using the Newcastle-Ottawa Scale and the quality of experimental studies using the Cochrane risk-of-bias tool. When possible, we will conduct a random-effects meta-analysis. If meta-analysis is not possible, we will conduct a narrative review. If a sufficient number of studies are found, we will perform subgroup analysis and assess outcomes measured by drug class, dose, trimester of exposure, and child's age and gender.

Results: We registered our review protocol with PROSPERO (International Prospective Register of Systematic Reviews; CRD42021275750), completed the literature search, and initiated title and abstract review in August 2021. We expect to finalize this review by April 2022.

Conclusions: Findings should identify the impact of prenatal antidepressant effects on stress regulation and distinguish it from the impact of prenatal exposure to maternal mood disturbances. This review should inform decisions about serotonergic antidepressant use during pregnancy.

Trial Registration: PROSPERO CRD42021275750; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=275750

International Registered Report Identifier (IRRID): PRR1-10.2196/33363

(JMIR Res Protoc 2022;11(3):e33363) doi:10.2196/33363

KEYWORDS

pregnancy; serotonergic medications; antidepressants; stress regulation; systematic review; meta-analysis

Introduction

Background

Up to 20% of mothers experience antenatal depression and approximately 30% of these women are treated with a serotonergic antidepressant during pregnancy [1,2]. Selective serotonin reuptake inhibitors (SSRI) are the most common serotonergic medications prescribed [3-7]. They readily cross the placenta and the fetal blood-brain barrier, potentially altering serotonin (or 5-hydroxytryptamine [5-HT]) signaling in the fetal brain [8-14], and such exposure has been paradoxically reported to be associated with an increased risk for anxiety, attention, and behavioral disorders in children of mothers with depression treated with an SSRI during pregnancy [15-17]. Importantly, childhood behaviors have been associated with altered indices of stress regulation [18-21], raising critical questions about whether prenatal exposure to serotonergic psychotropic medications alters stress reactivity, thereby contributing to an increased risk for behavioral disturbances.

Long before 5-HT becomes a neurotransmitter in the mature brain, it plays a role as a neurodevelopmental signal regulating cell growth and function [22,23]. In the fetal brain, 5-HT and its receptors are overexpressed and widespread in regions where they are absent in adults, pointing to a time-dependent specificity to 5-HT expression during development [24,25]. Early 5-HT alterations, either via pharmacological, genetic, or other manipulations, potentially alter these processes via the presynaptic, membrane-bound serotonin transporter protein (5-HTT), the target of SRI antidepressants. 5-HTT is a key regulator of brain 5-HT [26,27].

Serotonin is central to the development and function of two key stress response systems—the locus-coeruleus-norepinephrine (autonomic nervous system [ANS]) and the hypothalamic-pituitary-adrenal (HPA) systems [28-30], which may illustrate sites affected by prenatal exposure to serotonin reuptake inhibitors on stress responses [31-33]. The relationship between 5-HT and stress reactivity is bidirectional; stressors appear to alter 5-HT metabolism as well as bias how one copes with subsequent stressful challenges [32,34]. ANS activation leads to a rapid “flight or fight” response, in turn leading to increased cardiac activity (heart rate) and release of catecholamines (norepinephrine) [35]. Central to our understanding of how prenatal exposure to serotonergic medications such as antidepressants influences early brain development is understanding the diverse roles the neurotransmitter 5-HT plays in early brain development, stress regulation, and mental health [23,28,36]. Prenatal maternal mood disturbances, the very disorders that lead to antidepressant treatment, have also been shown to shape the development of the HPA axis [37,38].

5-HT and cardiovascular/autonomic stress regulation are highly interrelated via links between reflex control of parasympathetic outflow to the heart and other organs that involve central 5-HT1A receptors located in the vicinity of preganglionic vagal neurons. Further, 5-HT3 receptors are implicated in afferent regulation of central sympathetic and parasympathetic tone [39]. The development and function of the HPA stress response and the serotonergic regulatory systems are highly interrelated and exquisitely sensitive to the effects of early adverse experience [40,41]. Serotonin influences how an individual copes with subsequent social stressors and plays a role in mediating the effects of adverse experience [42]. Early differences in maternal care alter central 5-HT levels that change HPA axis stress function, reflected as an altered capacity to regulate stress responses [40,43].

Considering the importance of serotonin in neurodevelopment, it is conceivable that early changes to 5-HT, secondary to prenatal serotonergic medication exposure or maternal mood disorders, could have developmental consequences [26,44] and may modify the formation and function of key stress regulatory systems such as the ANS and HPA axis in ways that may affect subsequent responses to stress challenges and may have life-long implications for the offspring’s health and behavior. Understanding relationships between stress response systems and children’s behavior may provide essential insight into the developmental origins of physiological processes that contribute to disrupted behavior. Taken together, it is possible that prenatal exposure to serotonergic psychotropic medications used to manage mood disturbances during pregnancy could alter stress reactivity/regulation in offspring.

Objectives

Our study aims to assess the impact of prenatal exposure to serotonergic medications, and distinguish these effects from the impact of maternal mood disturbances on neonatal, infant, childhood, and adolescent indices of stress regulation.

Methods

Overview

We will adhere to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for reporting systematic reviews [45] and have used the PRISMA for

systematic review protocols (PRISMA-P) [46]. Our research protocol was designed a priori, defining methods for searching the literature, including and examining articles, and extracting and analyzing data.

Eligibility Criteria

Inclusion Criteria

This systematic review will consider studies that included pregnant people diagnosed with prenatal mood disorders (depression and/or anxiety) who were exposed to serotonergic medications at any point during pregnancy. We will assess stress regulation outcomes in the offspring and the way they relate to behavior. We will include monopharmacy use of antidepressants. We will include both singleton and multiple gestation pregnancies as well as both nulliparous and multiparous pregnancies. We will include intervention studies (randomized controlled trials, pre-post trials) and observational trials (case-control studies, cross-sectional studies, cohort studies, case reports or case series). We will only include full-text studies published in English or French. Studies that meet our inclusion criteria will be included in this review.

Exclusion Criteria

We will exclude polypharmacy use of multiple antidepressant medications from several classes, as well as animal studies, gray literature (including theses and dissertations), review studies, letters to the editor, conference abstracts, and posters. If we come across different studies that include the same population and outcome, the study that involves a longer follow-up will be included.

Outcome Measures

We will assess physiological outcomes and how they relate to each other. Specifically, our outcome includes the following physiological outcomes: cortisol, heart rate variability, salivary amylase, pupillary size, C-reactive protein (CRP), and immunological biomarkers (cytokines, chemokines, lymphokines, IL-6, etc).

Information Sources and Literature Search

We developed our search strategy with the consultation of a librarian and will search the literature by population and intervention (Table 1). The following databases will be searched: MEDLINE (Ovid), Embase (Ovid), CINAHL (EBSCOhost), PsycINFO (EBSCOhost), and ClinicalTrials.gov. We will use both Medical Subject Headings (MeSH) terms and keywords. Our search strategy for the MEDLINE database is outlined in Multimedia Appendix 1. We will adjust the MeSH terms and keywords used to accommodate the different databases' requirements and limitations. We will limit our search to studies published in English or French. We will not limit for year of publication and will include original studies of all study types.

Table 1. Eligibility criteria to be included in the review.

Item	Criteria
Population	Pregnant people diagnosed with antepartum depression, prenatal depression, or maternal mood disorders
Intervention/exposure	Exposure to a serotonergic drug: • SSRI[a]: citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline • SNRI[b]: desvenlafaxine, duloxetine, levomilnacipran, venlafaxine • Second-generation antipsychotics: aripiprazole, brexpiprazole, olanzapine, quetiapine, risperidone • Serotonin modulators: trazodone, vilazodone, vortioxetine • Tricyclic antidepressants: amitriptyline, clomipramine, desipramine, doxepin, imipramine, nortriptyline, trimipramine • 5-HT1A (serotonin receptor) agonist: buspirone
Comparison/control	None
Outcome	Physiological outcomes: cortisol, cardiac autonomic function (heart rate variability, pre-ejection period), salivary amylase, pupillary size, C-reactive protein, immunological biomarkers (cytokines, chemokines, lymphokines, IL-6).

[a]SSRI: selective serotonin reuptake inhibitor.

[b]SNRI: serotonin–norepinephrine reuptake inhibitor.

We will use the following keywords and MeSH terms: Pregnancy Trimesters/ or Pregnancy/ or Pregnancy Trimester, Third/ or pregnancy.mp. or Pregnancy Trimester, First/ or Pregnancy, or Pregnancy Trimester, Second/, pregnant.mp. or Pregnant Women/, or gestation.mp. or perinatal.mp. or prenatal.mp. or pregnant* AND Depression/ or depressive disorder/ or depressive disorder, major/ or depressive disorder, treatment-resistant/ or Anxiety Disorders/ or Anxiety/ or anxiety.mp., or Mood Disorders/, AND antidepressants.mp. or Antidepressive Agents/, serotonergic drugs.mp. or Serotonin Agents/, selective serotonin reuptake inhibitors.mp. or Serotonin Uptake Inhibitors/, citalopram.mp. or Citalopram/, escitalopram.mp. or Citalopram/, fluoxetine.mp. or Fluoxetine/, fluvoxamine.mp. or Fluvoxamine/, paroxetine.mp. or Paroxetine/, sertraline.mp. or Sertraline/, "Serotonin and Noradrenaline Reuptake Inhibitors"/ or Serotonin Uptake Inhibitors/ or serotonin-norepinephrine reuptake Inhibitor.mp., desvenlafaxine.mp., or Desvenlafaxine Succinate/, duloxetine.mp. or Duloxetine Hydrochloride/, levomilnacipran.mp. or Levomilnacipran/, venlafaxine.mp. or Venlafaxine Hydrochloride/, Antipsychotic Agents/ or Second generation antipsychotics.mp., aripiprazole.mp., or Aripiprazole/, brexpiprazole.mp., olanzapine.mp., or Olanzapine/, quetiapine.mp. or Quetiapine Fumarate,

risperidone.mp. or Risperidone/, Serotonin Modulators.mp., trazodone.mp., or Trazodone/, vilazodone.mp. or Vilazodone Hydrochloride/, vortioxetine.mp. or Vortioxetine/, tricyclic antidepressants.mp. or Antidepressive Agents, Tricyclic/, amitriptyline.mp. or Amitriptyline/, clomipramine.mp. or Clomipramine/, desipramine.mp. or Desipramine/, doxepin.mp. or Doxepin/, imipramine.mp. or Imipramine/, nortriptyline.mp. or Nortriptyline/, trimipramine.mp. or Trimipramine/ or buspirone.mp. or Buspirone/.

We will conduct a manual search of the journals *Psychoneuroendocrinology*, *Early Human Development*, *Neuroscience*, and *Neuroscience & Biobehavioral Reviews*, as well as a forward and backward citation search through Google Scholar [47] on all included papers to locate additional papers that may have been missed in our literature search.

Study Selection Process

Using our predetermined selection criteria, two authors (EZZ and AL) will independently screen all retrieved papers at level 1 (title and abstract) for inclusion in the study using Covidence, a screening and data extraction tool for systematic reviews [48]. Once a list of studies is determined, the selected papers will be reviewed at level 2 (full text) to select a final list of review studies. Screening questions can be found in Multimedia Appendix 2. At any point, authors will meet to discuss discrepancies and a third author (SH) will be consulted if disagreement occurs.

Data Collection Process

Two authors (EZZ and AL) will independently extract the following information from all included studies: year of publication, country, sample size, study design, study setting, trimester of pregnancy, drug exposure class, drug exposure generic name, drug dose, cortisol level (diurnal), cortisol level (stress challenge), heart rate variability, salivary amylase, pupillary size, CRP, cytokines, chemokines, lymphokines, IL-6, maternal depression, maternal depression diagnosis method, maternal anxiety, and maternal anxiety diagnosis method. Authors will meet to discuss discrepancies and a third author (SH) will be consulted if disagreement occurs. If needed, we will contact study corresponding authors for unpublished or missing data.

Quality and Risk of Bias Assessment

To assess the methodological quality of the included studies and their risk of bias, we will use different checklists and scales. Two authors (EZZ and AL) will independently screen each included paper according to its methodology. We will use the Newcastle-Ottawa Scale for observational cohort studies [49], the modified Newcastle-Ottawa Scale for observational cross-sectional studies [50], and the Cochrane risk-of-bias tool for randomized controlled trials and experimental studies [51].

We will use funnel plots to assess for publication bias [52]. In case of a publication bias, we will use the trim-and-fill method. We will remove ("trim") the studies that give rise to the funnel plot's asymmetry and then impute ("fill") the suggested missing studies based on the bias-corrected overall estimate [53].

Synthesis of Included Studies

We will pool studies based on their reported outcome and will present the characteristics of included studies both descriptively and in a table. Where possible, we will pool reported levels of cortisol, heart rate variability, salivary amylase, pupillary size, and CRP. We will calculate the Cochrane Q test (chi-square) and Higgins I^2 score to assess the statistical heterogeneity of effect size estimates across our included studies before running a meta-analysis. When meta-analysis is possible, we will calculate pooled mean differences for continuous data and perform a random-effects meta-analysis for dichotomous data [54]. If meta-analysis is not possible, we will conduct a narrative synthesis of the data. When a sufficient number of studies are found, we will perform subgroup analysis and assess outcomes measured by drug class, dose, trimester of exposure, and child's age and gender. We will consider method of assessment for pooling. We will perform all data analysis using Stata (version 15; StataCorp LLC).

Results

We registered our review protocol with PROSPERO (International Prospective Register of Systematic Reviews; CRD42021275750) [55], completed the literature search, and initiated title and abstract review in August 2021. We expect to finalize this review by April 2022.

Discussion

Antepartum depression is a common condition that is often treated with different serotonergic drugs, leading to altered central serotonin signaling in the developing brain. As serotonin plays a key role in fetal neurodevelopment that shapes key components of stress regulation pathways, understanding how prenatal exposure to these medications affects physiological stress responses could elucidate pathways to behavioral outcomes in the offspring of mothers with depression. This study will add to the existing body of knowledge by integrating data that will lead to new insights about early origins of mental health disorders and the risks and benefits of use of serotonergic medications. The impact of prenatal serotonergic medication exposure and early origins of mental health will be essential for both theoretical and clinical reasons, specifically to inform decisions about serotonergic medication use in pregnancy and to inform interventions that promote healthy child development.

Acknowledgments

EZZ is supported by a Vanier Canada Graduate Scholarship from the Canadian Institutes of Health Research (CIHR) and a University of British Columbia Killam Doctoral Scholarship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy for MEDLINE database.
[PNG File , 304 KB - resprot_v11i3e33363_app1.png]

Multimedia Appendix 2
Reviewers' literature search screening questions.
[DOCX File , 15 KB - resprot_v11i3e33363_app2.docx]

References

1. Millard SJ, Weston-Green K, Newell KA. The effects of maternal antidepressant use on offspring behaviour and brain development: Implications for risk of neurodevelopmental disorders. Neurosci Biobehav Rev 2017 Sep;80:743-765. [doi: 10.1016/j.neubiorev.2017.06.008] [Medline: 28629713]
2. Biaggi A, Conroy S, Pawlby S, Pariante CM. Identifying the women at risk of antenatal anxiety and depression: A systematic review. J Affect Disord 2016 Feb;191:62-77 [FREE Full text] [doi: 10.1016/j.jad.2015.11.014] [Medline: 26650969]
3. Oberlander TF, Warburton W, Misri S, Aghajanian J, Hertzman C. Neonatal outcomes after prenatal exposure to selective serotonin reuptake inhibitor antidepressants and maternal depression using population-based linked health data. Arch Gen Psychiatry 2006 Aug 01;63(8):898-906. [doi: 10.1001/archpsyc.63.8.898] [Medline: 16894066]
4. Ashley JM, Harper BD, Arms-Chavez CJ, LoBello SG. Estimated prevalence of antenatal depression in the US population. Arch Womens Ment Health 2016 Apr 21;19(2):395-400. [doi: 10.1007/s00737-015-0593-1] [Medline: 26687691]
5. Molenaar NM, Bais B, Lambregtse-van den Berg MP, Mulder CL, Howell EA, Fox NS, et al. The international prevalence of antidepressant use before, during, and after pregnancy: A systematic review and meta-analysis of timing, type of prescriptions and geographical variability. J Affect Disord 2020 Mar 01;264:82-89. [doi: 10.1016/j.jad.2019.12.014] [Medline: 31846905]
6. O'Connor E, Senger CA, Henninger ML, Coppola E, Gaynes BN. Interventions to Prevent Perinatal Depression: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA 2019 Dec 12;321(6):588-601. [doi: 10.1001/jama.2018.20865] [Medline: 30747970]
7. Brown HK, Ray JG, Wilton AS, Lunsky Y, Gomes T, Vigod SN. Association Between Serotonergic Antidepressant Use During Pregnancy and Autism Spectrum Disorder in Children. JAMA 2017 Apr 18;317(15):1544-1552. [doi: 10.1001/jama.2017.3415] [Medline: 28418480]
8. Goodnick PJ, Goldstein BJ. Selective serotonin reuptake inhibitors in affective disorders — I. Basic pharmacology. J Psychopharmacol Oxf Engl 1998;12(3 suppl B):5-20. [doi: 10.1177/0269881198012003021]
9. Holland J, Brown R. Neonatal venlafaxine discontinuation syndrome: A mini-review. Eur J Paediatr Neurol 2017 Mar;21(2):264-268. [doi: 10.1016/j.ejpn.2016.11.003] [Medline: 27931774]
10. Dubovicky M, Belovicova K, Csatlosova K, Bogi E. Risks of using SSRI / SNRI antidepressants during pregnancy and lactation. Interdiscip Toxicol 2017;10(1):30-34. [doi: 10.1515/intox-2017-0004]
11. Newport DJ, Calamaras MR, DeVane CL, Donovan J, Beach AJ, Winn S, et al. Atypical antipsychotic administration during late pregnancy: placental passage and obstetrical outcomes. Am J Psychiatry 2007 Aug;164(8):1214-1220. [doi: 10.1176/appi.ajp.2007.06111886] [Medline: 17671284]
12. Dalmizrak O, Kulaksiz-Erkmen G, Ozer N. The inhibition characteristics of human placental glutathione S-transferase-π by tricyclic antidepressants: amitriptyline and clomipramine. Mol Cell Biochem 2011 Sep 13;355(1-2):223-231. [doi: 10.1007/s11010-011-0858-6] [Medline: 21567209]
13. Oberlander TF. Fetal serotonin signaling: setting pathways for early childhood development and behavior. J Adolesc Health 2012 Aug;51(2 Suppl):S9-16. [doi: 10.1016/j.jadohealth.2012.04.009] [Medline: 22794534]
14. de Montigny C, Chaput Y, Blier P. Modification of serotonergic neuron properties by long-term treatment with serotonin reuptake blockers. J Clin Psychiatry 1990 Dec;51 Suppl B:4-8. [Medline: 2175309]
15. Borue X, Chen J, Condron BG. Developmental effects of SSRIs: lessons learned from animal studies. Int J Dev Neurosci 2007 Oct 07;25(6):341-347 [FREE Full text] [doi: 10.1016/j.ijdevneu.2007.06.003] [Medline: 17706396]
16. Rotem-Kohavi N, Williams LJ, Oberlander TF. Advanced neuroimaging: A window into the neural correlates of fetal programming related to prenatal exposure to maternal depression and SSRIs. Semin Perinatol 2020 Apr;44(3):151223. [doi: 10.1016/j.semperi.2020.151223] [Medline: 32122645]
17. Ornoy A, Koren G. SSRIs and SNRIs (SRI) in Pregnancy: Effects on the Course of Pregnancy and the Offspring: How Far Are We from Having All the Answers? Int J Mol Sci 2019 May 14;20(10):2370 [FREE Full text] [doi: 10.3390/ijms20102370] [Medline: 31091646]
18. Shirtcliff EA, Essex MJ. Concurrent and longitudinal associations of basal and diurnal cortisol with mental health symptoms in early adolescence. Dev Psychobiol 2008 Nov;50(7):690-703. [doi: 10.1002/dev.20336]

19. Gunnar M, Quevedo K. The neurobiology of stress and development. Annu Rev Psychol 2007 Jan;58(1):145-173. [doi: 10.1146/annurev.psych.58.110405.085605] [Medline: 16903808]
20. Gunnar MR, Vazquez DM. Low cortisol and a flattening of expected daytime rhythm: potential indices of risk in human development. Dev Psychopathol 2001 Sep 27;13(3):515-538. [doi: 10.1017/s0954579401003066] [Medline: 11523846]
21. Eisenberg N, Spinrad TL, Eggum ND. Emotion-related self-regulation and its relation to children's maladjustment. Annu Rev Clin Psychol 2010 Mar 01;6(1):495-525 [FREE Full text] [doi: 10.1146/annurev.clinpsy.121208.131208] [Medline: 20192797]
22. Azmitia E. Serotonin Neurons, Neuroplasticity, and Homeostasis of Neural Tissue. Neuropsychopharmacology 1999 Aug;21(2):33S-45S. [doi: 10.1016/s0893-133x(99)00022-6]
23. Whitaker-Azmitia PM, Druse M, Walker P, Lauder JM. Serotonin as a developmental signal. Behavioural Brain Research 1995 Dec;73(1-2):19-29. [doi: 10.1016/0166-4328(96)00071-x]
24. Huether G, Thömke F, Adler L. Administration of tryptophan-enriched diets to pregnant rats retards the development of the serotonergic system in their offspring. Developmental Brain Research 1992 Aug;68(2):175-181. [doi: 10.1016/0165-3806(92)90059-6]
25. Di Pino G, Moessner R, Lesch K, Lauder J, Persico A. Roles for Serotonin in Neurodevelopment: More than just Neural Transmission. CN 2004 Oct 01;2(4):403-417. [doi: 10.2174/1570159043359495]
26. Brummelte S, Mc Glanaghy E, Bonnin A, Oberlander T. Developmental changes in serotonin signaling: Implications for early brain function, behavior and adaptation. Neuroscience 2017 Feb 07;342:212-231 [FREE Full text] [doi: 10.1016/j.neuroscience.2016.02.037] [Medline: 26905950]
27. Houwing DJ, Buwalda B, van der Zee EA, de Boer SF, Olivier JDA. The Serotonin Transporter and Early Life Stress: Translational Perspectives. Front Cell Neurosci 2017 Apr 26;11:117 [FREE Full text] [doi: 10.3389/fncel.2017.00117] [Medline: 28491024]
28. St-Pierre J, Laurent L, King S, Vaillancourt C. Effects of prenatal maternal stress on serotonin and fetal development. Placenta 2016 Dec;48:S66-S71. [doi: 10.1016/j.placenta.2015.11.013]
29. Côté F, Fligny C, Fromes Y, Mallet J, Vodjdani G. Recent advances in understanding serotonin regulation of cardiovascular function. Trends Mol Med 2004 May;10(5):232-238. [doi: 10.1016/j.molmed.2004.03.007] [Medline: 15121050]
30. Kermorgant M, Pavy-Le TA, Senard J, Arvanitis D. Serotonergic Receptor 5-HT2A in the Cardiosympathovagal System. In: Guiard BP, Giovanni G, editors. 5-HT2A Receptors in the Central Nervous System. Switzerland: Springer International Publishing; 2018:137-145.
31. Fisher LA, Brown MR. 3 Central regulation of stress responses: regulation of the autonomic nervous system and visceral function by corticotrophin releasing factor-41. Baillière's Clinical Endocrinology and Metabolism 1991 Mar;5(1):35-50. [doi: 10.1016/s0950-351x(05)80095-3]
32. Chaouloff F, Berton O, Mormède P. Serotonin and Stress. Neuropsychopharmacology 1999 Aug;21(2):28S-32S. [doi: 10.1016/s0893-133x(99)00008-1]
33. Mitchell JB, Iny LJ, Meaney MJ. The role of serotonin in the development and environmental regulation of type II corticosteroid receptor binding in rat hippocampus. Developmental Brain Research 1990 Sep;55(2):231-235. [doi: 10.1016/0165-3806(90)90204-c]
34. Chaouloff F. Physiopharmacological interactions between stress hormones and central serotonergic systems. Brain Research Reviews 1993 Jan;18(1):1-32. [doi: 10.1016/0165-0173(93)90005-k]
35. Everly G, Lating J. The Anatomy and Physiology of the Human Stress Response. In: A Clinical Guide to the Treatment of the Human Stress Response. New York, NY: Springer; 2019:19-56.
36. Lv J, Liu F. The Role of Serotonin beyond the Central Nervous System during Embryogenesis. Front Cell Neurosci 2017 Mar 13;11:74 [FREE Full text] [doi: 10.3389/fncel.2017.00074] [Medline: 28348520]
37. Kapoor A, Dunn E, Kostaki A, Andrews M, Matthews SG. Fetal programming of hypothalamo-pituitary-adrenal function: prenatal stress and glucocorticoids. J Physiol 2006 Apr 01;572(Pt 1):31-44 [FREE Full text] [doi: 10.1113/jphysiol.2006.105254] [Medline: 16469780]
38. Glover V, O'Connor TG, O'Donnell K. Prenatal stress and the programming of the HPA axis. Neurosci Biobehav Rev 2010 Sep;35(1):17-22. [doi: 10.1016/j.neubiorev.2009.11.008] [Medline: 19914282]
39. Ramage AG. Central cardiovascular regulation and 5-hydroxytryptamine receptors. Brain Research Bulletin 2001 Nov;56(5):425-439. [doi: 10.1016/s0361-9230(01)00612-8]
40. Laplante P, Diorio J, Meaney MJ. Serotonin regulates hippocampal glucocorticoid receptor expression via a 5-HT7 receptor. Developmental Brain Research 2002 Dec;139(2):199-203. [doi: 10.1016/s0165-3806(02)00550-3]
41. Kofman O. The role of prenatal stress in the etiology of developmental behavioural disorders. Neuroscience & Biobehavioral Reviews 2002 Jun;26(4):457-470. [doi: 10.1016/s0149-7634(02)00015-5]
42. Weinstock M. Alterations induced by gestational stress in brain morphology and behaviour of the offspring. Progress in Neurobiology 2001 Dec;65(5):427-451. [doi: 10.1016/s0301-0082(01)00018-1]
43. Francis D, Diorio J, LaPlante P, Weaver S, Seckl JR, Meaney MJ. The role of early environmental events in regulating neuroendocrine development. Moms, pups, stress, and glucocorticoid receptors. Ann N Y Acad Sci 1996 Sep 20;794:136-152. [doi: 10.1111/j.1749-6632.1996.tb32517.x] [Medline: 8853600]

44. Rayen I, Gemmel M, Pauley G, Steinbusch HWM, Pawluski JL. Developmental exposure to SSRIs, in addition to maternal stress, has long-term sex-dependent effects on hippocampal plasticity. Psychopharmacology (Berl) 2015 Apr 12;232(7):1231-1244. [doi: 10.1007/s00213-014-3758-0] [Medline: 25304865]
45. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009 Jul 21;339:b2700-b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
46. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 01;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
47. Schroeder R. Pointing Users Toward Citation Searching: Using Google Scholar and Web of Science. portal: Libraries and the Academy 2007;7(2):243-248. [doi: 10.1353/pla.2007.0022]
48. Babineau J. Product Review: Covidence (Systematic Review Software). J Can Health Libr Assoc 2014 Aug 01;35(2):68-71. [doi: 10.5596/c14-016]
49. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010 Sep;25(9):603-605. [doi: 10.1007/s10654-010-9491-z] [Medline: 20652370]
50. Herzog R, Álvarez-Pasquin MJ, Díaz C, Del BJL, Estrada JM, Gil Á. Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. BMC Public Health 2013 Feb 19;13:154 [FREE Full text] [doi: 10.1186/1471-2458-13-154] [Medline: 23421987]
51. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
52. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol 2001 Oct;54(10):1046-1055. [doi: 10.1016/s0895-4356(01)00377-8] [Medline: 11576817]
53. Shi L, Lin L. The trim-and-fill method for publication bias: practical guidelines and recommendations based on a large database of meta-analyses. Medicine (Baltimore) 2019 Jun;98(23):e15987 [FREE Full text] [doi: 10.1097/MD.0000000000015987] [Medline: 31169736]
54. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Analysing data and undertaking meta-analyses. In: Cochrane Handbook for Systematic Reviews of Interventions. Chichester (UK): John Wiley & Sons; 2019:241-284.
55. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev 2012 Feb 09;1:2 [FREE Full text] [doi: 10.1186/2046-4053-1-2] [Medline: 22587842]

Abbreviations

5-HT: 5-hydroxytryptamine (serotonin)
5-HTT: serotonin transporter
ANS: autonomic nervous system
CRP: C-reactive protein
HPA: hypothalamic-pituitary-adrenal
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
PROSPERO: International Prospective Register of Systematic Reviews
SNRI: serotonin–norepinephrine reuptake inhibitor
SSRI: selective serotonin reuptake inhibitor

Edited by T Leung; submitted 03.09.21; peer-reviewed by J Hwang, C Weerth; comments to author 17.01.22; revised version received 19.01.22; accepted 26.01.22; published 28.03.22.

Please cite as:
Zusman EZ, Lavu A, Pawliuk C, Pawluski J, Hutchison SM, Platt RW, Oberlander TF
Associations Between Prenatal Exposure to Serotonergic Medications and Biobehavioral Stress Regulation: Protocol for a Systematic Review and Meta-analysis
JMIR Res Protoc 2022;11(3):e33363
URL: https://www.researchprotocols.org/2022/3/e33363
doi:10.2196/33363
PMID:35343913

Protocol

Associations Between Behavior Change Techniques and Engagement With Mobile Health Apps: Protocol for a Systematic Review

Madison Milne-Ives[1], BAS, MSc; Sophie Homer[2], MPsych, PGCAP, PhD; Jackie Andrade[2], BA, PhD; Edward Meinert[1,3,4], MA, MSc, MBA, MPA, PhD

[1]Centre for Health Technology, University of Plymouth, Plymouth, United Kingdom

[2]School of Psychology, Faculty of Health, University of Plymouth, Plymouth, United Kingdom

[3]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

[4]Harvard T H Chan School of Public Health, Harvard University, Boston, MA, United States

Corresponding Author:
Madison Milne-Ives, BAS, MSc
Centre for Health Technology
University of Plymouth
6 Kirkby Place, Room 2
Plymouth, PL4 6DN
United Kingdom
Phone: 44 1752600600
Email: madison.milne-ives@plymouth.ac.uk

Abstract

Background: Digitally enabled care along with an emphasis on self-management of health is steadily growing. Mobile health apps provide a promising means of supporting health behavior change; however, engagement with them is often poor and evidence of their impact on health outcomes is lacking. As engagement is a key prerequisite to health behavior change, it is essential to understand how engagement with mobile health apps and their target health behaviors can be better supported. Although the importance of engagement is emphasized strongly in the literature, the understanding of how different components of engagement are associated with specific techniques that aim to change behaviors is lacking.

Objective: The purpose of this systematic review protocol is to provide a synthesis of the associations between various behavior change techniques (BCTs) and the different components and measures of engagement with mobile health apps.

Methods: The review protocol was structured using the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) and the PICOS (Population, Intervention, Comparator, Outcome, and Study type) frameworks. The following seven databases will be systematically searched: PubMed, Embase, Cumulative Index to Nursing and Allied Health Literature, APA PsycInfo, ScienceDirect, Cochrane Library, and Web of Science. Title and abstract screening, full-text review, and data extraction will be conducted by 2 independent reviewers. Data will be extracted into a predetermined form, any disagreements in screening or data extraction will be discussed, and a third reviewer will be consulted if consensus cannot be reached. Risk of bias will be assessed using the Cochrane Collaboration Risk of Bias 2 and the Risk Of Bias In Non-Randomized Studies - of Interventions (ROBINS-I) tools; descriptive and thematic analyses will be conducted to summarize the relationships between BCTs and the different components of engagement.

Results: The systematic review has not yet started. It is expected to be completed and submitted for publication by May 2022.

Conclusions: This systematic review will summarize the associations between different BCTs and various components and measures of engagement with mobile health apps. This will help identify areas where further research is needed to examine BCTs that could potentially support effective engagement and help inform the design and evaluation of future mobile health apps.

Trial Registration: PROSPERO CRD42022312596; https://tinyurl.com/nhzp8223

International Registered Report Identifier (IRRID): PRR1-10.2196/35172

(JMIR Res Protoc 2022;11(3):e35172) doi:10.2196/35172

KEYWORDS

engagement; behavior change techniques; telemedicine; mobile apps

Introduction

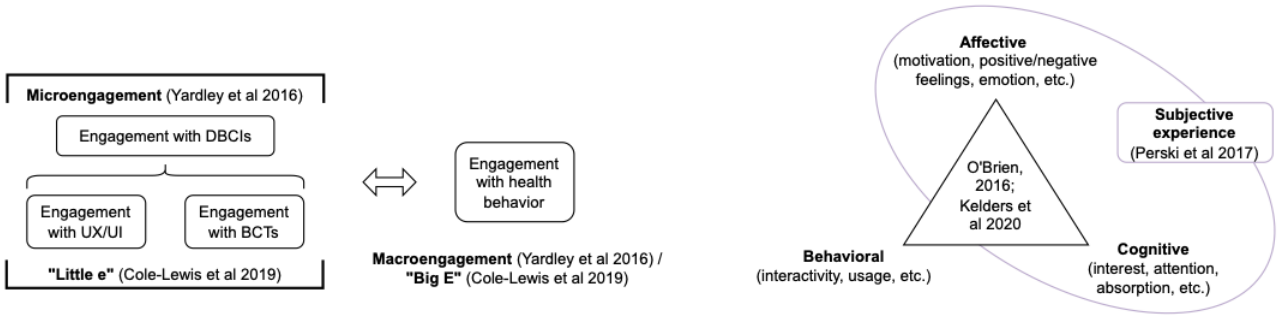
Background

This systematic review aims to provide an overview of how behavior change techniques (BCTs) [1] are associated with different components of engagement with mobile health apps. Effective engagement with digital health interventions is an essential factor influencing their ability to support positive behavior change. Although several models and frameworks conceptualizing engagement and its association with intervention impact have recently been published, a comprehensive understanding of how to develop digital health interventions that significantly impact health behavior and outcomes is still lacking [2]. This is a serious concern because although mobile health apps are frequently used to deliver health behavior change interventions [3], there is still a lack of evidence supporting their impact on behavior and health outcomes [4,5]. This lack of evidence necessitates an in-depth examination of the stages of engagement and behavior change so that particular barriers and blockers can be targeted. BCTs, "observable, replicable, and irreducible components" of behavior change interventions [1], provide a means of reliably classifying and testing potential strategies for altering behavior to address particular barriers. Understanding the associations between different BCTs, theoretical components of engagement, and measures of engagement will provide insight into how BCTs can be incorporated to improve and personalize the design of digital health interventions to support effective engagement.

Engagement with digital health interventions can be poor, which limits their potential impact. As health care service delivery is becoming increasingly digital and accessible through personal devices like smartphones and wearables [6,7], there is a need to ensure that these digital interventions are achieving their intended outcomes. The potential impact of digital interventions is limited by the extent of users' engagement with them [8-10]; a meta-analysis of engagement with digital mental health interventions found a significant positive association between engagement and mental health outcomes [11]. However, the variety in the definitions and measures of engagement means that reliable quantitative estimates of the relationship between engagement and outcomes are still lacking [8,11]. Maintaining engagement with digital health interventions is a common challenge. Studies on engagement with mobile health apps and wearable devices often observe poor long-term use [9,12,13] and high rates of attrition [8,14]. Although the duration of use is a commonly used indicator of engagement with a digital health intervention, its validity has been questioned because it only captures 1 component of engagement [2,10].

Inconsistency in the way engagement is defined and measured is one of the challenges associated with studying engagement [15,16]. The lack of a clear, comprehensive, and well-accepted conceptualization of engagement is a major gap, which several papers and reviews have recently tried to address [2,15,17,18]. Although various models and definitions of engagement have been proposed, there is a general consensus that engagement is a multifaceted concept [15-17,19,20]. These conceptual frameworks highlight the importance of considering cognitive, behavioral, and affective aspects of engagement [10,15,17,20], as well as examining different levels of engagement with digital behavior change interventions (DBCIs) and health behaviors [16,19] (see Figure 1). A key review defined engagement in terms of 2 key components, extent of usage and subjective experience [17]. Another paper emphasizes the importance of the relationship between engagement with the intervention and the target behavior by defining "effective engagement" as the level of engagement sufficient to achieve the aims of the intervention [21]. This highlights the crucial distinction between engagement with the intervention and engagement with the behavior, as frequent or indefinite engagement with the intervention may not be required to support sustained engagement with the behavior, as shown in Figure 1.

Figure 1. Summary of key theoretical concepts of engagement with digital health [15-17,20,21]. BCT: behavior change technique; DBCI: digital behavior change intervention; UI: user interface; UX: user experience.



Engagement with the intervention can be subdivided into engagement with the device or software and engagement with BCTs or "active ingredients" of the intervention [16,22] (see Figure 1, far left). As the DBCI is the proposed trigger for the behavior change, engagement with the health behavior is thought to depend on engagement with the DBCI [16,21]. However, the interconnected nature of engagement with the device, BCTs, and behavior makes it challenging to untangle the relationships between various stages and components of engagement and different BCTs. This is because BCTs can be used to influence users' engagement with the health behavior, for example, by including goal setting (BCT 1.1) or self-monitoring of behavior

(BCT 2.3) features to support users' engagement with physical activity. However, BCTs can also provide "feedback" to influence engagement with devices or with other BCTs included in the DBCI, for example, by using prompts or cues (BCT 7.1) such as app notifications to remind a user to engage with the app or with specific BCT-based features on the app.

Different BCTs are associated with different theoretical barriers to behavior (eg, capability, opportunity, and motivation) [23,24]. For instance, "instruction on how to perform the behavior" (BCT 4.1) is commonly used to support a "training" intervention function, which in turn can target barriers related to physical and psychological capability [24]. Given the different functions associated with BCTs, it seems likely that different BCTs will also have different relationships with the 3 main components of engagement (affective, cognitive, and behavioral). To improve engagement with DBCIs and target behaviors, it is essential to understand the relationships between BCTs and the various components of engagement and incorporate them into the design and evaluation of digital health interventions.

Rationale

The growing recognition of the importance of engagement in the design and evaluation of digital health interventions has led to an exponential increase in research concerning that topic in recent years. Given the accepted importance of engagement as a prerequisite for behavior change [17], several systematic reviews have examined various factors that could influence engagement with digital health interventions [25-28]. Among these, the analysis in 1 review [25] is structured around the COM-B (Capability, Opportunity, Motivation – Behavior) model, which is part of the Behavior Change Wheel theoretical framework [23]. The authors identified 26 different factors relating to capability, opportunity, and motivation that have been associated with uptake of and engagement with mobile health apps in the literature [25]. This provides a valuable, theory-based contribution to the understanding of factors affecting engagement with mobile health apps. However, despite including studies with either qualitative or quantitative (primarily system use data) measures of engagement and using a multifaceted definition of engagement [17], the review did not clarify how the influence of these factors varied for the different components and measures of engagement.

The importance of understanding the factors associated with engagement lies in their potential to inform designs that improve "effective engagement" with DBCIs and thereby better support behavior change and the associated positive health outcomes. Because engagement is a complex and multifaceted concept, it is important to understand how specific BCTs are related to different elements of engagement and which ones have the most influence on effective engagement and health outcomes [21]. As the best strategies for achieving effective engagement could differ among individuals, an understanding of how different BCTs are associated with different components of engagement would enable digital health interventions to be personalized to individuals, specific populations, or contexts, providing an opportunity to increase their health impact.

PROSPERO was searched using various combinations of the following keywords: engagement, digital health interventions, DBCIs, behavior change techniques, BCT, mobile health apps, mHealth, eHealth, and digital behavior change. None of the registered protocols aimed to examine the associations between BCTs and the different components of engagement; however, the search terms identified the PROSPERO preregistration for one of the previous reviews cited in this rationale [25], indicating that the search terms were appropriate.

Objectives

The main aim of the review is to provide a synthesis of the associations between BCTs and the different components of engagement (and their outcome measures) with mobile health apps in the literature. The following are the key objectives of this review: (1) to identify the BCTs being incorporated in the development of mobile health apps; (2) to identify the components of engagement that are being evaluated in studies on mobile health apps and how the different components are being measured; (3) to document the associations between specific BCTs and engagement component outcomes and outcome measures; and (4) to compare those associations across the included studies to hypothesize causal relationships between specific BCTs and specific components of engagement that can be empirically evaluated in future studies.

Methods

Overview

The PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) [29] and the PICOS (Population, Intervention, Comparator, Outcome, Study type) frameworks [30,31] will be used to structure this review and develop the search strategy. The PRISMA-P checklist is available in Multimedia Appendix 1. This review is registered on PROSPERO (registration number: CRD42022312596).

Eligibility Criteria

The PICOS framework is based on the research questions and is presented in Table 1.

Table 1. PICOS (Population, Intervention, Comparator, Outcome, and Study type) framework.

Framework component	Description
Population	Mobile health app users of any age (adults and children)
Intervention	Mobile health apps that explicitly use BCTs[a] in their design to target at least 1 of 5 key health categories established in the literature, including drug use, alcohol use, diet, physical activity, and mental health
Comparator	No comparator is required.
Outcomes	The primary outcome will be the qualitative or quantitative engagement outcomes measured (including any components of engagement specified by a theoretical framework). Secondary outcomes will include the BCTs included in the mobile health app, the measure(s) of engagement used by the study, and the behavioral and health outcomes reported.
Study types	Studies that evaluate engagement with at least 1 mobile health app that uses BCTs will be eligible (including randomized controlled trials, quantitative, qualitative, cohort, and case studies). Reviews, protocols, papers that describe interventions without evaluating them, and papers where full texts cannot be identified (eg, conference abstracts) will be excluded.

[a]BCT: behavior change technique.

Search Strategy

The search will be conducted in seven databases: PubMed, Embase, Cumulative Index to Nursing and Allied Health Literature, APA PsycInfo, ScienceDirect, Cochrane Library, and Web of Science. These databases were chosen because they were commonly searched in previous systematic reviews relating to engagement and digital health interventions, and they broadly cover topics related to digital technology, health, and behavior change. Keywords and MeSH (Medical Subject Headings) terms relating to engagement with digital health behavior change interventions were identified in an initial review of the literature and used to develop the search strategy. These search terms were expanded upon and grouped into three themes (see Table 2) to develop the following search structure: engagement (MeSH OR Keywords) AND mobile health apps (MeSH OR Keywords) AND behavior change (MeSH OR Keywords). Sample searches conducted in PubMed, Embase (Ovid), and Web of Science are included in Multimedia Appendix 2.

Table 2. Search terms.

Category	MeSH[a]	Keywords (in title or abstract)
Engagement	Treatment Adherence and Compliance OR Patient Participation OR Patient Compliance	Engagement OR adherence OR compliance OR maintenance OR acceptability OR satisfaction OR attention OR enjoyment OR interest OR affect OR flow OR "cognitive absorption" OR "subjective experience" OR immersion OR presence OR ((amount OR frequency OR duration OR depth OR breadth) NEAR/2 (use OR usage)) OR dose OR stickiness OR dropout OR "drop out" OR "drop-out" OR attrition
Mobile health apps	Telemedicine OR Mobile Applications	"mHealth" OR "mobile health" OR "eHealth" OR telehealth OR ((mobile OR phone OR smartphone OR cell OR mHealth OR "behavior change" OR "behavior change" OR digital) NEAR/2 (app OR apps OR application*))
Behavior change	Behavior Control	"behavior change techniques" or "behavior change techniques" or "BCT" or "behavior change technique" or "behavior change technique" or "behavioral change strategies" or "behavioral change strategies" or "behavior change wheel" or "behavior change wheel" or "behavioral theory" or "behavioral theory" or "behavior change theory" or "behavior change theory" or "health behavior change" or "behavior change" or "behavior change" or "digital behavior change intervention" or "digital behavior change intervention" or "DBCI" or "behavior change intervention"

[a]MeSH: Medical Subject Headings.

Inclusion Criteria

The review will include studies that evaluate theory-based mobile apps for health behavior change. Studies will be included if they evaluate at least 1 component or measure of engagement (quantitative or qualitative) with a mobile app that uses BCTs to influence health behavior. No restrictions will be placed on the type of health behavior or the sample population examined in the initial screening to ensure that all eligible studies are identified. If there are too many studies eligible after initial screening to conduct a thorough review, the number of studies will be restricted based on health behavior. This will limit included studies to those that focus on at least 1 of 5 key health categories, including drug use, alcohol use, diet, physical activity, and mental health [32,33], aligned with a previous review by the authors [5]. Studies with any type of sampled population will be eligible for inclusion, with no restrictions on age, gender, or country. Interventions with comparisons to control groups with no intervention, waiting list or irrelevant interventions, minimal interventions, usual care, other mobile apps, telemedicine, and internet-based or in-person interventions will be included. Studies with no comparators will also be included.

Exclusion Criteria

Studies involving mobile health apps that do not detail the BCTs included in the app will be excluded from the review. Studies that do not evaluate at least 1 measure of engagement, such as reviews, protocols, papers that describe interventions without evaluating them, and papers where full texts cannot be identified (eg, conference abstracts) will also be excluded.

Screening and Article Selection

The references returned by each database search will be exported into the citation management software EndNote X9 (Clarivate) so that duplicate references can be identified and removed. The screening will take place in three stages: (1) Keywords based on the search criteria will be entered into EndNote's search function over multiple passes to exclude any studies that are clearly ineligible (eg, protocols, reviews). (2) The titles and abstracts of the remaining references will be screened by 2 independent reviewers. (3) The full texts of the studies will be screened by 2 independent reviewers to determine the final set of included papers. Any disagreements between reviewers will be discussed until consensus; if consensus cannot be reached, a third reviewer will be consulted. Details of the screening and selection process will be recorded in a PRISMA flow diagram to ensure study reproducibility and the EndNote searches in stage 1 will be recorded and included in the review as an appendix.

Data Extraction

The full texts of all the articles included in the final set will be read by 2 independent reviewers to extract the required data mentioned in Table 3. As with the screening process, any disagreements will be discussed and resolved by involving a third reviewer if necessary.

Table 3. Article information and data extraction.

Article information	Data to be extracted
General study information	• Year of publication • Country of study • Sample demographics (including age, gender, target population) • Initial/intended sample size • Analyzed sample size • Study duration
Intervention	• App name • Operating platform (eg, iOS, Android) • Target health behavior • Specific aim of the intervention • Behavioral theory used in the design of the app (if any) • How the app was developed (eg, iterative design, experience-based co-design, etc) • Number of included behavior change techniques [1] • List of included behavior change techniques [1] • Intended purpose of included behavior change techniques (if specified) • Intended use (eg, dose and duration if specified)
Evaluation	• Component(s) of engagement examined • Engagement outcome measures • Effect of intervention on engagement outcomes (including engagement with specific behavior change techniques, the app, and the target health behavior) • Effect of intervention on behavior change outcomes • Effect of intervention on participant health outcomes

Quality Appraisal and Risk of Bias Assessment

The risk of bias of the studies will be evaluated by 2 independent reviewers using the Cochrane Collaboration Risk of Bias 2 tool for randomized controlled trials [34,35] and the Risk Of Bias In Non-Randomized Studies - of Interventions (ROBINS-I) tool for nonrandomized studies [36]. The GRADE (Grading of Recommendations, Assessment, Development and Evaluations) guidelines will be used to assess the strength of the body of evidence gathered during the review [37].

Data Analysis and Synthesis

The feasibility of conducting a meta-analysis will be examined when the data are extracted; however, a meta-analysis may not be possible owing to the expected variety of study aims, measures, and reported outcomes. The extracted data will be summarized by conducting a descriptive analysis to provide counts of the engagement components examined, outcome measures used, health behaviors targeted, and levels of evidence showing the effectiveness of BCTs for engagement, behavioral, and health outcomes. The associations between the inclusion of various BCTs and evidence of their effectiveness for various outcomes will be mapped. Any qualitative data reported will be examined by performing a thematic analysis to provide contextual data about the potential relationships between BCTs and certain components of engagement. The risk of bias in the studies will be considered in the synthesis.

Results

The full systematic review has not yet started, but it is expected to be completed and submitted for publication by May 2022.

Discussion

A systematic review of the literature on engagement with theoretically based mobile apps for health behavior change will contribute to the understanding of how BCTs fit into the multifaceted state and process of engagement. With the ubiquity of mobile health apps and the continuous growth of digitally enabled care [6], it is necessary to ensure that the mobile health apps being used are effective. A key component of the efficacy

of DBCIs is the extent to which the user engages effectively with the intervention to achieve the intended target behavior. An overview of the associations between BCTs and the different components and measures of engagement will inform the design and evaluation of mobile health apps. Based on the data, we will determine what conclusions can be drawn, identify the limitations of our systematic review, and propose key topics for future research.

Acknowledgments

This research was funded by Health Education England (grant AM1000393). The views expressed in the paper belong to the authors and not necessarily those of Health Education England or the University of Plymouth. The funding bodies were not involved in the study design, the writing, or the decision to submit the article for publication.

Authors' Contributions

The protocol was conceived and written by MMI with revisions from EM, SH, and JA.

Conflicts of Interest

EM is the Editor-in-Chief of JMIRx Med. The other authors report no conflicts of interest.

Multimedia Appendix 1
PRISMA-P (PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) checklist.
[DOC File , 97 KB - resprot_v11i3e35172_app1.doc]

Multimedia Appendix 2
Sample search strings.
[DOCX File , 19 KB - resprot_v11i3e35172_app2.docx]

References

1. 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]
2. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
3. Ferrara G, Kim J, Lin S, Hua J, Seto E. A focused review of smartphone diet-tracking apps: usability, functionality, coherence with behavior change theory, and comparative validity of nutrient intake and energy estimates. JMIR Mhealth Uhealth 2019 May;7(5):e9232 [FREE Full text] [doi: 10.2196/mhealth.9232] [Medline: 31102369]
4. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: a review of measurement, reporting, and validity. Psychiatr Serv 2019 Jul;70(7):538-544 [FREE Full text] [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
5. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. JMIR Mhealth Uhealth 2020 Mar;8(3):e17046 [FREE Full text] [doi: 10.2196/17046] [Medline: 32186518]
6. The NHS Long Term Plan.: NHS; 2019 Jun. URL: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/01/nhs-long-term-plan-june-2019.pdf [accessed 2021-10-18]
7. Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. Eur Heart J 2016 May;37(18):1428-1438 [FREE Full text] [doi: 10.1093/eurheartj/ehv770] [Medline: 26873093]
8. Yeager CM, Benight CC. If we build it, will they come? Issues of engagement with digital health interventions for trauma recovery. Mhealth 2018 Sep;4:37 [FREE Full text] [doi: 10.21037/mhealth.2018.08.04] [Medline: 30363749]
9. Grady A, Yoong S, Sutherland R, Lee H, Nathan N, Wolfenden L. Improving the public health impact of eHealth and mHealth interventions. Aust N Z J Public Health 2018 Apr;42(2):118-119. [doi: 10.1111/1753-6405.12771] [Medline: 29384248]
10. Torous J, Michalak EE, O'Brien HL. Digital health and engagement-looking behind the measures and methods. JAMA Netw Open 2020 Jul;3(7):e2010918 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10918] [Medline: 32678446]
11. Gan DZQ, McGillivray L, Han J, Christensen H, Torok M. Effect of engagement with digital interventions on mental health outcomes: a systematic review and meta-analysis. Front Digit Health 2021 Nov;3:764079 [FREE Full text] [doi: 10.3389/fdgth.2021.764079] [Medline: 34806079]
12. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep;21(9):e14567. [doi: 10.2196/14567]

13. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HWW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec;11(6):704-719 [FREE Full text] [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
14. Etminani K, Tao Engström A, Göransson C, Sant'Anna A, Nowaczyk S. How behavior change strategies are used to design digital interventions to improve medication adherence and blood pressure among patients with hypertension: systematic review. J Med Internet Res 2020 Apr;22(4):e17201 [FREE Full text] [doi: 10.2196/17201] [Medline: 32271148]
15. Kelders SM, van Zyl LE, Ludden GDS. The concept and components of engagement in different domains applied to eHealth: a systematic scoping review. Front Psychol 2020 May;11:926 [FREE Full text] [doi: 10.3389/fpsyg.2020.00926] [Medline: 32536888]
16. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions. JMIR Form Res 2019 Oct;3(4):e14052 [FREE Full text] [doi: 10.2196/14052] [Medline: 31603427]
17. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
18. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
19. Wannheden C, Stenfors T, Stenling A, von Thiele Schwarz U. Satisfied or frustrated? A qualitative analysis of need satisfying and need frustrating experiences of engaging with digital health technology in chronic care. Front Public Health 2021 Jan;8:623773 [FREE Full text] [doi: 10.3389/fpubh.2020.623773] [Medline: 33585391]
20. O'Brien H. Theoretical perspectives on user engagement. In: O'Brien H, Cairns P, editors. Why Engagement Matters. Cham: Springer International Publishing; 2016:1-26.
21. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
22. Hankonen N. Participants' enactment of behavior change techniques: a call for increased focus on what people do to manage their motivation and behavior. Health Psychol Rev 2021 Jun;15(2):185-194. [doi: 10.1080/17437199.2020.1814836] [Medline: 32967583]
23. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
24. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. Surrey: Silverback Publishing; 2014:1-329.
25. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020 May;22(5):e17572 [FREE Full text] [doi: 10.2196/17572] [Medline: 32348255]
26. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. J Med Internet Res 2020 Dec;22(12):e21687 [FREE Full text] [doi: 10.2196/21687] [Medline: 33295292]
27. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
28. Borghouts J, Eikey E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar;23(3):e24387 [FREE Full text] [doi: 10.2196/24387] [Medline: 33759801]
29. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015 Jan;350:g7647 [FREE Full text] [doi: 10.1136/bmj.g7647] [Medline: 25555855]
30. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club 1995 Nov;123(3):A-12. [Medline: 7582737]
31. Counsell C. Formulating questions and locating primary studies for inclusion in systematic reviews. Ann Intern Med 1997 Sep;127(5):380-387. [doi: 10.7326/0003-4819-127-5-199709010-00008] [Medline: 9273830]
32. Spring B, Moller AC, Coons MJ. Multiple health behaviours: overview and implications. J Public Health (Oxf) 2012 Mar;34(Suppl 1):i3-i10 [FREE Full text] [doi: 10.1093/pubmed/fdr111] [Medline: 22363028]
33. Conner M, Norman P. Health behaviour: current issues and challenges. Psychol Health 2017 Aug;32(8):895-906. [doi: 10.1080/08870446.2017.1336240] [Medline: 28612656]
34. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct;343:d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]

35. Risk of Bias 2 (RoB 2) tool. Cochrane Methods. URL: https://methods.cochrane.org/risk-bias-2 [accessed 2021-10-18]
36. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct;355:i4919 [FREE Full text] [doi: 10.1136/bmj.i4919] [Medline: 27733354]
37. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011 Apr;64(4):383-394. [doi: 10.1016/j.jclinepi.2010.04.026] [Medline: 21195583]

Abbreviations

BCT: behavior change technique
DBCI: digital behavior change intervention
MeSH: Medical Subject Headings
PICOS: Population, Intervention, Comparator, Outcome, and Study type
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
ROBINS-I: Risk Of Bias In Non-Randomized Studies - of Interventions

Edited by G Eysenbach; submitted 24.11.21; peer-reviewed by K Pfisterer, F Velayati; comments to author 09.12.21; revised version received 16.12.21; accepted 18.12.21; published 29.03.22.

Please cite as:
Milne-Ives M, Homer S, Andrade J, Meinert E
Associations Between Behavior Change Techniques and Engagement With Mobile Health Apps: Protocol for a Systematic Review
JMIR Res Protoc 2022;11(3):e35172
URL: https://www.researchprotocols.org/2022/3/e35172
doi:10.2196/35172
PMID:35348460

Protocol

PriSUD-Nordic—Diagnosing and Treating Substance Use Disorders in the Prison Population: Protocol for a Mixed Methods Study

Anne Bukten[1,2], PhD; Nicoline Toresen Lokdam[1,2], MA; Ingeborg Skjærvø[1,2], PhD; Thomas Ugelvik[3], PhD; Svetlana Skurtveit[1,4], PhD; Roman Gabrhelík[5], PhD; Torbjørn Skardhamar[6], PhD; Ingunn Olea Lund[4,7], PhD; Ingrid Amalia Havnes[2,8], PhD; Eline Borger Rognli[2], PhD; Zheng Chang[9], PhD; Seena Fazel[10], PhD; Christine Friestad[2,11], PhD; Morten Hesse[12], PhD; Johan Lothe[13], BA; Gerhard Ploeg[14], PhD; Anja J E Dirkzwager[15], PhD; Thomas Clausen[1], MD, PhD; Christian Tjagvad[1], PhD; Marianne Riksheim Stavseth[1,2], PhD

[1]Norwegian Centre for Addiction Research, University of Oslo, Oslo, Norway
[2]Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway
[3]Department of Criminology and Sociology of Law, University of Oslo, Oslo, Norway
[4]Department of Mental Disorders, Division of Mental and Physical Health, The Norwegian Institute of Public Health, Oslo, Norway
[5]Department of Addictology, First Faculty of Medicine, Charles University, Prague, Czech Republic
[6]Department of Sociology and Human Geography, University of Oslo, Oslo, Norway
[7]Department of Psychology, University of Oslo, Oslo, Norway
[8]Adult Psychiatry Unit, University of Oslo, Oslo, Norway
[9]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[10]Department of Psychiatry, University of Oxford, Oxford, United Kingdom
[11]University College of Norwegian Correctional Service, Lillestrøm, Norway
[12]Center for Alcohol and Drug Research, Aarhus University, Aarhus, Denmark
[13]WayBack, Foundation for Life After Imprisonment, Oslo, Norway
[14]Directorate of Norwegian Correctional Service, Lillestrøm, Norway
[15]Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam, Netherlands

Corresponding Author:
Anne Bukten, PhD
Norwegian Centre for Addiction Research
University of Oslo
Kirkeveien 166
Oslo, 0407
Norway
Phone: 47 23368971
Email: anne.bukten@medisin.uio.no

Abstract

Background: A large proportion of the prison population experiences substance use disorders (SUDs), which are associated with poor physical and mental health, social marginalization, and economic disadvantage. Despite the global situation characterized by the incarceration of large numbers of people with SUD and the health problems associated with SUD, people in prison are underrepresented in public health research.

Objective: The overall objective of the PriSUD (Diagnosing and Treating Substance Use Disorders in Prison)-Nordic project is to develop new knowledge that will contribute to better mental and physical health, improved quality of life, and better life expectancies among people with SUD in prison.

Methods: PriSUD-Nordic is based on a multidisciplinary mixed method approach, including the methodological perspectives of both quantitative and qualitative methods. The qualitative part includes ethnographic fieldwork and semistructured interviews. The quantitative part is a registry-based cohort study including national registry data from Norway, Denmark, and Sweden. The national prison cohorts will comprise approximately 500,000 individuals and include all people imprisoned in Norway, Sweden,

and Demark during the period from 2000 to 2019. The project will investigate the prison population during three different time periods: before imprisonment, during imprisonment, and after release.

Results: PriSUD-Nordic was funded by The Research Council of Norway in December 2019, and funding started in 2020. Data collection is ongoing and will be completed in the first quarter of 2022. Data will be analyzed in spring 2022 and the results will be disseminated in 2022-2023. The PriSUD-Nordic project has formal ethical approval related to all work packages.

Conclusions: PriSUD-Nordic will be the first research project to investigate the epidemiology and the lived experiences of people with SUD in the Nordic prison population. Successful research in this field will have the potential to identify significant areas of benefit and will have important implications for ongoing policy related to interventions for SUD in the prison population.

International Registered Report Identifier (IRRID): DERR1-10.2196/35182

(JMIR Res Protoc 2022;11(3):e35182) doi:10.2196/35182

KEYWORDS

substance use disorders; prison; criminal justice; epidemiology; mixed methods; harm reduction; treatment

Introduction

Globally, more than 30 million people are released from prisons each year [1], and the number is increasing [2,3]. Among people in prison, a large proportion have a history of drug use and substance use disorders (SUDs) [4,5]. A systematic review and meta-analysis found that the pooled prevalence estimate for SUD was 51% among women and 30% among men [4].

The harm caused by SUD, including the use of illicit substances as well as alcohol and other legal substances, is a significant contributor to the burden of disease [6]. Individuals with SUDs have a higher risk of premature death, ranging from 4-fold increased mortality among persons with alcohol use disorder [7] to 15-fold increased mortality among persons with opioid use disorder [8]. During 2015, 28 million years of healthy life (ie, disability-adjusted life years) were lost worldwide due to premature death and disability caused by drug use [7], with a heavier burden among socially disadvantaged groups, such as the prison population. People with SUD in prison thus constitute a group of people who are marginalized in terms of both their substance use and their incarceration, and they suffer disproportionately from poor physical and mental health, infectious diseases, social marginalization, and economic disadvantage [9].

A continuing challenge in public health is to provide services to the people who need them the most and who are the hardest to reach. Prisons may provide a unique opportunity for health interventions: a high proportion of people in prison have untreated SUD and, in prison, they are reachable for a predictable amount of time. The detection of mental health problems and SUDs, followed by adequate treatment and the introduction of harm reduction measures, may, from a public health perspective, represent a turning point in promoting SUD treatment in a highly disadvantaged group. However, the provision of health services in prison, worldwide, is characterized by large variations within a spectrum ranging from no health services to universal health coverage, with the Nordic countries being examples of the latter [10,11]. Where high-quality health services are offered to prisoners, prison is one of the few settings where health service agencies can engage in regular contact with marginalized populations that typically have precarious lifestyles when not imprisoned.

Advancing our knowledge of traditionally marginalized and understudied groups, such as people with SUD in prison, is essential to understanding social disparities in health. In addition, this is a precondition for planning the most appropriate interventions among people with SUD in prison. According to the World Health Organization (WHO), the public health importance of imprisonment is insufficiently recognized [12]. Despite the global situation characterized by incarceration of large numbers of people with SUD and the health problems associated with SUD, people in prison are underrepresented in public health research.

To address this challenge, the PriSUD (Diagnosing and Treating Substance Use Disorders in Prison)-Nordic project aims to investigate the epidemiology of SUD and explore the lived experiences of people with SUD in the Nordic prison population during three different periods: the time *before* imprisonment, the time *during* imprisonment, and the time *after* release.

To reach the aims of the PriSUD-Nordic study, we have put together a multidisciplinary research group that will analyze a wide range of existing, high-quality Nordic registry data combined with analysis of qualitative data based on ethnographic fieldwork and semistructured interviews. All the scientific aims have interlinking health and methodological perspectives as well as work packages (WPs). The WPs are related to (1) epidemiology (WP I), (2) risk assessment (WP II), and (3) qualitative methods (WP III; Table 1).

Table 1. Scientific aims, interlinked health and methodological perspectives, and work packages (WPs) for the periods prior to, during, and after imprisonment.

Time period and aim	Health perspective	Methodological perspective	WP[a]
Prior to imprisonment			
Investigate the prevalence of SUD[b] among the prison population before incarceration in Scandinavia	Determine the burden of SUD-related problems and selection of persons with SUD who are in prison	Does registry data capture the prevalence of SUD across the Scandinavian countries?	I
Investigate the prevalence of mental and physical health problems before incarceration among the prison population	Determine the burden of mental and physical health problems and selection of persons with such problems who are in prison	Does registry data capture the prevalence of mental and physical health problems across the Scandinavian countries?	I
During imprisonment			
Test a risk assessment tool based on OxRec[c] for identification of persons with SUD in need of treatment interventions	Identify persons in need of treatment based on routinely collected data	Improving risk assessment by testing and externally validating the OxRec instrument in the Nordic countries	II
Investigate and compare the availability of treatment interventions during incarceration for SUD and non-SUD populations	Determine if adequate treatment is offered in line with health policies and whether the availability of treatment for persons with SUD (first aim) differs from persons without SUD	Is registry data a valid source when estimating treatment availability in prison?	I
Investigate the nature, norms, mechanisms, and process of SUD treatment in Scandinavian prisons	Determine how SUD treatment is implemented and integrated into the prison context	How can qualitative analysis complement and support register data on health in prison research?	III
Postimprisonment			
Investigate and compare postrelease outcomes	Determine effects of in-prison treatment on postrelease outcomes regarding health, social welfare, and recidivism	How can variables from different countries be aligned to fully compare outcomes?	I
Investigate postrelease narratives among former prisoners	Describe challenges and opportunities for persons released from prison as perceived by participants in WayBack (a nonprofit foundation that helps prisoners return to society)	How can qualitative prison research guide, improve, and support implementation of health interventions in the postrelease context?	III
Test a risk assessment tool based on OxRec for identification of persons with a need for additional follow-up postrelease	Identify persons with a high risk of substance use postrelease to prevent relapse to substance use, crime, overdose, and other high-risk events	Improving risk assessment by testing and externally validating the OxRec instrument in the Nordic countries	II

[a]WP I is related to epidemiology, WP II is related to risk assessment, and WP III is related to qualitative methods.

[b]SUD: substance use disorder.

[c]OxRec: Oxford Risk of Recidivism Tool.

Methods

Overview

Prisons are situated in complex social, cultural, and political contexts and are dependent on social, structural, and historical factors. The experience of imprisonment is shaped by the characteristics of the individual prisoners as well as relational, structural, and regime factors. This makes a multidisciplinary mixed methods approach relevant, as it includes the methodological perspectives of epidemiological quantitative methods and ethnographic qualitative methods. Mixed methods enable investigators to integrate qualitative research and qualitative data conceptually and analytically with traditional epidemiological and quantitative research methods to facilitate translation. Mixed methods will help us understand not just whether an intervention works but how, why, for whom, and under what circumstances it works [13].

Epidemiological Approach (WP I)

A large part of the project will be based on data from nationwide public registries. This approach has several advantages: low costs, it covers the entire population, and it provides longitudinal data with controllable attrition [14-16]. The methodology—linking data through personal identification numbers to construct rich longitudinal data sets—is an important feature of PriSUD-Nordic. Typically, individual data are collected either in clinical studies, with limited follow-up time to measure posttreatment outcomes, or from registries only, excluding patient-reported outcomes. Treating these data sources as complimentary will harness the strengths of both. When performing multinational studies, problems concerning different data custodians are common: ethical and cross-jurisdictional data-sharing restrictions prevent the direct sharing of individual-level data. To account for these issues, PriSUD-Nordic will use a two-step, individual participant data meta-analysis (IPDMA) [17], described in detail in the Data Analysis Plan section.

Risk Assessment (WP II)

Big data and machine learning are hot topics in health and medical research [18]. With increasing amounts of data available, new opportunities for statistical analysis arise. One promising method is using risk assessment tools, often referred to in medicine as prognostic models, prediction models, prediction rules, or risk scores [19]. Using data that are routinely collected among the population in question allows risk assessments tools that are used as adjuncts to be employed as practical and easy-to-use guides; these can assist clinicians, other health personnel, or prison staff in decision-making, raising the ceiling of expertise, and potentially enabling more evidence-based approaches to delineate treatment pathways.

We intend to complement the epidemiological research with a translational approach in which we test an evidence-based and scalable risk assessment tool based on the Oxford Risk of Recidivism Tool (OxRec), which was developed by the Forensic Psychiatry and Psychology group at the University of Oxford [20]. This tool requires new external validations in Nordic samples. We will test the performance of this prediction tool in new cohorts of released prisoners and consider whether it needs recalibration. In addition, we will examine how such a tool, if the external validation is promising, can be translated into practice and, in particular, whether it can be used to identify prisoners at risk of reoffending who need additional substance use treatment or other interventions, such as more regular follow-up or links with community health services.

Qualitative Approach (WP III)

Qualitative health research methods are underused in public health research [21]. It is well established in penological research that prisons provide an environment often characterized by a series of lacks, including the lack of predictability, autonomy, and purpose [22,23]. When investigating the epidemiology of SUD among people in prison using longitudinal registry data, it is, therefore, highly relevant to take the specific context of the prison into account. Two methods will be at the center of the qualitative research: ethnographic fieldwork and semistructured interviews. The focus will be on identifying challenges to successful treatment in prison and on the prison as a specific arena for treatment interventions.

Qualitative methods can provide background and depth to epidemiological and statistical analysis, which may be helpful when developing hypotheses and research design. Qualitative analysis can also help validate or challenge the interpretation of quantitative results by developing knowledge of processes, mechanisms, and explanatory models behind the results. In addition, qualitative methods can inform the development and implementation of interventions, guidelines, and recommendations that may result from the research project. Finally, qualitative approaches will provide increased insight into the combination of, or conflict between, health and welfare-oriented goals (ie, treatment, increased health, and well-being) on the one hand, and penal goals in prison settings (ie, punishment, control, security, and retribution) on the other hand.

Research Opportunity in Nordic Prison Settings

The Nordic countries are in an ideal position to perform world-class quality research on substance-related public health challenges. This is partly because all the Nordic countries collect individual-level data in the form of various national registries, including rich health and social services data. The countries have publicly financed health care, available to those who need it regardless of their financial situation. The similarities in societal development across the Nordic countries makes the Nordic region ideal for comparative studies within health and SUD. By investigating postrelease outcomes regarding health, social welfare, and recidivism, the output from PriSUD-Nordic will help us understand what characterizes best-practice interventions.

The Nordic countries have enormous potential for synergy, with strong health care registers, publicly owned universities and university hospitals, and a high appreciation for medical research among the public and politicians. In the recent Nordic White Paper on Medical Research, national registries were identified as a specific area where coordinated actions and determined cooperation could bring the Nordic region into a unique, global leadership position [24]. Because of all the benefits, Norwegian and Nordic research councils are currently promoting the use of national registries [24]. PriSUD-Nordic will use the full potential of these registries by applying a methodology designed specifically for multinational studies.

Study Population and Data Sources

The Scandinavian prison cohorts will include approximately 500,000 individuals. The Norwegian prison cohort (approximately 100,000 individuals) includes all people in Norwegian prisons during the period from 2000 to 2019. The Danish prison cohort (approximately 250,000 individuals) includes people released from prison or on probation during the period from 2000 to 2012. The Swedish cohort (approximately 150,000 individuals) includes all people released from prison or on probation during the period from 2000 to 2013. All the national cohorts are drawn from the national prison registries and will be linked to (1) national cause of death registries, (2) national prescription databases (excluding the Norwegian cohort), (3) national patient registries, (4) police and crime registries, and (5) data on socioeconomic conditions. See Table 2 for a more detailed description of the national registries.

Table 2. Types of registries and data and their descriptions.

Type of registry or data	Description[a]
National prison registries	National prison registries include a range of personal data on people imprisoned in each country, including age, gender, convictions and sentences, and the actual time spent in prison. The registries also include some information on participation in correctional interventions.
National cause of death registries	National cause of death registries include the cause of death based on the ICD-10[b]. The registries include all residents of the country at the time of death. The registries are based on death certificates and information that is coded at a national level.
National prescription databases	National prescription databases contain information on all prescription drugs, whether reimbursed or not, dispensed by pharmacies to individual patients. All drugs are classified according to the Anatomical Therapeutic Chemical classification system.
National patient registries	National patient registries include information on all patients receiving hospital-level care in both inpatient and outpatient facilities and acute and emergency services for mental and somatic illnesses. The national patient registries also include birth date, county of residence, date of admission, date of discharge, and primary and secondary diagnoses, according to the ICD-10.
Police and crime registries	Police and crime registries include information on all registered criminal cases, including identified offenders. They provide data on several prosecuting decisions, such as formal charges leading to conviction, formal charges leading to acquittal, and fines. In addition, these registries contain dates for offences and convictions.
Data on socioeconomic conditions	Data on socioeconomic conditions include information on employment, income, and social benefits.

[a]The registry descriptions are intended as an overview and may vary across countries.

[b]ICD-10: International Classification of Diseases, Tenth Revision.

Data Analysis Plan

Evidence from randomized controlled trials is rarely available for studies on the treatment of SUDs, as the randomized controlled trial design has ethical challenges related to harmful substances. Therefore, most studies on SUDs are observational. In such studies, associations between exposure and outcome can be explained by true causation, reverse causation, or confounding. Methods to support causal inference in observational studies are required. The PriSUD-Nordic project will adopt newly developed causal inference methods, such as nearly saturated propensity score matching, instrumental variable methods, and inverse probability weighting [25].

The quantitative part of the study is a linked registry-based cohort study. All data will be analyzed using the two-step IPDMA method. Methodologically, it allows for consistent inclusion and exclusion criteria across countries and has been successfully applied to other studies [17,26]. This method has several advantages. Because data are initially analyzed locally, we remove hurdles associated with sharing sensitive individual-level data across local jurisdictions. Thus, local investigators are provided with an opportunity to convey the nuances of data. It allows for consistent adjustment for confounding factors that may explain differences in findings across countries; in addition, it increases the clinical relevance of findings by providing the opportunity to explore clinical questions that cannot be answered by the individual countries alone [17].

The interviews will be conducted following a semistructured interview guide, based on a scoping review of the current knowledge within the field. Some of the overall themes in the interview guide may include social relations, support and marginalization, health and addiction, and motivation for treatment. A combination of thematic and inductively developed analytic codes will be employed. The focus will be on identifying challenges to successful treatment and the prison as an arena for treatment interventions.

Ethics Approval

People in prison form a disadvantaged group in a coerced setting that carries a heavy burden of problems. This requires increased awareness of ethical boundaries by investigators and research personnel and highlights the need for scientific knowledge about the group and their situation. No prisoners will be denied treatment or experience a reduction in the quality of treatment as a result of the research project. All efforts to ensure that data are treated confidentially and in accordance with existing legislation for research data will be followed, and the project will be conducted according to the Declaration of Helsinki.

WP I and WP II will be based on registry data from Norway, Sweden, and Denmark. The Norwegian registry linkage has been approved by the Regional Committees for Medical and Health Research Ethics (REC ID 2012/1401, 29513), the Norwegian Centre for Research Data (NSD ID 847562), and the Data Protection Officer at the Faculty of Medicine at the University of Oslo. The Swedish data linkage has been approved by the Regional Ethical Review Board in Stockholm (Dnr 2013/862–31/5) and has already been obtained. In Denmark, registry data can be used for research without ethical approval. To overcome the obstacles associated with sharing individual-level data across jurisdictions, data will be analyzed separately in each country. All data will be stored according to local regulations. All dissemination resulting from the study will contain group-based information. Thus, no individual participants will be identifiable.

The national registry studies (WP I and WP II) have been approved for exemption from the consent requirement. WP III is based on ethnographic fieldwork and semistructured

interviews. This part of the project has been approved by the Norwegian Centre for Research Data (NSD ID 964221). Informed written consent must be obtained from all participants.

User Involvement

User involvement will be integrated into the steering committee and the research group, which is represented by clinicians, policy makers, and user organizations. The Norwegian user organization WayBack [27] provides services in prison and helps build prosocial networks postrelease to prevent relapse into crime and drug use. WayBack will play an essential role in planning the project, discussing research questions for the papers in the project, and recruiting participants to interview (WP III). WayBack will also be involved in our plans to disseminate the study results to participants and the relevant wider communities (eg, choosing what information and results to share, when to share them, and in what format).

Results

PriSUD-Nordic was funded by The Research Council of Norway in December 2019, and funding began in 2020. Data collection is ongoing and will be completed in the first quarter of 2022. Data will be analyzed in spring 2022 and the results will be disseminated in 2022-2023.

Discussion

The Potential Impact of the Proposed Research

If our project proves to be successful, the knowledge and outputs generated from this project can provide new insight in order to solve challenges related to the United Nations' Sustainable Development Goals. For instance, Goal 3 (Good Health And Well-being) includes Target 3.5 (Prevent and Treat Substance Abuse), which is to "strengthen the prevention and treatment of substance abuse, including narcotic drug abuse and harmful use of alcohol."

An essential part of PriSUD-Nordic is developing organizational and methodological novelty to run a multi-organizational, multidisciplinary, and multinational research study. PriSUD-Nordic aims to develop new knowledge that contributes to better mental and physical health, improved quality of life, and better life expectancies among people with SUD in prison. More specifically, we anticipate that the aims outlined in the proposal will impact (1) establishment of a new knowledge base, (2) identification of treatment gaps and potential discrimination in treatment, and (3) provision of best-practice interventions.

Establishing a New Knowledge Base

To prevent or promote anything effectively, we first need to *identify* the risk factors and characteristics associated with a higher likelihood of specific outcomes. In the case of harmful substance use as an outcome, prevention must consider that risk factors related to SUD vary according to *different social and political contexts* and *over the individual life course.* Preventing further development of SUD requires interventions to meet the target group's needs and consider vulnerable transitions in life, such as cycling in and out of prison.

Although several studies have described the prison population before entering prison [28,29], the complexity of SUD, along with other health-related problems and lifestyle factors, has not yet been addressed. There is an increasing recognition that lifestyle factors can be modified to improve health outcomes directly [30]. However, while many countries have established and implemented policies and interventions that address smoking, alcohol, nutrition, and physical activity among nonclinical groups, these lifestyle factors have been addressed less frequently among the prison population. Within these groups, unmet needs must be identified, and user-acceptable and user-accessible interventions must be further developed and implemented. The output from PriSUD-Nordic will be a new and improved knowledge base for fitting best-practice interventions for the prison population.

Identifying Treatment Gaps and Potential Discrimination in Treatment

Lives lost that are attributable to morbidity and mortality resulting from all causes of substance use have increased in the past decade [31], and data from the WHO suggest that only 7% of those with past-year SUDs received even minimally adequate treatment [32]. Among the prison population, the proportions are even smaller. This illustrates that mental health and the addiction field have lagged behind other areas of medicine in terms of resources for treatment and research, and the public health goal of reducing the world's drug problems cannot be achieved without addressing SUDs with the same scientific commitment with which physical problems are addressed [33].

However, underuse of treatment can also result from the extensive and deeply seated stigmatization of substance users. Despite having poor health, many do not seek or receive health services [34]. Others receive lower-quality services and are judged as least deserving of health care [35]. This aspect is highlighted in the United Nations General Assembly Special Session (UNGASS) resolution, which aims to eliminate stigma and discrimination toward individuals with SUDs [36]. By investigating the need and availability of treatment interventions during incarceration based on the preprison burden of disease, the output from PriSUD-Nordic could identify treatment gaps and potential discrimination in treatment offered.

Providing Best-Practice Interventions

According to the recent UNGASS resolution, a comprehensive public health approach should offer accessible, evidence-based prevention, treatment, and recovery options during and following incarceration [37]. According to a seminal Lancet report, receiving treatment in prison is a *human right* [38].

The Nordic countries are in an ideal position to perform world-class quality research on today's most pressing substance-related public health challenges. This is, in part, because all the Nordic countries collect individual-level data in the form of various national registries, including health and social services data. The Nordic similarities in terms of societal development makes the Nordic region ideal for comparative studies within health and SUD. Research across the Nordic countries and within Europe has a great advantage, as findings may be compared between nations. By investigating postrelease

outcomes regarding health, social welfare, and recidivism, the output from PriSUD-Nordic may contribute to the provision of best-practice interventions.

Acknowledgments

We would like to thank Ann Oldervoll for proofreading the manuscript. This work was supported by The Research Council of Norway (grant 301535).

Availability of Data and Materials

This population study will be based on individual-level data from the Norwegian Prison Registry, which is held by the Directorate of Norwegian Correctional Service, and the Norwegian Cause of Death Registry, which is held by the Norwegian Institute of Public Health. The ethical approval of this research project does not include permission to publicly share the raw data. Qualifying researchers can apply for access to relevant data through the Norwegian Institute of Public Health and the Directorate of Norwegian Correctional Service upon approval from the Regional Committees for Medical and Health Research Ethics.

Authors' Contributions

AB and MRS designed the PriSUD-Nordic protocol with substantial critical input from TU, NTL, and SF. AB drafted the study protocol, with critical input from all authors. Description of data sources and the study population and size were drafted by MRS and AB, with substantial input from MH and ZC. The statistical analyses section was drafted by MRS. The literature search was conducted by AB, with input from all authors. All authors were involved in reading and revising the paper. All authors critically read and approved the final version of the manuscript.

Conflicts of Interest

ZC and SF were part of the research team that developed the Oxford Risk of Recidivism (OxRec) tool.

Multimedia Appendix 1
Peer-review report.
[PDF File (Adobe PDF File), 23 KB - resprot_v11i3e35182_app1.pdf]

References

1. Policy Brief. HIV prevention, treatment and care in prisons and other closed settings: A comprehensive package of interventions. United Nations Office on Drugs and Crime. 2013 Jun. URL: https://www.unodc.org/documents/hiv-aids/HIV_comprehensive_package_prison_2013_eBook.pdf [accessed 2022-03-07]
2. Fazel S, Baillargeon J. The health of prisoners. Lancet 2011 Mar 12;377(9769):956-965. [doi: 10.1016/S0140-6736(10)61053-7] [Medline: 21093904]
3. Walmsley R. Word Prison Brief. World Prison Population List. 11th edition. London, UK: Institute for Criminal Policy Research; 2016. URL: https://www.prisonstudies.org/sites/default/files/resources/downloads/world_prison_population_list_11th_edition_0.pdf [accessed 2022-03-07]
4. Fazel S, Yoon IA, Hayes AJ. Substance use disorders in prisoners: An updated systematic review and meta-regression analysis in recently incarcerated men and women. Addiction 2017 Oct;112(10):1725-1739 [FREE Full text] [doi: 10.1111/add.13877] [Medline: 28543749]
5. United Nations Office on Drugs and Crime. World Drug Report 2019. New York, NY: United Nations; 2019 Jun. URL: https://www.un-ilibrary.org/content/books/9789210041744/read [accessed 2022-03-13]
6. Lim S, Vos T, Flaxman A, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
7. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: A meta-review. World Psychiatry 2014 Jun;13(2):153-160 [FREE Full text] [doi: 10.1002/wps.20128] [Medline: 24890068]
8. Degenhardt L, Bucello C, Mathers B, Briegleb C, Ali H, Hickman M, et al. Mortality among regular or dependent users of heroin and other opioids: A systematic review and meta-analysis of cohort studies. Addiction 2011 Jan;106(1):32-51. [doi: 10.1111/j.1360-0443.2010.03140.x] [Medline: 21054613]
9. Friestad C. Socio-economic status and health in a marginalized group: The role of subjective social status among prison inmates. Eur J Public Health 2010 Dec;20(6):653-658. [doi: 10.1093/eurpub/ckp242] [Medline: 20123684]
10. Pratt J. Scandinavian exceptionalism in an era of penal excess: Part I: The nature and roots of Scandinavian exceptionalism. Br J Criminol 2007 Dec 06;48(2):119-137. [doi: 10.1093/bjc/azm072]

11. Ugelvik T, Dullum J, editors. Penal Exceptionalism?: Nordic Prison Policy and Practice. New York, NY: Routledge; 2012:1-268.
12. van den Bergh B, Gatherer A, Fraser A, Moller L. Imprisonment and women's health: Concerns about gender sensitivity, human rights and public health. Bull World Health Organ 2011 Sep 1;89(9):689-694. [doi: 10.2471/blt.10.082842]
13. Rolfe DE, Ramsden VR, Banner D, Graham ID. Using qualitative health research methods to improve patient and public involvement and engagement in research. Res Involv Engagem 2018;4:49 [FREE Full text] [doi: 10.1186/s40900-018-0129-8] [Medline: 30564459]
14. Hovde Lyngstad T, Skardhamar T. Nordic register data and their untapped potential for criminological knowledge. Crime Justice 2011 Aug;40(1):613-645. [doi: 10.1086/658881]
15. Lund IO, Bukten A. Harm to others from substance use and abuse: The underused potential in nationwide registers. Subst Abus 2015 Oct 29;9(S2):33-38 [FREE Full text] [doi: 10.4137/sart.s23545]
16. Gabrhelík R, Handal M, Mravčík V, Nechanská B, Tjagvad C, Thylstrup B, et al. Opioid maintenance treatment in the Czech Republic, Norway and Denmark: A study protocol of a comparative registry linkage study. BMJ Open 2021 May 10;11(5):e047028 [FREE Full text] [doi: 10.1136/bmjopen-2020-047028] [Medline: 33972343]
17. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: Rationale, conduct, and reporting. BMJ 2010 Feb 05;340:c221. [doi: 10.1136/bmj.c221] [Medline: 20139215]
18. Krumholz HM. Big data and new knowledge in medicine: The thinking, training, and tools needed for a learning health system. Health Aff (Millwood) 2014 Jul;33(7):1163-1170 [FREE Full text] [doi: 10.1377/hlthaff.2014.0053] [Medline: 25006142]
19. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: What, why, and how? BMJ 2009 Feb 23;338:b375. [doi: 10.1136/bmj.b375] [Medline: 19237405]
20. Fazel S, Chang Z, Fanshawe T, Långström N, Lichtenstein P, Larsson H, et al. Prediction of violent reoffending on release from prison: Derivation and external validation of a scalable tool. Lancet Psychiatry 2016 Jun;3(6):535-543 [FREE Full text] [doi: 10.1016/S2215-0366(16)00103-6] [Medline: 27086134]
21. Isaacs A. An overview of qualitative research methodology for public health researchers. Int J Med Public Health 2014;4(4):318. [doi: 10.4103/2230-8598.144055]
22. Ugelvik T. Power and Resistance in Prison: Doing Time, Doing Freedom. New York, NY: Palgrave Macmillan; 2014.
23. Sykes GM. The Society of Captives: A Study of a Maximum Security Prison. Princeton, NJ: Princeton University Press; 1958.
24. Joint Nordic Registers and Biobanks: A Goldmine for Health and Welfare Research. NordForsk Policy Paper 5. Oslo, Norway: NordForsk; 2014. URL: http://norden.diva-portal.org/smash/get/diva2:785891/FULLTEXT02.pdf [accessed 2022-03-07]
25. Austin PC, Fine JP. Propensity-score matching with competing risks in survival analysis. Stat Med 2019 Feb 28;38(5):751-777 [FREE Full text] [doi: 10.1002/sim.8008] [Medline: 30347461]
26. Kivimäki M, Nyberg ST, Batty GD, Fransson EI, Heikkilä K, Alfredsson L, IPD-Work Consortium. Job strain as a risk factor for coronary heart disease: A collaborative meta-analysis of individual participant data. Lancet 2012 Oct 27;380(9852):1491-1497 [FREE Full text] [doi: 10.1016/S0140-6736(12)60994-5] [Medline: 22981903]
27. WayBack. URL: https://www.wayback.no/ [accessed 2022-03-13]
28. Pape H, Rossow I, Bukten A. Alcohol problems among prisoners: Subgroup variations, concurrent drug problems, and treatment needs. Eur Addict Res 2021;27(3):179-188. [doi: 10.1159/000511253] [Medline: 33321498]
29. Stewart AC, Cossar R, Dietze P, Armstrong G, Curtis M, Kinner SA, et al. Lifetime prevalence and correlates of self-harm and suicide attempts among male prisoners with histories of injecting drug use. Health Justice 2018 Oct 15;6(1):19 [FREE Full text] [doi: 10.1186/s40352-018-0077-2] [Medline: 30324532]
30. Kinner S, Young J. Understanding and improving the health of people who experience incarceration: An overview and synthesis. Epidemiol Rev 2018 Jun 01;40(1):4-11 [FREE Full text] [doi: 10.1093/epirev/mxx018] [Medline: 29860342]
31. United Nations Office on Drugs and Crime. World Drug Report 2017. Executive Summary Conclusions and Policy Implications. New York, NY: United Nations; 2017 May. URL: https://www.unodc.org/wdr2017/field/Booklet_1_EXSUM.pdf [accessed 2022-03-07]
32. Degenhardt L, Glantz M, Evans-Lacko S, Sadikova E, Sampson N, Thornicroft G, World Health Organization's World Mental Health Surveys Collaborators. Estimating treatment coverage for people with substance use disorders: An analysis of data from the World Mental Health Surveys. World Psychiatry 2017 Oct;16(3):299-307 [FREE Full text] [doi: 10.1002/wps.20457] [Medline: 28941090]
33. Volkow ND, Poznyak V, Saxena S, Gerra G, UNODC-WHO Informal International Scientific Network. Drug use disorders: Impact of a public health rather than a criminal justice approach. World Psychiatry 2017 Jun;16(2):213-214 [FREE Full text] [doi: 10.1002/wps.20428] [Medline: 28498572]
34. Meyer CL, Tangney JP, Stuewig J, Moore KE. Why do some jail inmates not engage in treatment and services? Int J Offender Ther Comp Criminol 2013 Jun 18;58(8):914-930. [doi: 10.1177/0306624x13489828]
35. Lloyd C. The stigmatization of problem drug users: A narrative literature review. Drugs 2012 Nov 29;20(2):85-95. [doi: 10.3109/09687637.2012.743506]

36. United Nations Office on Drugs and Crime. Outcome Document of the 2016 United Nations General Assembly Special Session on the World Drug Problem. New York, 19-21 April 2016: Our Joint Commitment to Effectively Addressing and Countering the World Drug Problem. New York, NY: United Nations; 2016 Jun. URL: https://www.unodc.org/documents/postungass2016/outcome/V1603301-E.pdf [accessed 2022-03-13]
37. Global Prison Trends. Special Focus 2015. Drugs and Imprisonment. London, UK: Penal Reform International; 2015. URL: https://www.unodc.org/documents/ungass2016/Contributions/Civil/PenalReform/Drugs_and_imprisonment_PRI_submission_UNGASS.pdf [accessed 2022-03-07]
38. Rubenstein LS, Amon JJ, McLemore M, Eba P, Dolan K, Lines R, et al. HIV, prisoners, and human rights. Lancet 2016 Sep 17;388(10050):1202-1214. [doi: 10.1016/S0140-6736(16)30663-8] [Medline: 27427457]

Abbreviations

IPDMA: individual participant data meta-analysis
OxRec: Oxford Risk of Recidivism Tool
PriSUD: Diagnosing and Treating Substance Use Disorders in Prison
SUD: substance use disorder
UNGASS: United Nations General Assembly Special Session
WHO: World Health Organization
WP: work package

Edited by T Derrick; submitted 30.11.21; this is a non–peer-reviewed article;accepted 06.12.21; published 23.03.22.

Please cite as:
Bukten A, Lokdam NT, Skjærvø I, Ugelvik T, Skurtveit S, Gabrhelík R, Skardhamar T, Lund IO, Havnes IA, Rognli EB, Chang Z, Fazel S, Friestad C, Hesse M, Lothe J, Ploeg G, Dirkzwager AJE, Clausen T, Tjagvad C, Stavseth MR
PriSUD-Nordic—Diagnosing and Treating Substance Use Disorders in the Prison Population: Protocol for a Mixed Methods Study
JMIR Res Protoc 2022;11(3):e35182
URL: https://www.researchprotocols.org/2022/3/e35182
doi:10.2196/35182
PMID:35320114

Protocol

Rehabilitation Using Mobile Health for Older Adults With Ischemic Heart Disease in the Home Setting (RESILIENT): Protocol for a Randomized Controlled Trial

John A Dodson[1], MPH, MD; Antoinette Schoenthaler[2], EdD; Greg Sweeney[3], DPT; Ana Fonceva[4], MS; Alicia Pierre[3], DPT; Jonathan Whiteson[5], MD; Barbara George[6], EdD; Kevin Marzo[7], MD; Wendy Drewes[8], BSN, MSN; Elizabeth Rerisi[8], MS; Reena Mathew[8], BSN, MSN; Haneen Aljayyousi[4], BA; Sarwat I Chaudhry[9], MD; Alexandra M Hajduk[10], MPH, PhD; Thomas M Gill[10], MD; Deborah Estrin[11], PhD; Lara Kovell[12], MD; Lee A Jennings[13], MSHS, MD; Samrachana Adhikari[14], PhD

[1]Geriatric Cardiology Program, Medicine and Population Health, Leon H Charney Division of Cardiology, NYU Grossman School of Medicine, New York, NY, United States

[2]Department of Population Health, Department of Medicine, NYU Grossman School of Medicine, New York, NY, United States

[3]Department of Rehabilitation Medicine, NYU Grossman School of Medicine, New York, NY, United States

[4]Leon H Charney Division of Cardiology, NYU Grossman School of Medicine, New York, NY, United States

[5]Department of Rehabilitation Medicine, Department of Medicine, NYU Grossman School of Medicine, New York, NY, United States

[6]Division of Cardiology, Department of Medicine, NYU Long Island School of Medicine, Mineola, NY, United States

[7]Department of Medicine, Division of Cardiology, NYU Long Island School of Medicine, Mineola, NY, United States

[8]Division of Cardiology, NYU Langone Hospital Long Island, Mineola, NY, United States

[9]Section of General Medicine, Yale University School of Medicine, New Haven, NY, United States

[10]Yale University School of Medicine, New Haven, CT, United States

[11]Cornell Tech and Weill Cornell Medicine, New York, NY, United States

[12]Department of Medicine, University of Massachusetts Medical School, UMass Memorial Medical Center, Worcester, MA, United States

[13]Reynolds Section of Geriatric Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, United States

[14]Department of Population Health, NYU Grossman School of Medicine, New York, NY, United States

Corresponding Author:
John A Dodson, MPH, MD
Geriatric Cardiology Program, Medicine and Population Health
Leon H Charney Division of Cardiology
NYU Grossman School of Medicine
Translational Research Building
227 East 30th Street
New York, NY, 10016
United States
Phone: 1 646 501 2714
Email: john.dodson@nyulangone.org

Abstract

Background: Participation in ambulatory cardiac rehabilitation remains low, especially among older adults. Although mobile health cardiac rehabilitation (mHealth-CR) provides a novel opportunity to deliver care, age-specific impairments may limit older adults' uptake, and efficacy data are currently lacking.

Objective: This study aims to describe the design of the rehabilitation using mobile health for older adults with ischemic heart disease in the home setting (RESILIENT) trial.

Methods: RESILIENT is a multicenter randomized clinical trial that is enrolling patients aged ≥65 years with ischemic heart disease in a 3:1 ratio to either an intervention (mHealth-CR) or control (usual care) arm, with a target sample size of 400 participants. mHealth-CR consists of a commercially available mobile health software platform coupled with weekly exercise therapist sessions to review progress and set new activity goals. The primary outcome is a change in functional mobility (6-minute walk distance),

which is measured at baseline and 3 months. Secondary outcomes are health status, goal attainment, hospital readmission, and mortality. Among intervention participants, engagement with the mHealth-CR platform will be analyzed to understand the characteristics that determine different patterns of use (eg, persistent high engagement and declining engagement).

Results: As of December 2021, the RESILIENT trial had enrolled 116 participants. Enrollment is projected to continue until October 2023. The trial results are expected to be reported in 2024.

Conclusions: The RESILIENT trial will generate important evidence about the efficacy of mHealth-CR among older adults in multiple domains and characteristics that determine the sustained use of mHealth-CR. These findings will help design future precision medicine approaches to mobile health implementation in older adults. This knowledge is especially important in light of the COVID-19 pandemic that has shifted much of health care to a remote, internet-based setting.

Trial Registration: ClinicalTrials.gov NCT03978130; https://clinicaltrials.gov/ct2/show/NCT03978130

International Registered Report Identifier (IRRID): DERR1-10.2196/32163

(JMIR Res Protoc 2022;11(3):e32163) doi:10.2196/32163

KEYWORDS

mobile health; cardiac rehabilitation; clinical trial; rehabilitation; cardiology; heart disease; ambulatory care; mHealth; health outcomes; older adults

Introduction

Among older adults with ischemic heart disease (IHD), participation in ambulatory cardiac rehabilitation (CR) remains low despite decades of evidence about its benefits. Recent estimates suggest that fewer than two-thirds of eligible patients are referred, and even among those referred, only half attend the first session [1-3]. In addition to barriers faced by the general population (eg, limited facilities, competing time demands, high out-of-pocket costs, and prolonged wait time for space), older adults face additional barriers including lack of transportation, physical limitations, and sensory impairments that make it especially difficult to use existing CR paradigms [4,5]. Therefore, although older adults may have the greatest potential to benefit from CR because of their higher risk of adverse IHD-related sequelae, they are also the least likely to participate [4,6,7].

Mobile health CR (mHealth-CR) for IHD, which involves the delivery of rehabilitation via portable electronic devices, has proliferated rapidly in recent years [5,8,9]. mHealth-CR programs differ but typically include exercise documentation, hemodynamic assessment, video education, and electronic communication with an exercise therapist; these may be standalone components or adjunct to traditional ambulatory CR programs [5,9]. Although mHealth-CR has the potential to increase engagement by reducing participation barriers, it remains largely untested among the older adult population. It is therefore unclear what proportion of older adults with IHD (who may benefit the most) are able to engage with mHealth-CR and whether mHealth-CR leads to better outcomes than usual care in this population. In this paper, we describe the rehabilitation using mobile health (mHealth) for older adults with IHD in the home setting trial, which we designed to address this knowledge gap.

Methods

Overview

The rehabilitation using mHealth for older adults with ischemic heart disease in the home setting (RESILIENT) trial (NCT03978130) is recruiting 400 participants with a hospital visit for IHD at 3 academic medical centers: the original 2 sites were New York University (NYU) Langone Health (New York, New York) and Yale New Haven Health (New Haven, Connecticut). The first participant was enrolled on September 1, 2020. A third site, University of Massachusetts (Worcester, Massachusetts), was added in March 2021 to enhance recruitment. For NYU Langone Health, participants are being enrolled at both the NYU Langone Medical Center (New York, New York) and the NYU Langone Hospital–Long Island Hospital (Mineola, New York). NYU Langone Health serves as the coordinating center for both study administration and data management. The primary objective of RESILIENT is to test whether mHealth-CR improves functional capacity, as measured by the 6-minute walk test (6MWT), compared with usual care. We hypothesize that 6MWD (6-minute walk distance) will show significant improvement among participants receiving the study intervention, compared with participants in the usual care arm. RESILIENT was designed using pragmatic trial principles [10], including broad eligibility, the use of existing staff (exercise therapists) to deliver the study intervention, and the inclusion of outcomes (eg, physical function, goal attainment, and quality of life) that have been repeatedly cited as important by older adults [11,12]. There are two study visits: the baseline visit will occur within 4 weeks of hospital discharge and the follow-up visit will occur 3 months after baseline. We chose 3 months for the duration of the study intervention to match the duration of typical CR programs and evaluate whether mHealth-CR promotes early functional recovery.

Eligibility Criteria

The phenotype of interest for RESILIENT is IHD, which is operationalized as a hospital visit for either acute myocardial infarction (AMI) or coronary revascularization (percutaneous

coronary intervention [PCI] or coronary artery bypass graft). We chose the hospital visit as the time of enrollment for two reasons: first, previous research has demonstrated that a serious medical illness or procedure can serve as a *motivational moment* for patients to adopt healthier lifestyles [13,14]; second, deconditioning often accompanies either hospital admission or procedural recovery [15]. As our focus is on understanding mHealth efficacy in older adults, only patients aged ≥65 years are eligible. The exclusion criteria (Textbox 1) were designed to minimize the risk of adverse events (eg, falls with exercise) and ensure that participants can comprehend the study intervention.

Textbox 1. Eligibility criteria for the rehabilitation at home using mobile health in older adults after hospitalization for ischemic heart disease trial.

Inclusion criteria

- Age ≥65 years
- Hospital visit for either acute myocardial infarction or coronary revascularization (percutaneous coronary intervention [PCI] or coronary artery bypass graft)

Exclusion criteria

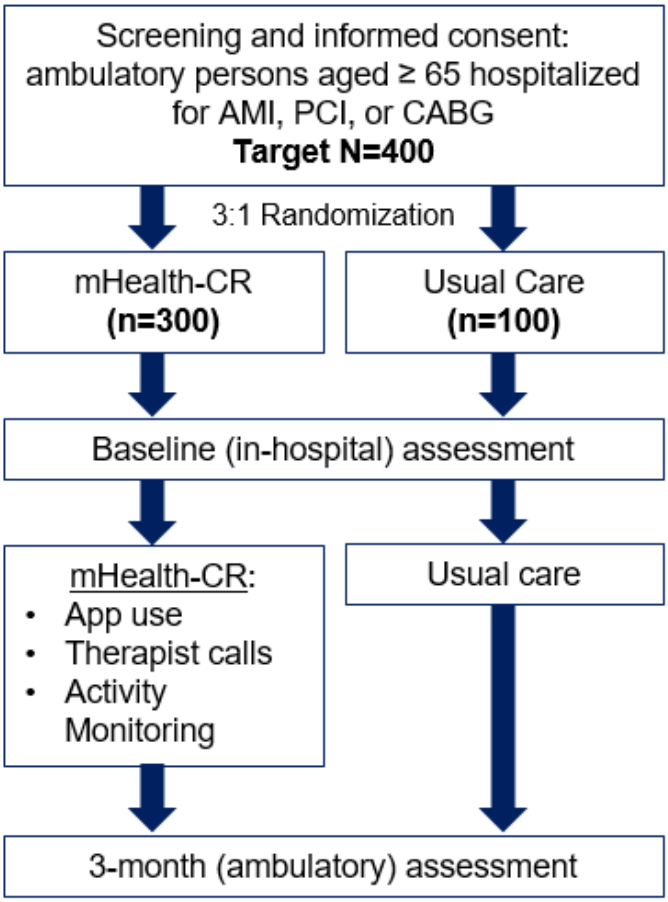
- Nonambulatory or regular use of walker for ambulation
- Moderate or severe cognitive impairment—defined as cognitive impairment that interferes with daily function
- Unable or unwilling to consent
- PCI-related groin hematoma that precludes brisk walking
- Incarcerated
- Unable to use mobile health software in English or Spanish
- Severe osteoarthritis or joint replacement within the last 3 months
- Parkinson disease or other progressive movement disorder
- Projected life expectancy of <3 months
- Clinical judgment concerning other safety or nonadherence issues
- Adverse event during the screening 6-minute walk test (drop in systolic blood pressure ≥15 mm Hg, chest pain, and ventricular arrhythmia)

Screening and Randomization

Participants are identified through daily electronic health record (EHR) screening of hospital lists with the index condition of AMI, elective PCI, or coronary artery bypass graft. Eligible cases are reviewed by a study physician (JAD, KM, SIC, LK) to ensure they meet the study criteria and are not a false-positive screen (eg, takotsubo cardiomyopathy may be screened by biomarkers, but this does not meet the phenotype of IHD). Potential participants are initially approached while still in hospital whenever possible or, if this is not feasible owing to off-hours procedures, within 48 hours of discharge. The baseline assessment is scheduled within 2 weeks of hospital discharge.

After informed consent and completion of the baseline assessment, including the 6MWT (performed in-person at the study site), enrolled participants are randomized through a function in REDCap (Research Electronic Data Capture; Vanderbilt University) [16] to either the intervention (mHealth-CR) or control (usual care) arm in a 3:1 ratio, using permuted block randomization with variable block sizes of 4 and 8 (Figure 1). The rationale for 3:1 randomization is based on trial efficiency; more participants are enrolled in mHealth-CR than in usual care to have an adequate sample size to understand daily engagement with the intervention. Randomization is stratified by study site to ensure balance across intervention and control groups, given the likely between-hospital population differences. The randomization code was created by the study statistician (SA), and randomization assignments are given by the study coordinator following the 6MWT completion.

Figure 1. Rehabilitation using mobile health for older adults with ischemic heart disease in the home setting (RESILIENT) study design overview. Participants will be randomized in a 3:1 manner to receive mobile health cardiac rehabilitation (mHealth-CR) versus usual care. A target of 400 participants will be enrolled to retain 320 with evaluable end points (accounting for 20% attrition between baseline and 3 months). AMI: acute myocardial infarction; CABG: coronary artery bypass graft; PCI: percutaneous coronary intervention.



Study Outcomes

Primary End Point (Efficacy)

The primary efficacy endpoint is the change in 6MWD, reflective of functional capacity, as measured by the 6MWT. The 6MWT is performed during the baseline visit and at the 3-month follow-up visit by a blinded clinical assessor (exercise therapist or research nurse). At baseline, blinding is maintained by randomization occurring after 6MWT completion in a separate space to ensure that the 6MWT assessor is not present. During follow-up, the randomization group is not disclosed to the assessor by the study staff, and the study documentation with this information is not accessible by the 6MWT assessor. The concept of walking to measure functional capacity was introduced by Balke [17] in 1963, and 6MWD has been used since the 1980s as a robust and reproducible outcome in patients with comorbid illness [18-20]. Among patients with IHD, 6MWD correlates with several clinically meaningful outcomes, including cardiovascular events [21], hospitalization [22], and death [21,23,24]. Changes in 6MWD frequently used by ambulatory CR programs as a measure of effectiveness in patients with IHD [25]. More broadly, changes in 6MWD have been used as an end point in trials focused on a range of cardiovascular conditions, including IHD, and are associated with outcomes including mortality and hospitalization [19,26-28]. The feasibility and safety of measuring 6MWD in patients with IHD, even during inpatient hospitalization, have been demonstrated by previous studies [29,30].

Secondary End Points (Efficacy)

There are 5 prespecified secondary efficacy end points.

Goal Attainment

Goal attainment is defined as whether a person's individual functional goals are achieved 3 months after hospital discharge and is measured using a 5-point goal attainment scale (GAS). Using the specific, measurable, achievable, realistic, and timely goal framework [11], the GAS describes the person's expected level of goal achievement over 3 months, ranging from no change (scored as −2) to much-better-than-expected change (scored as +2). Scales are dynamically set according to a person's needs, whereas the measurement of attainment is standardized. Goals are ascertained by study research coordinators, all of whom obtained structured training by a physician expert (LAJ) in goal attainment. Goal attainment, through goal-setting, is an especially important outcome in older adults who may begin an intervention with a variety of deficits (therefore necessitating individualized therapy toward realistic goals) [31].

Participant-Reported Health Status

Participant-reported health status will be measured using the 12-item Short Form Health Survey (general health status) questionnaire [32] and the Seattle Angina Questionnaire 7 (disease-specific health status) [33]. We have chosen these 2 instruments based on extensive validation and convenience of administration (<5 minutes). Changes between the 2 groups will be compared between baseline and 3 months.

Changes in Activities of Daily Living

Changes in activities of daily living (ADLs) are defined as any improvement or worsening in basic ADLs (BADLs) or instrumental ADLs (IADLs) over 3 months. BADLs are basic self-care behaviors, including feeding, toileting, bathing, dressing, transferring, and ambulating [34]. IADLs are activities that allow a person to live independently (eg, food preparation, medication management, transportation, shopping, managing finances, using the telephone, and housekeeping) [34].

Hospital Readmission

Hospital readmission is defined as an unanticipated overnight stay (including observation) in any hospital within 3 months of discharge. As these data are obtainable via the EHR, we also ascertain readmission events at 6 months and 1 year.

Death

Death is defined as death from any cause within 3 months of enrollment. Similar to readmission, we also ascertain death at 6 months and 1 year through the EHR. For both readmission and mortality, we acknowledge that we may not capture 6 months' or 1 year's events that occur outside of the study health systems, although external data linkages are becoming increasingly common.

Implementation End Point (Engagement)

We explicitly designed RESILIENT to enable the study of participant engagement with mHealth (an implementation end point) in addition to the efficacy end points. Our main measure of engagement is the weekly percent completion of the mHealth-CR program. Completion of mHealth-CR analyzed at weekly intervals allows us to determine distinct engagement trajectories throughout the 3-month study period. Weekly engagement is measured as the fraction of the following 11 elements completed each week: (1-7) daily entry of exercise data and relative perceived exertion (RPE); (8) completed weekly phone calls with exercise therapists; (9) at least one electronic communication with an exercise therapist; (10) watching an educational video (which varies by week); and (11) at least one home blood pressure (BP) measurement.

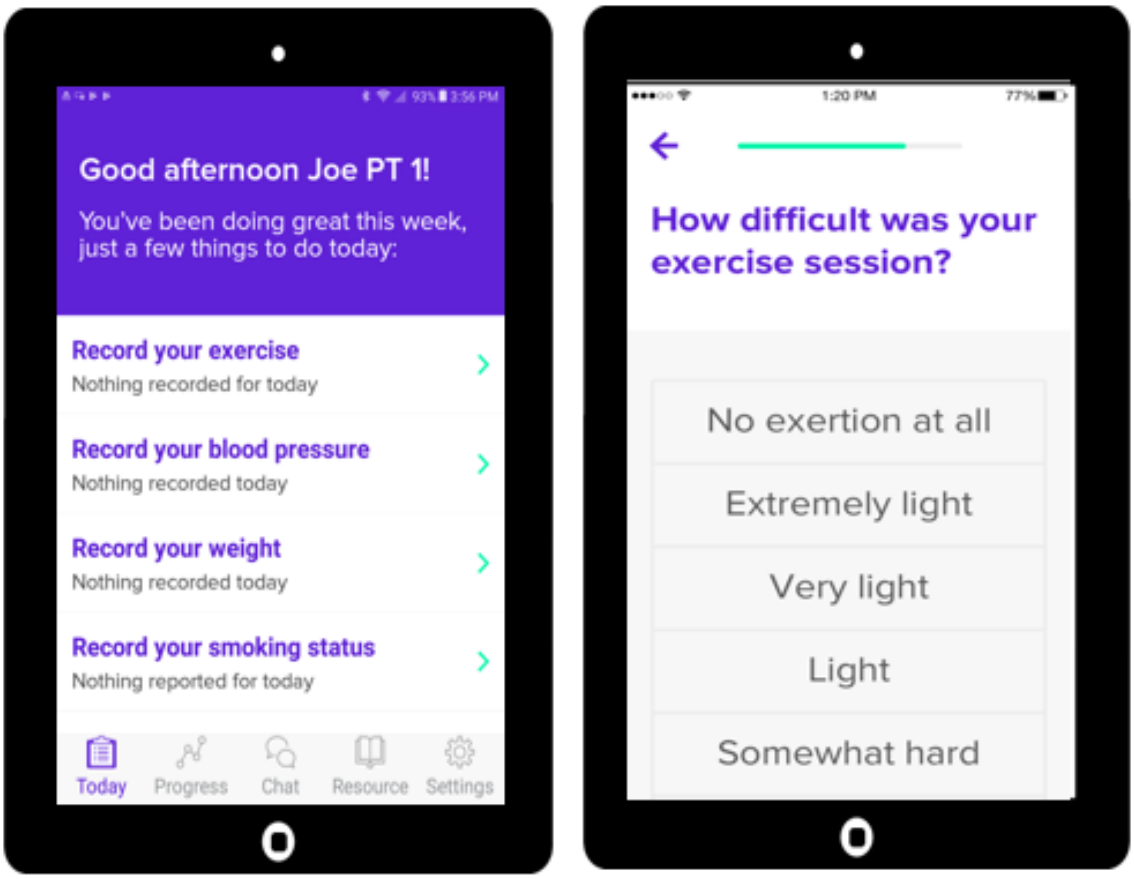
Study Intervention

Study participants randomized to the intervention (mHealth-CR) arm receive three components: (1) mHealth-CR software, (2) communication with an exercise therapist (in-hospital assessment or counseling followed by regular communication postdischarge), and (3) a wearable activity monitoring device. These components are designed to work in concert.

mHealth-CR Software

We have partnered with Moving Analytics, which has developed a commercial software platform to deliver mHealth-CR on portable electronic devices (Figure 2). To obviate barriers to portable electronic device ownership, participants will receive a tablet computer (Samsung Galaxy) with mHealth-CR software for the duration of the trial. Devices have cellular capability; therefore, home Wi-Fi access is not required. The software includes four components: (1) participant data entry of exercise duration and RPE; (2) a *chat* function where participants can communicate questions or concerns about symptoms to the exercise therapist (this is checked daily during weekdays); (3) weekly series of educational videos that are focused on secondary cardiovascular event prevention, addressing the following topics—introduction, understanding emotions, exercise guidelines, managing medications, having a heart-healthy diet, stress reduction, and smoking cessation; and (4) physiological measurement with an assessment of BP, heart rate, and physical activity. Participants are instructed by the exercise therapist to use the mHealth-CR app for at least 5 days each week. Training includes education on BP cuff use (cuff placement, seating position, and arm support). The same cuff (Omron HEM-9200T) is used for all the participants. Training also includes how to charge the tablet, how to enter the tablet, how to restart and switch the tablet on or off, how to close all windows, and how to return to the *home* page. All participants are also instructed to log-in to Moving Analytics with the research coordinator present and practice entering activity information. The software is available in English and Spanish. Engagement metrics (daily log-in, activity time on the app, number of exercise sessions logged, duration of exercise, and average weekly minutes of exercise) are captured as a component of the platform and can be analyzed retrospectively.

Figure 2. Study intervention. The intervention includes commercially available mobile health cardiac rehabilitation software from Moving Analytics where participants record their physical activity and self-rate difficulty of their exercise session. This is coupled with passive activity monitoring and weekly phone calls with an exercise therapist for a duration of 3 months.



Communication With the Exercise Therapist

The intervention arm participants meet the therapist immediately after the baseline interview. This visit includes education on cardiac risk factor management, assessment of baseline functional status, and an introduction to the mHealth-CR software. A personalized exercise program has been designed and includes alternating aerobic exercise (walking and stair climbing) and low-level isometric resistance training (upper body strength exercises using elastic bands). Participants are recommended to exercise for at least 5 out of 7 days per week, with an ideal goal of 150 minutes per week of moderate-intensity exercise [35]; however, recommendations are tailored based on functional limitations. The therapist identifies potential barriers to this plan and develops mitigation strategies. Given our target population (age ≥65 years), the therapist also assesses home safety during the baseline interview to engage the participant in removing fall hazards at home, and if there is an additional concern, they determine eligibility for a home physiotherapist evaluation after discharge; a similar safety assessment has been successfully implemented in the Strategies To Reduce Injuries and Develop Confidence in Elders study [36]. After this baseline visit, home exercise intensity is rated daily by participants on the mHealth-CR software, using RPE on the Borg Scale (target range 11-14) [37]. The exercise therapist then makes phone contact with the participants in week 1 and weekly (by phone) for the remainder of the study. Exercise recommendations are titrated during calls based on self-reported RPE and a review of activity data. As traditional ambulatory CR is currently the standard of care, participants will also receive a 1-page document containing information about traditional ambulatory CR, including the local facility phone number and recommendations to discuss with their cardiologist. In the interest of a pragmatic trial, referral to traditional ambulatory CR is not mandated for all participants but is left to the discretion of the treating inpatient or outpatient cardiologist (dependent on local practice patterns). Attendance to these programs is captured at 3 months by participant interviews and verified by an EHR review.

The intervention follows the United States' *Physical Activity Guidelines for Americans, second edition* [38], which has been endorsed by the American Heart Association. The specific guidance is as follows:

- At least 150 minutes of moderate-intensity exercise per week, or 75 minutes of vigorous-intensity aerobic physical activity, or an equivalent combination. Activity should be spread throughout the week (eg, 5 sessions of 30 minutes each).
- If 150 minutes of moderate-intensity activity is reached, participants will be encouraged to increase to 300 minutes of activity.

- Muscle-strengthening activities take place at least 2 days per week. For purposes of the RESILIENT trial, participants are provided with elastic resistance bands and trained on their use at the time of the baseline visit. They are trained to perform upper and lower body exercises using these bands. Each participant is provided with 3 levels of resistance bands to allow for progression and to be able to adjust resistance for the various exercises.
- For those unable to achieve at least 150 minutes of moderate-intensity exercise (eg, owing to functional limitations), lower exercise targets are adapted as endorsed by the guidelines. The principle of move more, sit less will is also recommended.
- Balance training is also incorporated into the treatment plan, as recommended by the guidelines, as they pertain to older adults.

Exercise therapists for the trial have at least a master's level training in exercise therapy. As clinically trained professionals, they may adapt the intervention based on individual study participants' physical or sensory limitations or specific rehabilitation needs. This concept is similar to traditional rehabilitation, in which the intervention is individualized.

Wearable Activity Monitoring Device

Participants are provided with a Fitbit Inspire or Fitbit Inspire 2 wearable wrist device (Fitbit Inc). This is a commercially available product that measures physical activity based on the number of steps per day. Activity is categorized (based on step count) as sedentary or mildly, moderately, or vigorously active. Heart rate information is also collected by the Inspire 2 model, which was adopted after the first 26 participants. Data are automatically uploaded daily to Moving Analytics and are viewable by both the study participant and the exercise therapist. Weekly phone calls with the exercise therapist include a review of activity data, including the percentage of time spent for each category and total daily step count.

Usual Care

In accordance with current guidelines [2], participants in both study arms receive information about ambulatory CR at the time of hospital discharge, with the phone number of the program at their respective hospital and guidance to discuss this option with their cardiologist. Therefore, receipt of the mHealth-CR intervention does not preclude participants from attending traditional ambulatory CR. However, in the interest of a pragmatic clinical trial, we do not provide additional incentives (eg, transportation to ambulatory CR) that may further reduce traditional barriers. Similarly, we do not mandate referral to ambulatory CR at the time of hospital discharge because this does not reflect current care patterns at study sites (in current practice, a referral is performed by the outpatient cardiologist). Attendance at traditional CR is captured at the 3-month follow-up visit.

Study Visits

All participants undergo a baseline visit and a 3-month ambulatory visit, led by a research coordinator who measures the elements listed in Table 1. The baseline visit, which lasts up to 2 hours, occurs either in the hospital or within 2 weeks of hospital discharge. The follow-up visit lasts up to 1 hour. Between visits, all participants undergo telephone assessments of BADLs and IADLs at 1 and 2 months to capture dynamic changes in these measures. Participants in the intervention arm also receive a baseline visit by an exercise therapist (up to 1 hour), regular phone calls (20 minutes), mHealth software, and a Fitbit activity monitoring device. Phone calls follow a structured template to ensure consistency of the study intervention. All study participants receive a US $25 ClinCard payment after completion of the baseline visit and an additional US $90 ClinCard payment after completion of the 3-month follow-up visit.

Table 1. Timeline for study participants.

Study arm	Baseline (in hospital)	Home activities	3 months (ambulatory)

Intervention and control arms[a]	• In-person assessment
	  • Demographics
	  • Height, weight, blood pressure
	  • 6MWT[b]
	  • Health status (SF-12[c] and SAQ-7[d])
	  • Activities of daily living
	  • Cognition (MOCA[e])
	  • Goal attainment scaling (GAS[f])
	  • Depression (PHQ-9[g])
	  • Frailty elements[h]
	• Chart abstraction
	  • Comorbidities, medications, and laboratory values
		• Monthly activities of daily living assessment	• In-person assessment
			  • Weight and blood pressure
			  • 6-6MWT
			  • Health status (SF-12 and SAQ-7)
			  • Activities of daily living
			  • Goal attainment scaling (GAS)
			  • Depression (PHQ-9)
			  • Frailty elements
			  • Hospital readmissions
			• Chart abstraction
			  • Hospital readmission (verification)[i] and attendance at traditional cardiac rehabilitation

Intervention arm	• Exercise therapist assessment
	  • Education on cardiac risk factor management
	  • Ascertainment of home environment and mobility barriers
	  • Introduction to the mobile health–cardiac rehabilitation software platform
	  • Personalized exercise plan
		• Daily therapist-directed activity (walking and upper extremity resistance training)
		• Daily mHealth data entry
		• Weekly therapist phone call (counseling or activity review)
		• Weekly video education
		• Weekly blood pressure
		• Fitbit activity tracking and review
			• System Usability Scale

[a]Intervention and control participants will also receive referral to traditional (ambulatory) cardiac rehabilitation at hospital discharge but not mandated or facilitated attendance of first visit. Usual first ambulatory cardiac rehabilitation visit at New York University and Yale takes place within 4 weeks.

[b]6MWT: 6-minute walk test (this will be performed by a blinded research nurse).

[c]SF-12: 12-item Short Form Health Survey.

[d]SAQ-7: Seattle Angina Questionnaire 7.

[e]MOCA: Montreal Cognitive Assessment.

[f]GAS: goal attainment scale.

[g]PHQ-9: Patient Health Questionnaire 9.

[h]On the basis of the 3/5 criteria: unintentional weight loss, weak grip strength (dynamometer), exhaustion, slow gait, and low physical activity.

[i]Hospital readmission will also be ascertained at 6 and 12 months through electronic health record review.

Treatment fidelity is monitored based on principles outlined by the National Institutes of Health's Behavior Change Consortium (Multimedia Appendix 1) and as described in a review by Borrelli [39]. Specifically, encounters for the first 50 intervention participants are audiotaped and rated by 2 study investigators (JAD and AS) using a structured tool to prevent protocol deviations by exercise therapists. After the first 50 participants, we will review a random sample of 20% of audiotaped encounters. If an individual exercise therapist falls below the a priori performance criterion (ie, a rating below the midpoint of the structured tool) based on an ongoing review of this random sample, the individual remediation will take place through a 1:1 feedback session with an expert in behavioral interventions (AS), who is part of the study team. In addition, a return to 50% monitoring may be warranted. The components evaluated include (1) length of encounter, (2) number of elements covered (eg, review of exercise activity, review of log-in frequency, addressing barriers to activity, and planning exercise for the next week), and (3) nonspecific factors (empathy and communication style) that may influence the success of the intervention.

Study Management

The Steering Committee consists of the principal investigator (JAD), a biostatistician (SA), and coinvestigator (AS). Moreover, 3 data safety monitoring board (DSMB) members (2 cardiologists and 1 biostatistician) have also been appointed by the National Institute on Aging. The DSMB meets biannually to review recruitment and monitor participant safety.

Participant Safety

The safety end point includes (1) fall-related injury (operationalized as any fall requiring acute medical care); (2) hospitalization for acute coronary syndrome; and (3) hospitalization for unstable arrhythmia. Separately, among intervention participants, study staff monitor potential exercise-related adverse events on an ongoing basis. Points of contact include weekly phone calls with the exercise therapist and electronic communication via the mHealth app, which is

checked daily. All adverse events are reported to the principal investigator and the DSMB. To reduce the likelihood of these events, participants complete the baseline 6MWT before randomization; if any adverse event occurs during the 6MWT (eg, a drop in systolic BP ≥15 mm Hg, chest pain, or ventricular arrhythmia), participants are deemed ineligible for the trial. Other exclusion criteria (severe osteoarthritis, recent joint replacement, and moderate or severe cognitive impairment) are also intended to minimize risk. Previous studies on home-based CR have reported that adverse events are uncommon [40,41].

Statistical Analysis

General

Statistical comparisons will be performed after enrollment of the full study sample, using 2-sided significance tests and 2-sided CIs; no interim comparative analyses are planned. We will begin all analyses with descriptive summary statistics and graphical displays of all variables, with attention to assessing balance in these characteristics by study group assignment and by assessing the distribution of variables relevant to the choice of statistical tests.

Primary End Point Analysis

We will assess the difference in 6MWD by calculating difference scores for each participant and comparing the intervention and usual care groups with independent group *t* tests (2-tailed), allowing for unequal variances. We will also regress 3-month 6MWD on a binary indicator of the treatment group, with adjustment for a baseline 6MWD and the stratification factor (enrollment site). Although randomization should obviate the need for additional adjustment, we will explore whether adjustment for participant-level characteristics (eg, demographic factors or referral to ambulatory CR) is necessary, using the change-in-estimate criterion. We realize that engagement with the mHealth-CR program may also affect attendance and engagement with structured ambulatory CR programs. We intend to explore this as a mediator of the effect of the assignment to the intervention arm. We will use structural equation models to estimate the direct and indirect effects of mHealth, where the direct effect is that of mHealth-CR on the 6MWD, and the indirect effect is mediated through the attendance for structured ambulatory CR.

Secondary End Point Analysis

Analysis of secondary efficacy end points will proceed in a similar fashion to that of the primary end point.

For *goal attainment*, goals are set with participants at baseline, and *goal attainment* will be assessed at 3 months using a score ranging from −2 to +2 for each participant. To preserve the ordinal nature of the data, we will calculate a median GAS score and then compare the treatment and control groups using a Wilcoxon–Mann–Whitney rank-sum test [42]. We will also compare the percentage of participants in the treatment and control groups who met their expected level of goal attainment (defined as a score of 0, +1, or +2) using a chi-square test.

Health status at 3 months will be assessed using linear regression with adjustment for baseline levels and a binary treatment indicator; if necessary, health status scores will be log-transformed to improve the approximation to normality. ADLs (BADLs and IADLs) will be assessed using longitudinal models for monthly scores, with indicators to incorporate time, a binary indicator of treatment group, and patient-level random effects to accommodate repeated assessments within individuals; if monthly scores are not approximately normally distributed, suitable transformations will be sought. Hospital readmission and death will be evaluated using Kaplan-Meier estimates, tested with log-rank statistics, and investigated using Cox proportional hazard models with adjustment for confounders if necessary. As with the models for the primary end point, in each model described, we will assess the need for adjustment for confounders using the change-in-estimate criterion.

Engagement Analysis

Among participants offered mHealth-CR, we will conduct latent class analysis to identify engagement profiles and explore whether these factors indicate membership in a class; these models will use maximum likelihood estimation, implemented with the iterative expectation–maximization algorithm, to identify a latent class solution for the set of indicators. We will evaluate the model fit using the G^2 statistic and compare models with the likelihood-difference test for nested models and the Akaike and Bayesian information criteria for nonnested models. Although we have identified four potential classes a priori (sustained engagement, disengagement, re-engagement, and deteriorating engagement), we will use the parametric bootstrap likelihood ratio test to select the optimal number of classes supported by the data in conjunction with the Akaike information criteria and Bayesian information criteria. The output of the model will be a set of *item-response* probabilities giving the likelihood of a particular characteristic within each latent class and a set of posterior predicted probabilities of latent class membership; uncertainty in predicting class membership will be summarized using the odds of correct classification diagnostic tool.

Although the engagement analysis is largely exploratory (given the paucity of data on mHealth-CR engagement), guided by literature related to engagement in other technologies, we will test whether the trajectory classes differ based on the following characteristics: age (≥80 years), sex, race or ethnicity, comorbidity burden (≥2 chronic medical conditions), frailty, social support (based on living alone), and depressive symptoms (based on Patient Health Questionnaire 9).

Power Considerations

We designed our sample size to detect a clinically meaningful difference between treatment arms in our primary efficacy end point, which is the change in the 6MWD from baseline to 3 months. A recent meta-analysis [25] found an average difference in the 6MWD of approximately 60 meters before and after traditional CR. Perera et al [43] determined that this degree of change in the 6MWD is clinically meaningful based on its relationship to other health status measures. However, Minneboo et al [27] also estimated an improvement in the 6MWD in a control group by approximately 35 m. Therefore, we require adequate power to detect a difference among groups with <25 m in change in the 6MWD (ie, the difference between a 60-m improvement in the intervention arm and a 35-m improvement

in the usual care arm); this is the same amount estimated by Gremeaux et al [44] as a minimal clinically important difference in patients with IHD. Assuming a conservative SD estimate of 60 m [25], 320 participants provide approximately 90% power to detect a difference among groups of 25 m, using a 2-sided, 0.05-level test; there is 80% power to detect a difference as small as 22 m. This sample size also provides ≥80% power to detect reasonable effect sizes for the secondary end points. On the basis of a projected attrition rate of 20% dropout in each arm, our target sample size is 400 participants, who will be randomized in a 3:1 allocation. Presuming 20% attrition, this would result in approximately 320 participants with evaluable end points (240 allocated to the intervention arm and 80 allocated to the usual care arm).

Study Response Owing to the COVID-19 Pandemic

New York was the first epicenter in the United States of the COVID-19 outbreak, with the first case being reported on January 3, 2020. Several weeks later, RESILIENT was closed to new participant enrollment by a university-wide mandate applicable to all clinical trials not related to COVID-19. The NYU Langone Main campus opened to enrollment on June 1, 2020, whereas NYU Langone Hospital—Long Island and Yale New Haven Health opened later (July 27, 2020, and July 20, 2020, respectively). Despite reopening, recruitment remained slow during the remainder of 2020 owing to a combination of slow resumption of normal clinical activities (eg, elective PCI scheduling), lower than expected hospital admissions for AMI, and patient fears of returning to the medical center for study visits.

In response to the pandemic, we made several changes to the original protocol. First, as originally designed, much of the mHealth-CR platform relied on recommendations to walk to achieve physical activity targets. This walking typically occurred outdoors or in large indoor spaces (eg, shopping malls); however, during the pandemic, many participants expressed fear of COVID-19 infection through being in public. Accordingly, with DSMB approval, we provided participants with access to several home exercise videos developed by an exercise therapist affiliated with the study (Dr Patrice Hazan). These videos are assigned by study exercise therapists weekly and include warm-ups, three levels of aerobic workouts (beginner, intermediate, and advanced), and stretch routines. Second, participants are not offered cost transportation reimbursement for the baseline (if occurring within 2 weeks of discharge rather than while in hospital) and 3-month study visits, in addition to the regular reimbursement for participation, to alleviate fears of needing to take public transportation. Finally, the University of Massachusetts was added as a study site to accomplish our recruitment goal.

Ethics Approval

The study operates on a single Institutional Review Board mechanism (sIRB), and was approved by the NYU School of Medicine Institutional Review Board. The NYU Institutional Review Board study number is 18-02017 for RESILIENT.

Results

As of December 2021, the RESILIENT trial has enrolled 116 participants. Enrollment is projected to continue until October 2023. The trial results are expected to be reported in 2024.

Discussion

The RESILIENT study will evaluate whether mHealth-CR improves functional mobility in older adults with IHD and a range of secondary outcomes, including goal attainment, health status, and hospital readmission. To our knowledge, RESILIENT is the largest trial to date for mHealth-CR in an older adult population. We designed this study in light of an aging US population that faces many current impediments to attending traditional ambulatory CR, including transportation barriers and physical impairments, coupled with widespread dissemination of mobile technologies that enable the delivery of CR at home.

Despite the promise of mHealth-CR, definitive evidence of its efficacy among older adults is lacking. Although a recent systematic review reported that mHealth-CR programs led to similar functional mobility and better adherence compared with traditional ambulatory CR programs, most trial participants were <65 years of age [45]. Furthermore, although mobile technology use has increased among the older adult population, it still lags considerably compared with younger patients, and there are many residual barriers to technology adoption, including utility cost (frustration with technology and resistance to change), physical limitations (vision impairment and arthritis), and cognitive challenges (poor memory and impaired reasoning) [46]. Any of these barriers may preclude successful patient engagement with mHealth-CR and limit its effectiveness in older adults. Similar factors may also lead to early termination of mHealth-CR even after successful initial engagement, a phenomenon that has been documented with other mHealth interventions targeted at physical activity [47]. Additional priorities for older patients (eg, clinical visits with multiple specialists) may also compete for patients' time and limit the number of mHealth-CR sessions completed.

With these barriers in mind, we designed RESILIENT as a hybrid intervention in which the use of the mHealth app is supported by a baseline visit and weekly phone calls by an exercise therapist. These clinical encounters serve to establish rapport between the exercise therapist and study participant, evaluate functional limitations and home safety, and address person-specific barriers to technology use. Although the coupling of human intervention with mHealth technology may limit the scalability of the intervention (as the success of RESILIENT is partially dependent on the proficiency of exercise therapists working at the study sites) and may make it difficult to disentangle the effects of human interaction versus the direct benefit of the mHealth platform, we feel that this tradeoff is necessary to provide the intervention with the best chance of success in the context of the target population. We have built-in auditing of these encounters and direct feedback to ensure the fidelity of the study intervention.

Another decision we made in the design of the RESILIENT trial was to enable both intervention and usual care arm participants to receive traditional ambulatory CR. An alternative strategy, which has been adopted by some trials, would have been a head-to-head study of mHealth-CR versus ambulatory CR. However, we feel that denying older adults access to traditional ambulatory CR in light of limited information about mHealth-CR's efficacy was not in accordance with the standard of care. In the context of our study, mHealth-CR may therefore serve to reinforce behaviors learned in ambulatory CR among those who attend or as a replacement for those who cannot. Accordingly, we will analyze the heterogeneity of the treatment effect among these 2 subgroups.

We have attempted to minimize bias in the RESILIENT trial through randomized treatment allocation, blinded assessment of the primary end point, and a multicenter design that includes a diverse population. However, we acknowledge that there are several potential sources of residual bias. First, as with any clinical trial, there is selection bias as people who agree to enroll are likely to be more motivated than the general patient population, typically with higher health literacy and a lower burden of chronic illness. Second, there is the potential for transfer bias whereby there may be a differential loss to follow-up in the intervention versus control arms. To minimize this possibility, participants in the control arm receive regular calls from the research coordinator to maintain a connection to the study, and the 3-month visit is scheduled on the same day as a clinical encounter whenever possible to minimize barriers to follow-up.

The generalizability of the RESILIENT trial may be limited by factors including a limited number of study sites, availability of the software only in English or Spanish, and use of a proprietary software from a single company. Furthermore, our intervention pairs the expertise of exercise therapists with an mHealth-CR platform, and positive findings should not be construed as the software platform being effective as a standalone product. We designed the intervention to couple in-person contact with the capabilities of mHealth to guide and reinforce healthy behaviors. In our opinion, this pairing of technology with human interaction—especially in an older adult population that may have limited technological proficiency—provides the best chance of success.

In summary, the RESILIENT trial will generate important evidence about the efficacy of mHealth-CR among older adults in domains including functional mobility, health status, and goal attainment. Moreover, patterns of engagement with mHealth-CR (eg, sustained engagement, declining engagement, and persistent low engagement) will be analyzed to understand the characteristics that predict different trajectories. These findings will help in designing future precision approaches to mHealth implementation and in understanding which patients are likely to engage. This knowledge is especially important in light of the COVID-19 pandemic, which has shifted much of health care to a remote, internet-based setting.

Acknowledgments

This rehabilitation using mobile health for older adults with ischemic heart disease in the home setting trial is supported by the National Institute on Aging (R01AG062520). REDCap (Research Electronic Data Capture) is supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through grant UL1 TR00000. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. TMG is supported by the Yale Claude D Pepper Older Americans Independence Center (P30AG021342).

Conflicts of Interest

None declared.

Multimedia Appendix 1
National Institutes of Health reporter data.
[PDF File (Adobe PDF File), 1614 KB - resprot_v11i3e32163_app1.pdf]

References

1. Mazzini MJ, Stevens GR, Whalen D, Ozonoff A, Balady GJ. Effect of an American Heart Association Get With the Guidelines program-based clinical pathway on referral and enrollment into cardiac rehabilitation after acute myocardial infarction. Am J Cardiol 2008 Apr 15;101(8):1084-1087. [doi: 10.1016/j.amjcard.2007.11.063] [Medline: 18394437]
2. Thomas RJ, King M, Lui K, Oldridge N, Piña IL, Spertus J, American Association of CardiovascularPulmonary Rehabilitation/American College of Cardiology/American Heart Association Cardiac Rehabilitation/Secondary Prevention Performance Measures Writing Committee. AACVPR/ACC/AHA 2007 performance measures on cardiac rehabilitation for referral to and delivery of cardiac rehabilitation/secondary prevention services. Circulation 2007 Oct 02;116(14):1611-1642. [doi: 10.1161/CIRCULATIONAHA.107.185734] [Medline: 17885210]
3. Aragam KG, Dai D, Neely ML, Bhatt DL, Roe MT, Rumsfeld JS, et al. Gaps in referral to cardiac rehabilitation of patients undergoing percutaneous coronary intervention in the United States. J Am Coll Cardiol 2015 May 19;65(19):2079-2088 [FREE Full text] [doi: 10.1016/j.jacc.2015.02.063] [Medline: 25975470]

4. Dunlay SM, Witt BJ, Allison TG, Hayes SN, Weston SA, Koepsell E, et al. Barriers to participation in cardiac rehabilitation. Am Heart J 2009 Nov;158(5):852-859 [FREE Full text] [doi: 10.1016/j.ahj.2009.08.010] [Medline: 19853708]
5. Varnfield M, Karunanithi M, Lee C, Honeyman E, Arnold D, Ding H, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. Heart 2014 Nov;100(22):1770-1779 [FREE Full text] [doi: 10.1136/heartjnl-2014-305783] [Medline: 24973083]
6. Grace S, Shanmugasegaram S, Gravely-Witte S, Brual J, Suskin N, Stewart D. Barriers to cardiac rehabilitation: does age make a difference? J Cardiopulm Rehabil Prev 2009;29(3):183-187 [FREE Full text] [doi: 10.1097/HCR.0b013e3181a3333c] [Medline: 19471138]
7. Witt BJ, Jacobsen SJ, Weston SA, Killian JM, Meverden RA, Allison TG, et al. Cardiac rehabilitation after myocardial infarction in the community. J Am Coll Cardiol 2004 Sep 01;44(5):988-996 [FREE Full text] [doi: 10.1016/j.jacc.2004.05.062] [Medline: 15337208]
8. Widmer RJ, Allison TG, Lerman LO, Lerman A. Digital health intervention as an adjunct to cardiac rehabilitation reduces cardiovascular risk factors and rehospitalizations. J Cardiovasc Transl Res 2015 Jul;8(5):283-292 [FREE Full text] [doi: 10.1007/s12265-015-9629-1] [Medline: 25946990]
9. Widmer RJ, Allison TG, Lennon R, Lopez-Jimenez F, Lerman LO, Lerman A. Digital health intervention during cardiac rehabilitation: a randomized controlled trial. Am Heart J 2017 Jun;188:65-72. [doi: 10.1016/j.ahj.2017.02.016] [Medline: 28577682]
10. Ford I, Norrie J. Pragmatic trials. N Engl J Med 2016 Aug 04;375(5):454-463. [doi: 10.1056/NEJMra1510059] [Medline: 27518663]
11. Tinetti ME, Naik AD, Dodson JA. Moving from disease-centered to patient goals-directed care for patients with multiple chronic conditions: patient value-based care. JAMA Cardiol 2016 Apr 01;1(1):9-10 [FREE Full text] [doi: 10.1001/jamacardio.2015.0248] [Medline: 27437646]
12. Fried T, McGraw S, Agostini JV, Tinetti ME. Views of older persons with multiple morbidities on competing outcomes and clinical decision-making. J Am Geriatr Soc 2008 Oct;56(10):1839-1844 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.01923.x] [Medline: 18771453]
13. Sivarajan Froelicher ES, Miller NH, Christopherson DJ, Martin K, Parker KM, Amonetti M, et al. High rates of sustained smoking cessation in women hospitalized with cardiovascular disease: the Women's Initiative for Nonsmoking (WINS). Circulation 2004 Feb 10;109(5):587-593. [doi: 10.1161/01.CIR.0000115310.36419.9E] [Medline: 14769679]
14. DeBusk RF, Miller NH, Superko HR, Dennis CA, Thomas RJ, Lew HT, et al. A case-management system for coronary risk factor modification after acute myocardial infarction. Ann Intern Med 1994 May 01;120(9):721-729. [doi: 10.7326/0003-4819-120-9-199405010-00001] [Medline: 8147544]
15. Krumholz HM. Post-hospital syndrome--an acquired, transient condition of generalized risk. N Engl J Med 2013 Jan 10;368(2):100-102 [FREE Full text] [doi: 10.1056/NEJMp1212324] [Medline: 23301730]
16. Harris P, Taylor R, Minor B, 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 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
17. Balke B. A simple field test for the assessment of physical fitness. Rep 63-6. Rep Civ Aeromed Res Inst US 1963 Apr:1-8. [Medline: 14131272]
18. O'Keeffe ST, Lye M, Donnellan C, Carmichael DN. Reproducibility and responsiveness of quality of life assessment and six minute walk test in elderly heart failure patients. Heart 1998 Oct;80(4):377-382 [FREE Full text] [doi: 10.1136/hrt.80.4.377] [Medline: 9875117]
19. Bittner V, Weiner DH, Yusuf S, Rogers WJ, McIntyre KM, Bangdiwala SI, et al. Prediction of mortality and morbidity with a 6-minute walk test in patients with left ventricular dysfunction. SOLVD Investigators. JAMA 1993 Oct 13;270(14):1702-1707. [Medline: 8411500]
20. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002 Jul 01;166(1):111-117. [doi: 10.1164/ajrccm.166.1.at1102] [Medline: 12091180]
21. Beatty AL, Schiller NB, Whooley MA. Six-minute walk test as a prognostic tool in stable coronary heart disease: data from the heart and soul study. Arch Intern Med 2012 Jul 23;172(14):1096-1102 [FREE Full text] [doi: 10.1001/archinternmed.2012.2198] [Medline: 22710902]
22. Shah MR, Hasselblad V, Gheorghiade M, Adams KF, Swedberg K, Califf RM, et al. Prognostic usefulness of the six-minute walk in patients with advanced congestive heart failure secondary to ischemic or nonischemic cardiomyopathy. Am J Cardiol 2001 Nov;88(9):987-993. [doi: 10.1016/s0002-9149(01)01975-0]
23. Cacciatore F, Abete P, Mazzella F, Furgi G, Nicolino A, Longobardi G, et al. Six-minute walking test but not ejection fraction predicts mortality in elderly patients undergoing cardiac rehabilitation following coronary artery bypass grafting. Eur J Prev Cardiol 2012 Dec;19(6):1401-1409. [doi: 10.1177/1741826711422991] [Medline: 21933832]
24. La Rovere MT, Pinna GD, Maestri R, Olmetti F, Paganini V, Riccardi G, et al. The 6-minute walking test and all-cause mortality in patients undergoing a post-cardiac surgery rehabilitation program. Eur J Prev Cardiol 2015 Jan;22(1):20-26. [doi: 10.1177/2047487313502405] [Medline: 23970071]

25. Bellet RN, Adams L, Morris NR. The 6-minute walk test in outpatient cardiac rehabilitation: validity, reliability and responsiveness--a systematic review. Physiotherapy 2012 Dec;98(4):277-286. [doi: 10.1016/j.physio.2011.11.003] [Medline: 23122432]
26. O'Connor CM, Whellan DJ, Lee KL, Keteyian SJ, Cooper LS, Ellis SJ, et al. Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. JAMA 2009 Apr 08;301(14):1439-1450 [FREE Full text] [doi: 10.1001/jama.2009.454] [Medline: 19351941]
27. Minneboo M, Lachman S, Snaterse M, Jørstad HT, Ter Riet G, Boekholdt SM, RESPONSE-2 Study Group. Community-based lifestyle intervention in patients with coronary artery disease: the RESPONSE-2 trial. J Am Coll Cardiol 2017 Jul 18;70(3):318-327 [FREE Full text] [doi: 10.1016/j.jacc.2017.05.041] [Medline: 28705312]
28. Alharbi M, Gallagher R, Kirkness A, Sibbritt D, Tofler G. Long-term outcomes from Healthy Eating and Exercise Lifestyle Program for overweight people with heart disease and diabetes. Eur J Cardiovasc Nurs 2016 Feb;15(1):91-99. [doi: 10.1177/1474515114557222] [Medline: 25344059]
29. Oliveira GU, Oliveira Carvalho V, de Assis Cacau LP, de Araújo Filho AA, de Cerqueira Neto ML, da Silva WM, et al. Determinants of distance walked during the six-minute walk test in patients undergoing cardiac surgery at hospital discharge. J Cardiothorac Surg 2014 May 31;9(1):95 [FREE Full text] [doi: 10.1186/1749-8090-9-95] [Medline: 24885130]
30. Peixoto TC, Begot I, Bolzan DW, Machado L, Reis MS, Papa V, et al. Early exercise-based rehabilitation improves health-related quality of life and functional capacity after acute myocardial infarction: a randomized controlled trial. Can J Cardiol 2015 Mar;31(3):308-313. [doi: 10.1016/j.cjca.2014.11.014] [Medline: 25633911]
31. Hurn J, Kneebone I, Cropley M. Goal setting as an outcome measure: a systematic review. Clin Rehabil 2006 Sep;20(9):756-772. [doi: 10.1177/0269215506070793] [Medline: 17005500]
32. 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]
33. Chan PS, Jones PG, Arnold SA, Spertus JA. Development and validation of a short version of the Seattle angina questionnaire. Circ Cardiovasc Qual Outcomes 2014 Sep;7(5):640-647 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.114.000967] [Medline: 25185249]
34. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc 1983 Dec;31(12):721-727. [doi: 10.1111/j.1532-5415.1983.tb03391.x] [Medline: 6418786]
35. 2008 physical activity guidelines for Americans. U.S. Department of Health and Human Services. URL: https://health.gov/paguidelines/guidelines/chapter5.aspx [accessed 2022-01-17]
36. Bhasin S, Gill T, Reuben D, Latham NK, Gurwitz JH, Dykes P, et al. Strategies to Reduce Injuries and Develop Confidence in Elders (STRIDE): a cluster-randomized pragmatic trial of a multifactorial fall injury prevention strategy: design and methods. J Gerontol A Biol Sci Med Sci 2018 Jul 09;73(8):1053-1061 [FREE Full text] [doi: 10.1093/gerona/glx190] [Medline: 29045582]
37. Lennon O, Carey A, Gaffney N, Stephenson J, Blake C. A pilot randomized controlled trial to evaluate the benefit of the cardiac rehabilitation paradigm for the non-acute ischaemic stroke population. Clin Rehabil 2008 Feb;22(2):125-133. [doi: 10.1177/0269215507081580] [Medline: 18212034]
38. Physical Activity Guidelines for Americans, 2nd edition. Washington, DC: U.S. Department of Health and Human Services; 2018.
39. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. J Public Health Dent 2011;71(s1):S52-S63 [FREE Full text] [doi: 10.1111/j.1752-7325.2011.00233.x] [Medline: 21499543]
40. Jolly K, Lip GY, Taylor RS, Raftery J, Mant J, Lane D, et al. The Birmingham Rehabilitation Uptake Maximisation study (BRUM): a randomised controlled trial comparing home-based with centre-based cardiac rehabilitation. Heart 2009 Jan;95(1):36-42. [doi: 10.1136/hrt.2007.127209] [Medline: 18332063]
41. Ades PA, Pashkow FJ, Fletcher G, Pina IL, Zohman LR, Nestor JR. A controlled trial of cardiac rehabilitation in the home setting using electrocardiographic and voice transtelephonic monitoring. American Heart Journal 2000 Mar;139(3):543-548. [doi: 10.1016/S0002-8703(00)90100-5]
42. Krasny-Pacini A, Hiebel J, Pauly F, Godon S, Chevignard M. Goal attainment scaling in rehabilitation: a literature-based update. Ann Phys Rehabil Med 2013 Apr;56(3):212-230 [FREE Full text] [doi: 10.1016/j.rehab.2013.02.002] [Medline: 23562111]
43. Perera S, Mody S, Woodman R, Studenski S. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc 2006 May;54(5):743-749. [doi: 10.1111/j.1532-5415.2006.00701.x] [Medline: 16696738]
44. Gremeaux V, Troisgros O, Benaïm S, Hannequin A, Laurent Y, Casillas J, et al. Determining the minimal clinically important difference for the six-minute walk test and the 200-meter fast-walk test during cardiac rehabilitation program in coronary artery disease patients after acute coronary syndrome. Arch Phys Med Rehabil 2011 Apr;92(4):611-619. [doi: 10.1016/j.apmr.2010.11.023] [Medline: 21440707]
45. Wongvibulsin S, Habeos EE, Huynh PP, Xun H, Shan R, Porosnicu Rodriguez KA, et al. Digital health interventions for cardiac rehabilitation: systematic literature review. J Med Internet Res 2021 Feb 08;23(2):e18773 [FREE Full text] [doi: 10.2196/18773] [Medline: 33555259]

46. Searcy RP, Summapund J, Estrin D, Pollak JP, Schoenthaler A, Troxel AB, et al. Mobile health technologies for older adults with cardiovascular disease: current evidence and future directions. Curr Geri Rep 2019 Jan 28;8(1):31-42. [doi: 10.1007/s13670-019-0270-8]
47. Recio-Rodriguez JI, Agudo-Conde C, Martin-Cantera C, González-Viejo MN, Fernandez-Alonso MD, Arietaleanizbeaskoa MS, EVIDENT Investigators. Short-term effectiveness of a mobile phone app for increasing physical activity and adherence to the mediterranean diet in primary care: a randomized controlled trial (EVIDENT II study). J Med Internet Res 2016 Dec 19;18(12):e331 [FREE Full text] [doi: 10.2196/jmir.6814] [Medline: 27993759]

Abbreviations

6MWD: 6-minute walk distance
6MWT: 6-minute walk test
ADL: activity of daily living
AMI: acute myocardial infarction
BADL: basic activity of daily living
BP: blood pressure
CR: cardiac rehabilitation
DSMB: data safety monitoring board
EHR: electronic health record
GAS: goal attainment scale
IADL: instrumental activity of daily living
IHD: ischemic heart disease
mHealth: mobile health
mHealth-CR: mobile health cardiac rehabilitation
NYU: New York University
PCI: percutaneous coronary intervention
REDCap: Research Electronic Data Capture
RESILIENT: rehabilitation using mobile health for older adults with ischemic heart disease in the home setting
RPE: relative perceived exertion

Edited by G Eysenbach; submitted 16.07.21; peer-reviewed by S Martin, M Östbring, A Beatty; comments to author 16.09.21; revised version received 29.10.21; accepted 08.12.21; published 03.03.22.

Please cite as:
Dodson JA, Schoenthaler A, Sweeney G, Fonceva A, Pierre A, Whiteson J, George B, Marzo K, Drewes W, Rerisi E, Mathew R, Aljayyousi H, Chaudhry SI, Hajduk AM, Gill TM, Estrin D, Kovell L, Jennings LA, Adhikari S
Rehabilitation Using Mobile Health for Older Adults With Ischemic Heart Disease in the Home Setting (RESILIENT): Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2022;11(3):e32163
URL: https://www.researchprotocols.org/2022/3/e32163
doi:10.2196/32163
PMID:35238793

Protocol

Toward Successful Implementation of Artificial Intelligence in Health Care Practice: Protocol for a Research Program

Petra Svedberg[1], PhD; Julie Reed[1], PhD; Per Nilsen[1,2], PhD; James Barlow[1,3], PhD; Carl Macrae[1,4], PhD; Jens Nygren[1], PhD

[1]School of Health and Welfare, Halmstad University, Halmstad, Sweden

[2]Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

[3]Centre for Health Economics and Policy Innovation, Imperial College Business School, London, United Kingdom

[4]Centre for Health Innovation, Leadership and Learning, Nottingham University Business School, Nottingham, United Kingdom

Corresponding Author:
Jens Nygren, PhD
School of Health and Welfare
Halmstad University
Box 823
Halmstad, 30118
Sweden
Phone: 46 035167100
Email: jens.nygren@hh.se

Abstract

Background: The uptake of artificial intelligence (AI) in health care is at an early stage. Recent studies have shown a lack of AI-specific implementation theories, models, or frameworks that could provide guidance for how to translate the potential of AI into daily health care practices. This protocol provides an outline for the first 5 years of a research program seeking to address this knowledge-practice gap through collaboration and co-design between researchers, health care professionals, patients, and industry stakeholders.

Objective: The first part of the program focuses on two specific objectives. The first objective is to develop a theoretically informed framework for AI implementation in health care that can be applied to facilitate such implementation in routine health care practice. The second objective is to carry out empirical AI implementation studies, guided by the framework for AI implementation, and to generate learning for enhanced knowledge and operational insights to guide further refinement of the framework. The second part of the program addresses a third objective, which is to apply the developed framework in clinical practice in order to develop regional capacity to provide the practical resources, competencies, and organizational structure required for AI implementation; however, this objective is beyond the scope of this protocol.

Methods: This research program will use a logic model to structure the development of a methodological framework for planning and evaluating implementation of AI systems in health care and to support capacity building for its use in practice. The logic model is divided into time-separated stages, with a focus on theory-driven and coproduced framework development. The activities are based on both knowledge development, using existing theory and literature reviews, and method development by means of co-design and empirical investigations. The activities will involve researchers, health care professionals, and other stakeholders to create a multi-perspective understanding.

Results: The project started on July 1, 2021, with the Stage 1 activities, including model overview, literature reviews, stakeholder mapping, and impact cases; we will then proceed with Stage 2 activities. Stage 1 and 2 activities will continue until June 30, 2026.

Conclusions: There is a need to advance theory and empirical evidence on the implementation requirements of AI systems in health care, as well as an opportunity to bring together insights from research on the development, introduction, and evaluation of AI systems and existing knowledge from implementation research literature. Therefore, with this research program, we intend to build an understanding, using both theoretical and empirical approaches, of how the implementation of AI systems should be approached in order to increase the likelihood of successful and widespread application in clinical practice.

International Registered Report Identifier (IRRID): PRR1-10.2196/34920

(JMIR Res Protoc 2022;11(3):e34920) doi:10.2196/34920

KEYWORDS

process evaluation; complex intervention; implementation; knowledge exchange; health policy; organizational change; capacity building; qualitative methods; framework analysis

Introduction

Background

In many high-income countries, policy makers, authorities, and care providers have great expectations that the uptake of information technology (IT) innovations in health care will contribute to improved efficiency and quality of health care as well as improved clinical and health outcomes [1,2]. Today, artificial intelligence (AI), as an IT innovation, holds significant promise for enhancing health care [3]. However, as for most other types of IT innovations, the uptake of AI in health care is still at an early stage [4]. Even though there are ample examples of successful implementation of innovations in health care practice, there are often considerable challenges to implement new technology in health care [5]. While implementation science has advanced our knowledge about barriers to implementing such innovations and provided guidance about what strategies can be used to overcome these barriers, this knowledge has not yet been applied for understanding or supporting the implementation of AI in health care [6].

Despite the extensive and still increasing interest in and research on AI in health care, it is evident that there is a huge knowledge gap on how to tackle implementation challenges and how to successfully plan to increase the likeliness of sustainable adoption in practice [7-9]. To date, research on AI in health care has mostly focused on algorithm development, proof-of-concept evaluations, and technical, legal, and ethical challenges [3,8,10-12]. The research is primarily published within the fields of computer science, engineering, and medical informatics [7] as well as the clinical specialties of oncology, neurology, radiology, and cardiology [3].

A variety of viewpoint articles, commentaries, and guidelines have been produced, discussing the potential of AI but also the challenges and aspects to consider in its development and introduction in health care [8,11,13,14]. However, such articles are mainly part of a scientific discourse built on an underdeveloped evidence base. This reduces their value as guidance of implementation initiatives in practice. The research literature, in the form of systematic reviews, that is relevant to implementation-related issues has mostly addressed aspects of regulation, privacy, and legal issues [15,16]; ethics [9,16,17]; clinical and patient outcomes [18-20]; and economic impact [21]. These studies point to the importance of undertaking more implementation research to study AI implementation in real-world clinical settings.

A few empirical studies with robust methodology, such as randomized controlled studies, have investigated the effects of implementation of AI technology in practice [20], but there are no AI-specific implementation theories, models, or frameworks that could provide guidance for how to translate the potential of AI into daily health care practices [22]. Thus, there is currently a paucity of knowledge concerning several key issues, including barriers and facilitators to successful implementation of AI in health care; what strategies might be used to support AI implementation; how the use of AI might change existing clinical workflows, roles, and responsibilities; or how the infrastructure of management and governance should be constructed to be effective [23].

Many barriers to optimal implementation of new innovations in health care have been identified. Among these is the alignment of the innovation with the setting and contextual circumstances in which the innovation will be used, and the degree of adaptation to the needs and wants of the stakeholders that the innovation is intended to support. Such barriers often result in poor program fidelity and a lack of sustainability in change behavior at individual, organizational, and system levels from the innovation [24,25]. A general dilemma is how to involve both early adopters and the majority of health care professionals. While early adopters usually are fairly easy to engage and are motivated in the development and use of new innovations, the large group of health care professionals represents much greater variation with regard to their willingness to be involved and in motivation to integrate the use of innovations in their everyday practice [26]. There is also a general lack of strategic knowledge on how to ensure that innovations can be safely and effectively integrated into local infrastructures for routine use and embedded into clinical workflows [27].

In order to meet these challenges, implementation models, frameworks, and strategies are needed that contain multiple components embedded in the context of application [28,29]. In particular, coproduction has been stressed as a key factor for successful implementation of innovations in health care in order to develop effective strategies and to ensure that value is created [30]. This requires knowledge about both barriers and facilitators that influence an innovation's use and implementation strategies that are designed to overcome the identified barriers and maximize the use of facilitating factors. Successful implementation usually requires an active change process aimed to achieve both organizational- and individual-level use of the intervention as designed. However, implementation is often a critical process between an organizational decision to adopt and support an innovation and the professional's willingness and ability to use it in their daily work. In order for AI to be successfully introduced to change clinical practice, we need to understand current practices and the contexts in which those practices are conducted, as well as how AI would fit with or change those ongoing practices and processes. However, the experiences of the professionals and patients who use a particular AI application are often overlooked [12]. It is, therefore, important to understand how care is actually delivered, how data are currently used to inform health care, and how new technologies impact individual and organizational

decision-making processes and alter the roles, responsibilities, and relationships that shape clinical work.

Given that health care practitioners have an ethical and legal duty of care to their patients and are responsible for clinical recommendations and decision-making, transparency regarding how clinical decisions are made—both with and without the use of AI—is important. This will necessitate the implementation of systems that encourage health care professionals to interact with AI in ways that augment their clinical decision-making without compromising their primary responsibilities and duties to patient care [31]; as well, implemented systems will need to ensure that quality and safety are appropriately governed and assured [32]. In order to achieve successful implementation, we need to close the gap between how work is imagined and what is actually taking place, and we need to build accurate, evidence-based, and shared understanding of what is really happening. Thus, more rigorous and empirical implementation studies of AI in health care, underpinning strong theory and real-world understanding, are urgently needed.

Objectives and Aims

This paper provides an outline for the first 5 years of a program lasting 8 years, using implementation and improvement frameworks and co-design between researchers, health care professionals, and stakeholders. There are two objectives for this initial period of the program and a third objective for the final period (not covered by this paper):

1. To develop a theoretically informed framework for AI implementation in health care that can be applied to facilitate such implementation in routine health care practice.
2. To carry out empirical AI implementation studies guided by the AI implementation framework, thus producing evidence on the value of the framework and generating learning for enhanced knowledge and operational insights to guide further refinement of the framework.
3. To practically apply the developed framework in clinical practice to develop regional capacity to provide the practical resources, competencies, and organizational structure required for AI implementation.

Methods

Setting

This research program is part of a regional and national initiative to build infrastructure to support the implementation of AI into practice. Together, Halmstad University (HU) and Region Halland (RH) have appointed the implementation of AI in health care as a prioritized cooperation area for research and innovation with the aim to accomplish more information-driven care. Together with a broad cluster of business partners, HU and RH have established a research center for information-driven care called CAISR (Center for Applied Intelligent Systems Research) Health, with funding provided by the Knowledge Foundation (a Swedish funding agency).

The CAISR Health research center builds on multidisciplinary collaboration between academics with expertise in data analytics, digital service, health economics, health care implementation, and health management. The center is unique in that it brings all these competencies together with partners in regional and municipal health care and industry in a joint undertaking to promote the use of information-driven care approaches in clinical practice. Emphasis has been placed on research to achieve a broad and deep understanding of how AI can be successfully implemented in health care. AI, and its implementation, in this context is broadly defined as the use, primarily, of machine learning but also other sophisticated computational techniques on health care data to support and improve clinical workflows, processes, and systems to improve quality and optimize resource use. This includes several layers within the sociotechnical ecosystem around the technology, dealing with (1) generating, cleaning, and labeling data; (2) developing models and verifying, assuring, and auditing AI tools and algorithms; (3) incorporating AI outputs into clinical decisions and resource allocation; and (4) the shaping of new organizational structures, roles, and practices. This implies that the implementation of AI extends beyond any specific intelligent technology and encompasses the whole sociotechnical system that surrounds and supports a particular technology. Given this, the focus of implementation is this broader AI system; therefore, AI is hereafter referred to as AI systems.

Realizing the ambition of successfully implementing AI systems in health care thus requires more than merely technological development. It also depends on knowledge about social, cultural, organizational, and implementation challenges; how information-driven care can be supported by AI systems; and what value can be created from different perspectives throughout the health care system.

Several major investments have been made in infrastructure and various forms of collaboration to achieve the ambition of high-quality research on and development of information-driven care.

One major investment was a research group for health care improvement that was developed at HU, with national and international collaborations with academic partners. The unit is interdisciplinary and combines applied and theoretical approaches, using both qualitative and quantitative methods, and is built up through extensive collaboration with users, regions, municipalities, and industry. The research focuses on questions about how health innovations, in the form of interventions supported by digital services and health data analysis, can be developed, implemented, and evaluated to provide health care organizations with knowledge and support to achieve high-quality care and improved health outcomes for particular groups.

Another major investment was a national effort that was initiated to establish an innovation environment for information-driven care to improve Swedish health care. The initiative, which is funded by the Swedish innovation agency Vinnova, aims to develop health care to become more information driven, individualized, and scalable through implementation and use of AI systems. This is done in close collaboration with public, private, and academic parties. Both HU and RH have key roles

as innovators, and the environment involves both the Swedish Association of Local Authorities and Regions and AI Sweden.

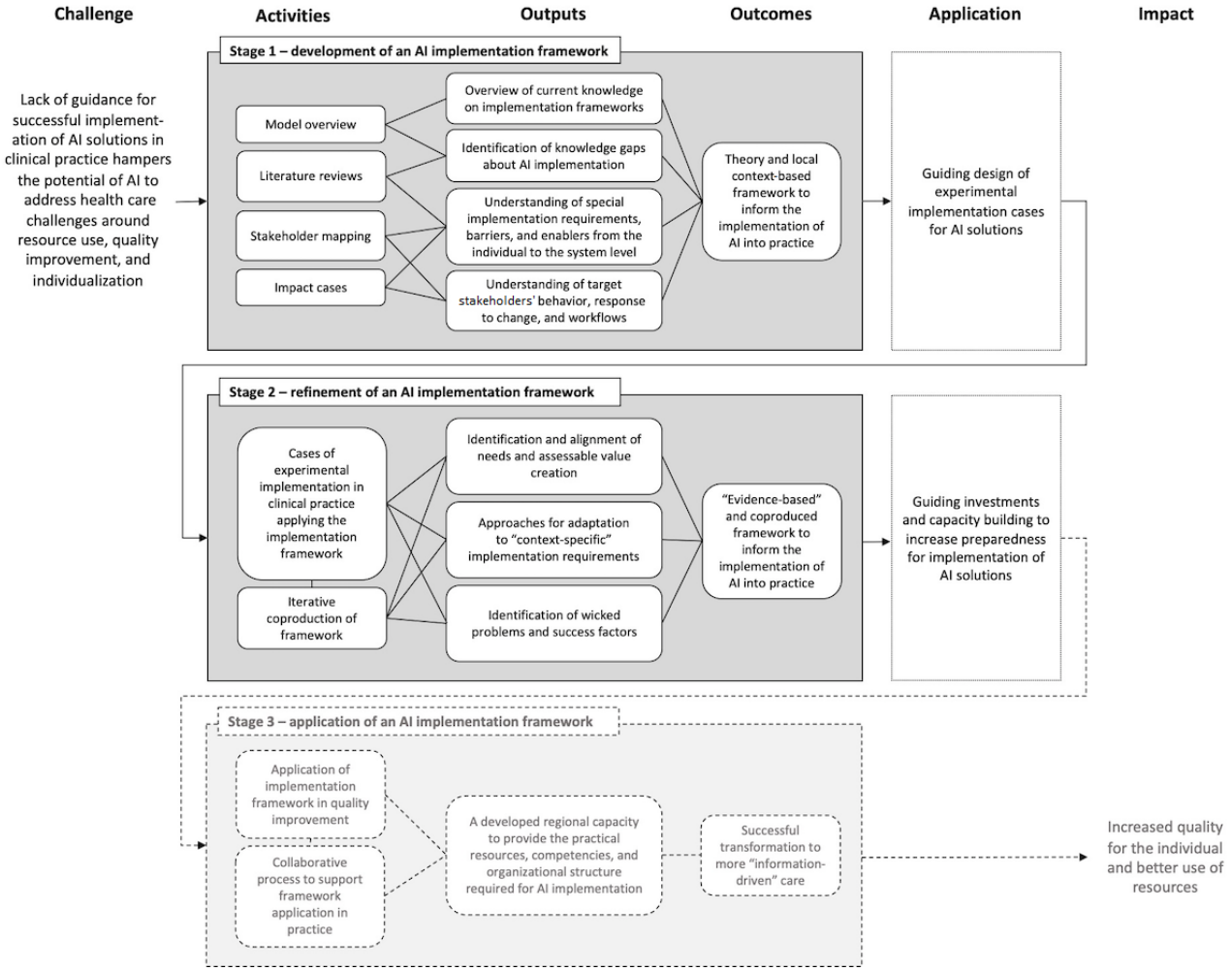
An additional major investment was an innovation arena, Leap for Life, for information-driven care that was established at HU with the purpose to act as a driving force for collaborations, innovations, and change within health care, from regional, national, and international perspectives. Leap for Life has its premises at HU and was built in partnership with RH and all the municipalities in Halland.

The final major investment was a strategic health care analysis and research platform, Regional Healthcare Information Platform (RHIP), which was established to enable agile management and analysis of the clinical and administrative data of every consumer of regional health care; the platform has had public funding since 2009 [33]. The platform is a structured, filtered, and pseudoanonymized far-reaching subset of the different data warehouses of patient and administrative data collected from over 20 different regional IT systems and national registers. It is designed to facilitate rapid analysis for clinical and management research and evaluation purposes. The Center for Information-Driven Care (CIDD) is a core facility within RH with responsibility for analyzing, simulating, and following up the effects of changes to the health care system at the system level and combining assessments of quality and costs. CIDD's mission is to perform system analyses, based on machine learning models; to be able to support change in and assure the quality of health care processes; to manage and develop the analytical platform and its methods; and to support research in the area. A research platform in the form of a Health Data Center (HDC), with a similar structure and content as RHIP, is located at HU and was established to enable the use of the content in RHIP for research and development purposes in collaboration with academia and the public and private sectors.

Research Design

This overall program can be described using a logic model [34] and will contribute to the development of a methodological framework [35] for planning, facilitating, and evaluating implementation of AI systems in health care and to support capacity building for its use in practice (Figure 1). The logic model is based on common challenges within health care systems regarding achieving and maintaining quality and optimizing the use of available resources, at the same time as the system undergoes a structural transformation with increased digitization and individualization of health care processes, workflow, and organization. A combination of activities has been identified as necessary to achieve the outputs required to accomplish the desired change within the scope of predefined outcome goals.

Figure 1. Logic model for the program, including activities identified as necessary to achieve the intended outputs and outcomes in relation to the desired impact and the defined challenge. AI: artificial intelligence.



The logic model has been developed to address issues about implementation through further understanding of the social, cultural, and organizational challenges for implementation and the potential value that can be created from drawing on different perspectives throughout the health care system. In our case, the context for the use of the logic model is AI systems in health care, but it could be equally relevant for studying, planning, facilitating, and evaluating implementation of any other type of technology in health care.

The logic model is divided into three time-separated stages, with a focus on theory-driven and coproduced framework development. The activities are based on both knowledge development, using existing theory and literature reviews, and method development by means of co-design and empirical investigations. The activities involve researchers, health care professionals, and other stakeholders, thus creating a multi-perspective understanding of how the implementation of AI systems should be approached in order to increase likelihood of successful implementation and application in clinical practice.

The framework development spans across the logic model. Workstreams will address the following: (1) identification of evidence to inform the methodological framework (first stage), (2) development of the methodological framework (first stage), and (3) evaluation and refinement of the methodological framework through empirical studies (second stage) [35].

The third stage of the logic model will go on to focus on the practical application of the developed framework in clinical practice. This stage lies further ahead in time and requires work that is more closely integrated with quality-improvement work in clinical practice. It is, therefore, beyond the scope of the first phase of the research program and is not addressed further in this protocol (Figure 1).

Stage 1: Development of an AI Systems Implementation Framework (Time Frame 2021-2023)

Outline

The first stage of this research program focuses on exploratory research with the aim to generate an overview and understanding of the area of interest (ie, "implementation of AI systems in health care"). This will be achieved through the following activities:

1. Development of a preliminary version of a theoretical framework that accounts for different perspectives in implementation in health care in the current literature.
2. Literature reviews of potentially relevant existing theories, models, and frameworks for guiding AI systems implementation in health care and empirically based experiences from implementing AI systems in health care.
3. Mapping of stakeholder perspectives and local awareness and boundaries for AI systems implementation.
4. Theory development based on empirical findings and insights from several initiatives for tentative or pilot implementation of AI systems in health care practice.

Model Overview

The first activity is the framing of perspectives on implementation to be incorporated in a theoretical framework based on current literature.

Design and Aim

This activity will investigate aspects on implementation and improvement found in models and frameworks in relation to health care [35]. This will give an overview of the different multidisciplinary perspectives on implementation and improvement that need to be considered in order to structure work for successful implementation of new AI systems in health care.

Data Collection and Analysis

The investigation will be informed by frameworks describing different perspectives on implementation in health care: service-centric, innovator-centric, and evaluation-centric perspectives. These three perspectives represent the different interests of the major stakeholder groups that we will be engaging (ie, health care professionals and patients, industry partners, and academics). In our aim to co-design a framework that engages all of these groups, we believe it is important to understand the perspectives of individual groups, to shape and understand how the different perspectives come together, and to identify where synergies and tensions are likely to exist.

We have identified three available frameworks that represent these three perspectives, and these will form a basis for initial conceptual development of the framework, which will subsequently be informed by wider literature in each of these fields and empirical data specific to AI implementation. The frameworks we have selected represent these diverse viewpoints; they are leading frameworks in their respective fields and are based on extensive literature review and empirical study. The frameworks are as follows: the Successful Healthcare Improvements From Translating Evidence in Complex Systems (SHIFT-Evidence) framework [29]; the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework [36]; and the process evaluation of complex interventions (PECI) framework [37].

The service-centric SHIFT-Evidence framework contains three strategic principles: (1) "act scientifically and pragmatically," (2) "embrace complexity," and (3) "engage and empower." These provide actionable guidance to inform the implementation and improvement process. In comparison to other leading implementation frameworks, SHIFT-Evidence offers a service-centric framework for conceptualizing the work required for successful implementation as comprised of an ongoing and multifaceted process of intervening in complex health care systems [38]. The framework makes clear that innovation implementation alone is unlikely to achieve improvements in care unless the wider interdependent issues are understood and addressed, including how the service is run and the behaviors of staff and patients. The framework emphasizes the real-world and often messy environments of service providers and service users that need to be the foundation of real-world implementation if it is to succeed in achieving improvements. Therefore, this framework is a sensible choice

to ensure that a holistic approach is taken to AI systems implementation.

The innovator-centric NASSS framework comprises six domains: (1) the condition or illness, (2) the technology, (3) the value proposition, (4) the adopter system (ie, comprising professional staff, patients, and lay caregivers), (5) the organizations, and (6) the wider institutional and societal context; a seventh domain considers interactions and adaptations over time [36].

The NASSS framework adopts an innovator-centric view of the spread of technology and, therefore, complements the service-centric view of the SHIFT-Evidence framework. In addition, the NASSS framework focuses on the specific challenges related to technological innovations. While this is not specific to AI systems, there is considerable overlap in the influencing factors of general technological interventions (eg, the types of data generated, the knowledge needs to use the technology, the technology supply model, understanding the value proposition for the supply side [ie, developer] and demand side [ie, patient and service], the regulatory environment, and the position of professional bodies).

The evaluation-centric PECI framework [37] provides actionable guidance to evaluate the implementation process. It does so by examining implementation factors (ie, the process through which interventions are delivered, and what is delivered in practice), mechanisms of impact (ie, the intermediate mechanisms through which intervention activities produce intended or unintended effects), and contextual factors (ie, factors external to the intervention that may influence its implementation, or whether its mechanisms of impact act as intended).

Data on different aspects within these frameworks will be extracted from the original manuscripts for each implementation framework. Assessment will consider both the number of related aspects and the breadth of issues they address. Thereafter, a literature search of implementation frameworks or models for health care will be conducted; identified literature will then be deductively analyzed to extend the understanding of the concepts identified in the initial analysis and to potentially identify, describe, and amend further concepts. This is an iterative process; after deductively synthesizing the new data, they will be discussed with key experts, researchers, and potential framework users (eg, health care and industry partners) for refinements [35].

Literature Reviews

The second activity is the completion of two scoping reviews investigating the following: (1) frameworks and models used to guide implementation of AI systems in health care and (2) empirical reports of AI systems implementation.

Design and Aim

Two scoping reviews based on the framework by Arksey and O'Malley [39] will be conducted in four stages: identification of relevant literature, selection of studies, charting of data, and synthesizing results. Scoping reviews are well suited to fields that are not yet comprehensively reviewed [39,40]. This methodology allows for summarizing the current state of the art, to map key findings, to identify research gaps, and to make recommendations for future research and practices [40]. By using search terms that capture general aspects of implementation, all three perspectives (ie, service-centric, innovator-centric, and evaluation-centric perspectives) will be covered in the reviews. This will allow for evaluation of how these perspectives have influenced current research on AI implementation in health care and where there is agreement or conflicting knowledge between the three perspectives.

Two reviews will be performed with separate but interrelated aims. The first review will investigate the existence and use of AI-specific implementation frameworks and identify which ones have been used to understand and support AI systems implementation in health care. The second review will investigate the existing empirical research on AI implementation and what lessons can be learned from this research for potential application in the development of an AI-specific framework for implementation in health care. This will allow us to map and synthesize barriers and facilitators for successful implementation into health care practice.

Data Collection and Analysis

Electronic databases—Cochrane, EBM Reviews, Embase, MEDLINE, and PsycInfo—will be searched to identify publications that were published in the last 10 years. The reviews will focus on studies published in English and investigating issues concerning AI systems implementation in health care. Two reviewers will independently review the titles and abstracts of the identified papers on the basis of inclusion and exclusion criteria using the Rayyan web platform. Any disagreements will be resolved by involving a third independent reviewer. Following the title and abstract review stage, full texts of identified papers will be obtained for more thorough review. The data will be extracted from the final set of studies using data extraction forms, including bibliographic details of the study, population and age group studied, geographical location, contexts where the studies were carried out, the dimensions of the implementation that were explored, and the specific AI systems of interest. Results will be collated, summarized, and reported deductively using the frameworks (ie, service-centric, innovator-centric, and evaluation-centric) investigated during the first activity.

Stakeholder Mapping

The third activity is the mapping of stakeholders and local awareness and boundaries for AI systems implementation.

Design and Aim

This activity has an exploratory qualitative approach [41,42] in order to understand contextual aspects regarding requirements, barriers, and enabling factors for the introduction of AI systems and, thus, their development and plan for implementation. The aim is to inform the planning process for AI systems implementation, to ensure that developed strategies to support suitable implementation approaches are based on stakeholder perspectives, and to avoid potential barriers to AI systems integration in clinical practice. The studies seek to answer the following research questions:

- How can stakeholder perspectives be used to help understand risks and opportunities in relation to AI systems implementation in health care?
- What assumptions are made by different stakeholders about the system and people? What opportunities and risks do the stakeholders perceive?
- How is the potential adoption of AI systems conceptualized by different stakeholders who are responsible for or impacted by AI systems?

Data Collection and Analysis

Key individuals, groups, or both who can affect or are affected by the implementation of AI systems in health care will be identified based on a high-level managerial starting point, and will continue to be identified through a snowball recruitment procedure. Recruitment will begin based on different settings, starting at the regional health care strategic management level and continuing with two frontline health departments, within which AI systems are intended to be developed and implemented. Within these settings, data in relation to implementation of AI systems will be collected from the management level, development level, and clinical practices. Individual interviews will be conducted with stakeholders representing different needs, experiences, interests, mandates, and responsibilities and will continue until the informants do not identify any further types of stakeholder perspectives to be included. This procedure will allow for informant perspectives to be represented, including from health care professionals, managers and quality developers, IT technicians, politicians, and patients, among others.

The interviews will be based on an interview guide structured around the following: (1) the roles and previous experiences the informants have concerning the application of AI systems in practice, (2) the opportunities and problems that need to be solved and considered to create strategies to support suitable introduction of AI systems, (3) beliefs and attitudes concerning the possibilities of using AI systems to support health care improvements, and (4) the obstacles, opportunities, and facilitating factors that need to be considered to enable AI systems to fit into existing processes, methods, and systems.

The analysis will be both inductive and deductive and will be structured by qualitative content analysis [41] and stakeholder mapping approaches [43]. Qualitative content analysis largely focuses on the subject and context, exploring differences between and similarities within codes and categories [41,42]. The analysis method is chosen based on the need to structure and condense various aspects of implementation that are described by the stakeholders based on the material collected. Interview data will be analyzed by (1) identifying stakeholders and their efforts in the ecosystem of services and their interest versus influence, (2) investigating different stakeholders' perspectives and assumptions regarding AI systems in health care (ie, inductive qualitative content analysis), and (3) investigating how the potential adoption of AI systems is conceptualized by different stakeholders (ie, deductive qualitative content analysis) using the frameworks (ie, service-centric, innovator-centric, and evaluation-centric) investigated during the first activity.

Qualitative content analysis, both with an inductive and deductive approach, consists of several analysis steps [41]. Data analysis will be led by one researcher, but he or she will collaborate with other researchers in order to increase the credibility and trustworthiness of the interpretation [42]. We will seek agreement between the researchers and continuously discuss how well the codes, categories, and themes represent the data. The researchers will work both individually and together during the analysis process until consensus is reached.

Impact Cases

The fourth activity is developing theory based on experiences from several initiatives for tentative or pilot implementations of AI systems in practice.

Design and Aim

This activity takes an exploratory qualitative approach to understand the different competencies, knowledge, perspectives, and logics that are represented by the different included or affected stakeholders, their professional roles, and their organizations in relation to the introduction of AI systems in clinical practice. The study will provide information on the preimplementation social, professional, and organizational context and structures that exist in the different arenas of the health system and will provide a base analysis of current practical, professional, organizational, and clinical structures and arrangements. The aim is to inform the planning process of AI systems implementation and to ensure that developed strategies to support suitable implementation approaches are based on the different social, cultural, practical, and organizational boundaries that will affect the integration of AI systems in clinical practice, as well as on how those social structures change through the process of AI systems implementation. The cases will provide a foundation for answering the following research questions:

- What are the special implementation requirements, barriers, and enablers specific to AI systems in health care?
- What are the target stakeholders' behaviors, responses to change, and new workflows as a result of introducing AI systems in health care?
- How do social structures and roles change through the process of AI systems implementation, particularly in relation to implementation requirements, barriers, and enablers?

Data Collection and Analysis

In this phase, a wide variety of data will be included in a qualitative thematic analysis [44] in the form of stakeholder interviews and meeting notes, as well as observations, plans, and reports from cases. Data will be collected on a set of concrete and specific cases of potential, tentative, or ongoing implementation of AI systems. Included cases will consist of the research, development, and innovation projects carried out during the project period within the region's and university's research and development activities. The cases will mainly be identified through the steering group for the research center CAISR Health, which includes leadership representatives from the university, the region, and business partners.

The purposively selected projects will allow for the collection of a diverse body of data on a wide range of practical aspects of AI system implementation, including data collection in formative studies (ie, through observations, interviews, and surveys), design and development studies (ie, interviews, focus groups, workshops, and usability and feasibility evaluations), and intervention and implementation studies (ie, interviews, quantitative process and effect evaluations, and health economy assessments). The data analysis method will be chosen based on the need to structure and condense the various aspects of implementation that can be drawn from a variety of data sources from different types of cases. In combination with the analysis from the third activity, this analysis will help identify the different types of barriers and enablers for AI system implementation as well as the social and organizational processes that maintain and create them.

Stage 2: Refinement of an AI Systems Implementation Framework (Time Frame 2023-2026)

Outline

The second stage of this research program focuses on implementation research with the aim of further identifying and aligning needs with assessable value creation in the implementation of AI systems. The objective is to develop a better understanding of the approaches that can be taken to adapt implementation strategies and AI systems to context-specific implementation requirements. Developing such approaches depends on exploration of the nature of wicked problems and success factors that can hinder or facilitate successful implementation in practice. This will be achieved through the development and execution of several implementation projects, together with partners from public and private health care and companies, and through an iterative process for coproduction of a refined framework.

Implementation Studies

Design and Aim

Each implementation study will be initiated and designed based on conversations and workshops with representatives from public and private care providers, companies that develop or apply AI technology and service development, and researchers. A network of such actors has been established through the formation of CAISR Health and will be expanded and developed through, among other approaches, the built-up infrastructure that is linked to the national innovation environment for information-driven care, Leap for Life, RHIP, CIDD, and HDC. The design and aim of each study will be adapted based on each study's specific objectives and context. For each study, the study design will be based on the structure and requirements of the implementation framework developed in Stage 1. The cases for the studies will be selected and used to explore and develop the implementation framework, with each case chosen to enable in-depth examination of particular aspects of the implementation framework and the broader sociotechnical system around AI technology, particularly with regard to the support and improvement of clinical workflows, processes, and systems.

Data Collection and Analysis

For each implementation study, data will be collected through qualitative and quantitative methods (project documentation, interviews, evaluations, etc) to examine the extent to which the implementation was carried out according to the plan developed based on the implementation framework. Data collection will focus on capturing aspects such as what deviations were made and why they were made; which aspects of the implementation worked well and what problems arose; how governance, management, and assurance processes were established around, and as part of, these new AI systems, and the impacts these processes had on AI system implementation and use; and whether implementation achieved the desired change, whether it contributed to an increased value, and, if so, for whom. Each implementation study will function as a separate project with its own project management and goals.

The analysis of the continuously collected data will be based on the grounded theory methodology [45]. In this methodology, the researcher explores how individuals construct meaning and actions, and the researcher aims to examine the context, existing structures, hierarchies of power, networks, and relationships in which the actions and processes take place. Analysis will be conducted by iteratively coding and categorizing the qualitative data that are collected, in order to develop higher-order concepts and explanations of the processes and challenges of AI implementation. The methodology is suitable because it enables exploration and understanding of various actions and processes involved in AI implementation. The aim is to develop a theoretical model that can inform policy and practice with regard to AI implementation. We will follow the quality criteria for constructivist grounded theory (ie, credibility, resonance, originality, and usefulness) [45].

By running all implementation studies as part of a larger program collaboration over several years, the joint documentation around each implementation study will contribute knowledge from many different perspectives and contexts. This will provide a foundation for continuous refinement and development of the framework for AI systems implementation developed in Stage 1.

Iterative Coproduction of Framework

Design and Aim

The constructivist grounded theory approach will be used [45] to develop, revise, and refine the previously generated framework (Stage 1) through the analysis of new empirical data and experiences of stakeholders who have worked practically with the framework in the implementation of AI systems in practice. This will be done to examine practices, roles, role relationships, and the situated and material organizational aspects of using the framework and implementing AI. The main objective is to iteratively refine the framework developed in Stage 1 in coproduction with stakeholders, in order to produce an evidence-based and coproduced framework that accounts for appropriate strategies that will enable the production of relevant value in practice.

Data Collection and Analysis

The grounded theory approach will guide iterative data collection and analysis of a range of documentation from implementation studies, such as project plans, interviews, observations, surveys, workshops, and meeting notes. This process will inform the interpretation and refinement of the framework. Grounded theory has been chosen, as it focuses on human experiences and actions in social contexts [45]; the theory is appropriate, since the phenomenon of implementing AI systems in practice is a complex area that involves many different people, actors, social structures, and processes. Members of a core team will be identified from the research group and from the different implementation studies to work together throughout the coproduction process in order to leverage different perspectives, experiences, and insights. The coproduction process involving the core team will be interactive and iterative, and will involve various management functions and stakeholders with experiences of the implementation of AI systems in health care. Particular focus will be placed on the interplay between theory and empiricism, in order to explain novel empirical data and how this data can modify and challenge the existing concepts and theoretical constructions in previous iterations and versions of the framework generated in Stage 1.

Ethics and Dissemination

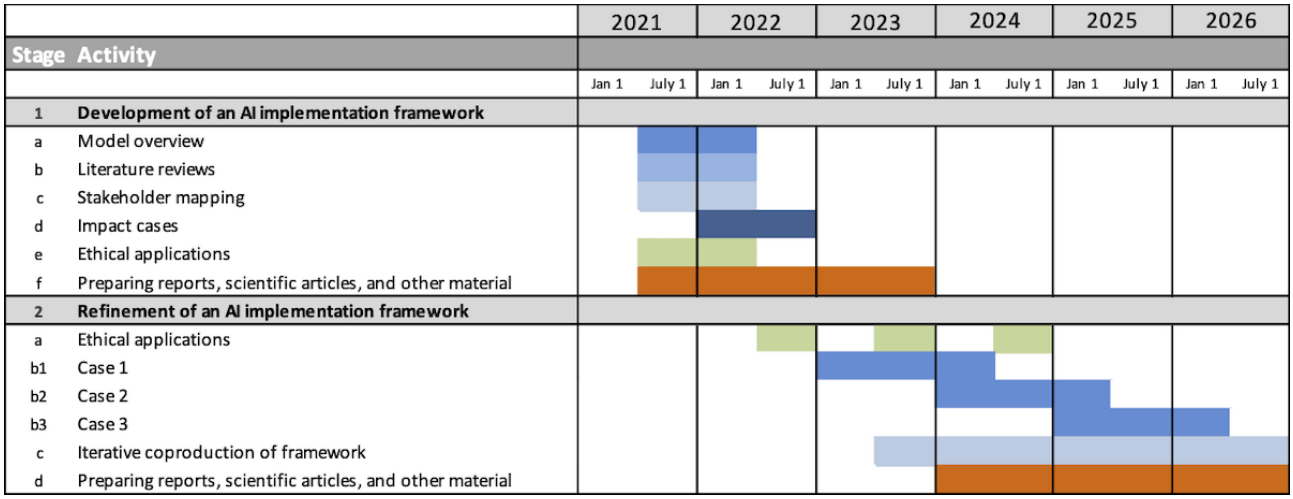
Ethical approval will be applied for regarding all empirical parts of the program. Site-specific approvals will be obtained for each site prior to commencing study activities. The study conforms to the principles outlined in the Declaration of Helsinki [46] and will fulfill the following requirements for research: information, consent, confidentiality, and safety of the participants guided by the ethical principles of autonomy, beneficence, nonmaleficence, and justice. All participants will receive written and verbal information about studies in which they are directly or indirectly involved. Participants will also be given information about the voluntary nature of the studies, confidentiality, and the ability to withdraw their consent at any time without having to justify why. All personal data will be registered according to the General Data Protection Regulation (GDPR2016/679), and the data will be stored in accordance with the Archive Act in Sweden (SFS1990:782).

The results of this program will be communicated to the included participants and partners, and key findings will be fed back to sites to enable refinement of strategies for implementation of AI systems in health care. The results will be disseminated via publications in peer-reviewed journals and presentations at national and international conferences.

Results

The purpose of Stages 1 and 2 is to develop a framework that specifically describes how implementation of AI systems in health care should be approached to increase the likelihood of successful implementation and use in routine health care practice. In Stage 1, starting in July 2021, the first version of the implementation framework started to be developed based on the initial investigation of current knowledge on implementation and improvement frameworks and models in health care (Figure 2). The evolution of the framework will thereafter be guided by findings from the literature reviews, stakeholder interviews, and impact cases to supplement it with current scientific evidence and relevant contextual conditions. In Stage 2, starting in July 2023, several implementation cases will be implemented with different sizes, time frames, and scopes of ambition (Figure 2). These implementation studies will be the means for developing the framework for AI systems implementation, which is, therefore, seen as the main product under Stage 2. The framework will be under continuous development during Stage 2 and will be reported in scientific articles and other formats for more practical application. The understanding and knowledge from Stage 2 can thereafter be used as the foundation for investments in a developed regional capacity to increase the practical resources, competencies, and organizational structure required for AI systems implementation, in collaboration with academia, industry partners, and health care. The planning for the integration of the outcomes from this research program will take place in parallel with the work on Stages 1 and 2 to allow for continuation and translation of the research to strategic investments in quality improvement in clinical practice (Stage 3 in Figure 1). This work will involve representatives from the university, the political and operational management teams of the regional and municipal health care systems, strategic partners from the business community, and representatives from various patient-specific interest groups.

Figure 2. Project timeline. AI: artificial intelligence.



Discussion

There is an immediate need to understand the health care service pathways and processes for AI systems to ensure that their impact in widespread practice fulfills the promise of transforming health care data into meaningful and actionable insights that support decision-making, optimize care processes, and provide high-quality patient care. Knowledge generated from implementation science and improvement science could be useful for understanding and developing methods, models, and frameworks needed to promote the uptake of research findings on AI systems into health care practices. The knowledge and expertise in implementation research in health care, in general, is advancing, but this is mainly held by research experts in the field. This knowledge is not widely accessible or used by people who could benefit from it, meaning that implementation efforts are not as successful as they could be, and common mistakes are repeated. There is a need to advance theory and empirical evidence regarding implementation requirements of AI systems in health care, and an opportunity to bring together insights from research on the development, introduction, and evaluation of AI systems and existing knowledge from implementation research literature. Therefore, with this research program, we intend to build an understanding, using both theoretical and empirical approaches, of how implementation of AI systems should be approached in order to increase the likelihood of successful and widespread application in clinical practice.

Acknowledgments

This project has received funding from the Swedish Innovation Agency and the Knowledge Foundation.

Conflicts of Interest

CM is National Professional Advisor for Patient Safety at the Care Quality Commission, UK.

References

1. Barlow J, Bayer S, Curry R. Implementing complex innovations in fluid multi-stakeholder environments: Experiences of 'telecare'. Technovation 2006 Mar;26(3):396-406. [doi: 10.1016/j.technovation.2005.06.010]
2. Webster A, Gardner J. Aligning technology and institutional readiness: The adoption of innovation. Technol Anal Strateg Manag 2019 Apr 02;31(10):1229-1241. [doi: 10.1080/09537325.2019.1601694]
3. Mehta N, Pandit A, Shukla S. Transforming healthcare with big data analytics and artificial intelligence: A systematic mapping study. J Biomed Inform 2019 Dec;100:103311 [FREE Full text] [doi: 10.1016/j.jbi.2019.103311] [Medline: 31629922]
4. De Nigris S, Craglia M, Nepelski D, Hradec J, Gomez-Gonzales E, Gomez Gutierrez E, et al. AI Watch: AI Uptake in Health and Healthcare 2020, EUR 30478 EN. Luxembourg: Publications Office of the European Union; 2020. URL: https://publications.jrc.ec.europa.eu/repository/handle/JRC122675 [accessed 2022-02-19]
5. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
6. Shaw J, Rudzicz F, Jamieson T, Goldfarb A. Artificial intelligence and the implementation challenge. J Med Internet Res 2019 Jul 10;21(7):e13659 [FREE Full text] [doi: 10.2196/13659] [Medline: 31293245]
7. Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial intelligence in health care: Bibliometric analysis. J Med Internet Res 2020 Jul 29;22(7):e18228 [FREE Full text] [doi: 10.2196/18228] [Medline: 32723713]

8. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019 Jan;25(1):30-36 [FREE Full text] [doi: 10.1038/s41591-018-0307-0] [Medline: 30617336]
9. Murphy K, Di Ruggiero E, Upshur R, Willison DJ, Malhotra N, Cai JC, et al. Artificial intelligence for good health: A scoping review of the ethics literature. BMC Med Ethics 2021 Feb 15;22(1):14 [FREE Full text] [doi: 10.1186/s12910-021-00577-8] [Medline: 33588803]
10. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: Past, present and future. Stroke Vasc Neurol 2017 Dec;2(4):230-243 [FREE Full text] [doi: 10.1136/svn-2017-000101] [Medline: 29507784]
11. Horgan D, Romao M, Morré SA, Kalra D. Artificial intelligence: Power for civilisation - and for better healthcare. Public Health Genomics 2019;22(5-6):145-161 [FREE Full text] [doi: 10.1159/000504785] [Medline: 31838476]
12. Alhashmi S, Alshurideh M, Kurdi B, Salloum S. A systematic review of the factors affecting the artificial intelligence implementation in the health care sector. In: Proceedings of the 1st International Conference on Artificial Intelligence and Computer Vision. Cham, Switzerland: Springer; 2020 Presented at: 1st International Conference on Artificial Intelligence and Computer Vision; April 8-9, 2020; Cairo, Egypt p. 37-49. [doi: 10.1007/978-3-030-44289-7_4]
13. Lin SY, Mahoney MR, Sinsky CA. Ten ways artificial intelligence will transform primary care. J Gen Intern Med 2019 Aug;34(8):1626-1630 [FREE Full text] [doi: 10.1007/s11606-019-05035-1] [Medline: 31090027]
14. Recht MP, Dewey M, Dreyer K, Langlotz C, Niessen W, Prainsack B, et al. Integrating artificial intelligence into the clinical practice of radiology: Challenges and recommendations. Eur Radiol 2020 Jun;30(6):3576-3584. [doi: 10.1007/s00330-020-06672-5] [Medline: 32064565]
15. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: Focus on clinicians. J Med Internet Res 2020 Jun 19;22(6):e15154 [FREE Full text] [doi: 10.2196/15154] [Medline: 32558657]
16. Gooding P, Kariotis T. Ethics and law in research on algorithmic and data-driven technology in mental health care: Scoping review. JMIR Ment Health 2021 Jun 10;8(6):e24668 [FREE Full text] [doi: 10.2196/24668] [Medline: 34110297]
17. Beil M, Proft I, van Heerden D, Sviri S, van Heerden PV. Ethical considerations about artificial intelligence for prognostication in intensive care. Intensive Care Med Exp 2019 Dec 10;7(1):70 [FREE Full text] [doi: 10.1186/s40635-019-0286-6] [Medline: 31823128]
18. Choudhury A, Asan O. Role of artificial intelligence in patient safety outcomes: Systematic literature review. JMIR Med Inform 2020 Jul 24;8(7):e18599 [FREE Full text] [doi: 10.2196/18599] [Medline: 32706688]
19. Fernandes M, Vieira SM, Leite F, Palos C, Finkelstein S, Sousa JM. Clinical decision support systems for triage in the emergency department using intelligent systems: A review. Artif Intell Med 2020 Jan;102:101762. [doi: 10.1016/j.artmed.2019.101762] [Medline: 31980099]
20. Yin J, Ngiam KY, Teo HH. Role of artificial intelligence applications in real-life clinical practice: Systematic review. J Med Internet Res 2021 Apr 22;23(4):e25759 [FREE Full text] [doi: 10.2196/25759] [Medline: 33885365]
21. Wolff J, Pauling J, Keck A, Baumbach J. The economic impact of artificial intelligence in health care: Systematic review. J Med Internet Res 2020 Feb 20;22(2):e16866 [FREE Full text] [doi: 10.2196/16866] [Medline: 32130134]
22. Gama F, Tyskbo D, Nygren J, Barlow J, Reed J, Svedberg P. Implementation frameworks for artificial intelligence translation into health care practice: Scoping review. J Med Internet Res 2022 Jan 27;24(1):e32215 [FREE Full text] [doi: 10.2196/32215] [Medline: 35084349]
23. Macrae C. Learning from the failure of autonomous and intelligent systems: Accidents, safety, and sociotechnical sources of risk. Risk Anal 2021 Nov 23:1-27 [FREE Full text] [doi: 10.1111/risa.13850] [Medline: 34814229]
24. Eldh AC, Almost J, DeCorby-Watson K, Gifford W, Harvey G, Hasson H, et al. Clinical interventions, implementation interventions, and the potential greyness in between: A discussion paper. BMC Health Serv Res 2017 Jan 07;17(1):16 [FREE Full text] [doi: 10.1186/s12913-016-1958-5] [Medline: 28061856]
25. Lennox L, Maher L, Reed J. Navigating the sustainability landscape: A systematic review of sustainability approaches in healthcare. Implement Sci 2018 Feb 09;13(1):27 [FREE Full text] [doi: 10.1186/s13012-017-0707-4] [Medline: 29426341]
26. Miake-Lye I, Mak S, Lam CA, Lambert-Kerzner AC, Delevan D, Olmos-Ochoa T, et al. Scaling beyond early adopters: A content analysis of literature and key informant perspectives. J Gen Intern Med 2021 Feb;36(2):383-395 [FREE Full text] [doi: 10.1007/s11606-020-06142-0] [Medline: 33111242]
27. Kirchner JE, Smith JL, Powell BJ, Waltz TJ, Proctor EK. Getting a clinical innovation into practice: An introduction to implementation strategies. Psychiatry Res 2020 Jan;283:112467 [FREE Full text] [doi: 10.1016/j.psychres.2019.06.042] [Medline: 31488332]
28. von Thiele Schwarz U, Nielsen K, Edwards K, Hasson H, Ipsen C, Savage C, et al. How to design, implement and evaluate organizational interventions for maximum impact: The Sigtuna Principles. Eur J Work Organ Psychol 2020 Aug 26;30(3):415-427. [doi: 10.1080/1359432x.2020.1803960]
29. Reed JE, Howe C, Doyle C, Bell D. Simple rules for evidence translation in complex systems: A qualitative study. BMC Med 2018 Jun 20;16(1):92 [FREE Full text] [doi: 10.1186/s12916-018-1076-9] [Medline: 29921274]
30. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opipari-Arrigan L, et al. Coproduction of healthcare service. BMJ Qual Saf 2016 Jul;25(7):509-517 [FREE Full text] [doi: 10.1136/bmjqs-2015-004315] [Medline: 26376674]

31. Lysaght T, Lim HY, Xafis V, Ngiam KY. AI-assisted decision-making in healthcare: The application of an ethics framework for big data in health and research. Asian Bioeth Rev 2019 Sep;11(3):299-314 [FREE Full text] [doi: 10.1007/s41649-019-00096-0] [Medline: 33717318]
32. Macrae C. Governing the safety of artificial intelligence in healthcare. BMJ Qual Saf 2019 Jun;28(6):495-498. [doi: 10.1136/bmjqs-2019-009484] [Medline: 30979783]
33. Ashfaq A, Lönn S, Nilsson H, Eriksson JA, Kwatra J, Yasin ZM, et al. Data Resource Profile: Regional healthcare information platform in Halland, Sweden. Int J Epidemiol 2020 Jun 01;49(3):738-739f. [doi: 10.1093/ije/dyz262] [Medline: 31930310]
34. McLaughlin JA, Jordan GB. Logic models: A tool for telling your programs performance story. Eval Program Plann 1999 Mar;22(1):65-72. [doi: 10.1016/s0149-7189(98)00042-1]
35. McMeekin N, Wu O, Germeni E, Briggs A. How methodological frameworks are being developed: Evidence from a scoping review. BMC Med Res Methodol 2020 Jun 30;20(1):173 [FREE Full text] [doi: 10.1186/s12874-020-01061-4] [Medline: 32605535]
36. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
37. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]
38. Reed J, Howe C, Doyle C, Bell D. Successful Healthcare Improvements From Translating Evidence in complex systems (SHIFT-Evidence): Simple rules to guide practice and research. Int J Qual Health Care 2019 Apr 01;31(3):238-244 [FREE Full text] [doi: 10.1093/intqhc/mzy160] [Medline: 30085160]
39. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
40. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
41. 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]
42. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
43. Wärnestål P, Nygren JM. Building an experience framework for a digital peer support service for children surviving from cancer. In: Proceedings of the 12th International Conference on Interaction Design and Children. 2013 Presented at: 12th International Conference on Interaction Design and Children; June 24-27, 2013; New York, NY p. 269-272. [doi: 10.1145/2485760.2485794]
44. Kolko J. Exposing the Magic of Design: A Practitioner's Guide to the Methods and Theory of Synthesis. Oxford, UK: Oxford University Press; 2010.
45. Charmaz K. Constructing Grounded Theory. 2nd edition. London, UK: SAGE Publications; 2014.
46. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. 2018. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2022-02-14]

Abbreviations

AI: artificial intelligence
CAISR: Center for Applied Intelligent Systems Research
CIDD: Center for Information-Driven Care
HDC: Health Data Center
HU: Halmstad University
IT: information technology
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability
PECI: process evaluation of complex interventions
RH: Region Halland
RHIP: Regional Healthcare Information Platform
SHIFT-Evidence: Successful Healthcare Improvements From Translating Evidence in Complex Systems

Edited by G Eysenbach; submitted 12.11.21; peer-reviewed by S Nilsson, C Kirkland; comments to author 03.01.22; revised version received 14.01.22; accepted 19.01.22; published 09.03.22.

Please cite as:
Svedberg P, Reed J, Nilsen P, Barlow J, Macrae C, Nygren J
Toward Successful Implementation of Artificial Intelligence in Health Care Practice: Protocol for a Research Program
JMIR Res Protoc 2022;11(3):e34920
URL: https://www.researchprotocols.org/2022/3/e34920
doi:10.2196/34920
PMID:35262500

Protocol

Level of Physical Activity, Sedentary Behavior, and Sleep in the Child and Adolescent Population in the Autonomous Community of the Basque Country (6-17 Years Old): Protocol for the Mugikertu Study

Arkaitz Larrinaga-Undabarrena[1,2*], MSc; Neritzel Albisua[3,4*], PhD; Xabier Río[1,2*], PhD; Garazi Angulo-Garay[1], MSc; Xabier González-Santamaria[1*], MSc; Iker Etxeberria Atxa[4*], MSc; Gorka Martínez de Lahidalga Aguirre[4*], MSc; Malen Ruiz de Azua Larrinaga[4*], MSc; Aitor Martínez Aguirre-Betolaza[1,5*], PhD; Ilargi Gorostegi-Anduaga[5*], PhD; Sara Maldonado-Martín[5*], PhD; Juan Aldaz Arregui[6*], PhD; Myriam Guerra-Balic[7*], PhD; Mikel Bringas[8*], PhD; José Ramón Sánchez Isla[9,10*], PhD; Aitor Coca[1*], PhD

[1]Faculty of Education and Sport, University of Deusto, Bilbao, Spain

[2]Department of Physical Activity and Health, Osasuna Mugimendua Kontrola SL Mugikon, Bilbao, Spain

[3]Mondragón Unibertsitatea, Arrasate, Spain

[4]Athlon Cooperative Society, Arrasate, Spain

[5]Faculty of Education and Sport, University of the Basque Country (UPV/EHU), Vitoria, Spain

[6]Faculty of Law, University of the Basque Country (UPV/EHU), Vitoria, Spain

[7]Faculty of Psychology, Education and Sport Sciences - Blanquerna, University Ramon Llull, Barcelona, Spain

[8]Directorate of Physical Activity and Sports, Basque Government, Vitoria, Spain

[9]Health and Consumption Department, Bilbao Council/Ayuntamiento de Bilbao, Bilbao, Spain

[10]Faculty of Health and Sciences, University of Deusto, Bilbao, Spain

[*]these authors contributed equally

Corresponding Author:
Arkaitz Larrinaga-Undabarrena, MSc
Faculty of Education and Sport
University of Deusto
Avenida de las Universidades, 24
Bilbao, 48007
Spain
Phone: 34 944139003 ext 3411
Email: a.larrinaga@deusto.es

Abstract

Background: Physical inactivity and sedentary behavior are increasingly common problems in the general population, which can lead to overweight, obesity, diabetes, cardiovascular disease, and decreased motor and cognitive capacity among children and adolescents. Establishing healthy habits in childhood on the basis of the World Health Organization's 2020 Physical Activity Guidelines is essential for proper physical, motor, and cognitive development.

Objective: The primary aim of this study is to describe the level of physical activity (PA), sedentary behavior, and sleep of the child and adolescent population from 6 to 17 years of age in the Basque Autonomous Community (BAC). Our secondary aim is to establish a starting point for future research and intervention protocols to improve the existing reality.

Methods: This cross-sectional study aims to recruit 1111 children and adolescents, aged 6 to 17 years from the BAC in a representative random sample. Participants will wear the ActiGraph WGT3X-BT triaxial accelerometer for 7 consecutive days in their nondominant wrist, and fill out a habit diary log of PA, mobility, and sleep routine. PA intensities, sedentary behavior, and sleep parameters (total bedtime, total sleep time, and sleep efficiency) will be calculated from raw accelerometer data using SPSS (IBM Corp). Participants will be randomly selected.

Results: The results of this study intend to demonstrate significant differences in PA levels in different age and gender groups since the volume of school PA in the BAC decreases as the age of the schoolchildren increases. The total study sample includes

1111 participants. In April 2021, up to 50% of the sample size was reached, which is expected to increase to 100% by April 2022. This sample will allow us to analyze, discuss, compare, and assess the reality of the school population, in a sensitive period of adherence to behavior patterns, using data from the geographical and administrative area of the BAC. This study will provide a realistic insight into PA levels among children and adolescents in the BAC. It will also offer scientific contributions on the positive relationship between PA levels and sleep quality in this population.

Conclusions: This study might highlight the need for the promotion of cross-sectional policies so that children and adolescents may increase their levels of PA, thus improving both the school environment and positive healthy behavior.

Trial Registration: ISRCTN Registry ISRCTN65573865; https://www.isrctn.com/ISRCTN65573865

International Registered Report Identifier (IRRID): DERR1-10.2196/31325

(JMIR Res Protoc 2022;11(3):e31325) doi:10.2196/31325

KEYWORDS

physical activity; sedentary behavior; sleep; Basque Autonomous Community; accelerometry; adolescents; children; healthy behavior; mobility

Introduction

Background

Physical inactivity (PI) and sedentary behavior are highly prevalent around the world and are associated with high morbidity and mortality [1]. In recent years, there has been an increase in the levels of sedentary lifestyle among the child and adolescent population, which is associated with lifestyle and has an impact on increasing obesity, diabetes, cardiovascular disease, cognitive capacities, and mental health [2,3].

In children and adolescents, physical activity (PA) improves physical fitness (both cardiorespiratory and muscular conditions), cardiometabolic health (blood pressure, dyslipidemia, obesity, glucose, and insulin resistance), bone health, cognitive ability (academic performance and executive function), and mental health (reducing symptoms of depression) [2,4,5].

The World Health Organization (WHO) recommends that children and adolescents between 5 and 17 years of age engage in at least an average of 60 minutes of PA per day, involving moderate to vigorous intensity (MVPA), mainly aerobic PA, with a minimum of 3 days a week of vigorous-intensity PA to limit sedentary behavior [2].

In Europe, approximately 80% of children and adolescents do not comply with the daily PA recommendations indicated by the WHO [6-8]. Hence, PI and sedentary lifestyle require intervention from an early age. If not, it may be too late as PA behaviors established in youth are maintained in adulthood [9]. It is estimated that in 2013, PI cost global health systems US $53.8 billion in direct health care, of which US $31.2 billion was allocated to the public sector, US $12.9 billion to the private sector, and US $9.7 billion for households [10]. However, it seems clear that intervention to promote PA must take place at an early age, since establishing habits and behavior changes are much easier and more achievable in childhood and adolescence than in adulthood [11].

In addition, PA is bidirectionally related to a shorter duration of sleep and an improvement in sleep efficiency, since PA contributes to an increase in sleep patterns [12]. Hence, physical exercise is already starting to be prescribed to improve sleep quality [13,14]. However, more national- and international-level studies are required among children and adolescents with objective measures to corroborate these findings [12].

Interventions promoting PA among children and adolescents must be supported by evidence that provides us the necessary information to do so efficiently. This study will allow us to analyze different realities (at different times, school settings, and ages), ultimately facilitating the design of policies and programs to promote PA and health on the basis of stated evidence [15]. Moreover, in the Basque Autonomous Community (BAC), there is a lack of evidence regarding the level of PA among children [16].

To determine the level of PA in the school population aged between 6 and 17 years in the BAC, and considering the benefits of PA for children and adolescents [17,18], this protocol aims to determine whether this group meets the established parameters of daily PA recommended by the WHO [2]. We hypothesize that >50% of the children do not meet the WHO recommendations, and that the level of PA in boys is greater than that in girls.

Objectives

The main objective of this study is to describe the level of PA, sedentary behavior, and sleep of the child and adolescent population aged 6 to 17 years in the BAC.

The secondary objectives of this study are to (1) contribute to the theoretical, conceptual, and methodological development of research devoted to the study of healthy behavior and well-being of school-age children; (2) monitor and compare healthy behaviors and characteristics of the social contexts in which school-age children develop; (3) develop collaboration with organizations and associations to activate initiatives aimed at promoting health among the school population; (4) promote and support the creation of a network of local professionals to integrate active and healthy behaviors during childhood and adolescence; and (5) strengthen the international research network in the field of promoting PA among school-age children.

Methods

Methods Overview

This protocol shows potential in terms of collaborative, interorganizational, and collaborative work among the 3 Basque universities (Basque Country University, Mondragon University, and University of Deusto) along with 2 companies (Athlon Cooperative Society and Osasuna Mugimendua Kontrola Ltd). They are innovative and proactive in the field of PA program intervention for health among different groups and social strata, creating new educational programs at the universities as well as new socio-sanitary interventions prescribing PA for different populations. Thus, they work for transversal and interdepartmental support of the PA and sports, educational innovation, and public health departments of the Basque government.

Participants and Selection Criteria

According to data extracted from the Basque Statistics Institute (EUSTAT) [19], as of October 16, 2020, the reference population to extract the sample residents in the BAC, aged between 6 and 17 years (born between 2003 and 2014), is 254,093 people, of whom 130,645 are boys and 123,448 are girls.

The corresponding sample size, considering that we are referencing a universal population greater than 100,000 people (referred as infinite for sample size calculation) is 1111 participants, at a confidence level of 2σ of 95.5%, an error limit of 3%, and heterogeneity of 50% [20]. To ensure a proportionate distribution of the sample in all age groups, territories, public and private centers, as well as girls and boys, these 1111 people were selected in accordance with the distribution detailed in Table 1. This sample is defined for all the assessment tests that comprise the project.

There is a proportional and random stratification based on province and county, age, gender, education network (public or private), and medea (socioeconomic level based on the deprivation index per census section, which enables the identification of sections with socioeconomic conditions), along with some inclusion and exclusion criteria that are listed in Textbox 1. Schools were selected randomly on the basis of the above-mentioned criteria.

Table 1. Sample size and the current sample (N=1111).

Grade	Araba province				Bizkaia province				Gipuzkoa province				Total			
	Male		Female		Male		Female		Male		Female		Male		Female	
	Sample, n	Current sample, n	Sample, n	Current sample, n	Sample, n	Current sample, n	Sample, n	Current sample, n	Sample, n	Current sample, n	Sample, n	Current sample, n	Sample, n	Current sample, n	Sample, n	Current sample, n
First primary	8	9	7	8	22	6	21	6	14	1	14	1	44	16	43	15
Second primary	7	8	7	8	22	5	21	5	15	1	14	1	44	14	42	14
Third primary	8	8	7	7	23	5	22	5	16	1	15	1	48	14	44	13
Fourth primary	8	8	8	8	24	6	23	6	16	1	16	1	48	15	47	15
Fifth primary	8	6	8	7	24	6	23	6	17	0	16	1	49	12	47	14
Sixth primary	8	6	7	6	24	6	23	8	17	1	16	0	49	13	45	14
First secondary	8	7	7	6	25	4	24	5	17	0	16	1	50	11	47	12
Second secondary	7	6	7	6	25	4	22	5	17	1	16	0	48	11	46	11
Third secondary	7	8	7	7	24	4	22	10	17	0	16	1	48	12	45	18
Fourth secondary	7	8	7	8	24	10	23	8	17	1	16	0	47	19	45	16
First high school	7	8	7	7	24	8	23	14	16	1	16	1	48	17	45	22
Second high school	7	8	7	7	24	3	22	3	17	1	15	1	48	12	44	11
Total	90	90	86	85	285	67	269	81	196	9	186	9	571	166	540	175

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Belonging to the student body of a participating school or institute
- Having the authorization to participate through a signed informed consent form by the parents or legal guardians of the child or adolescent

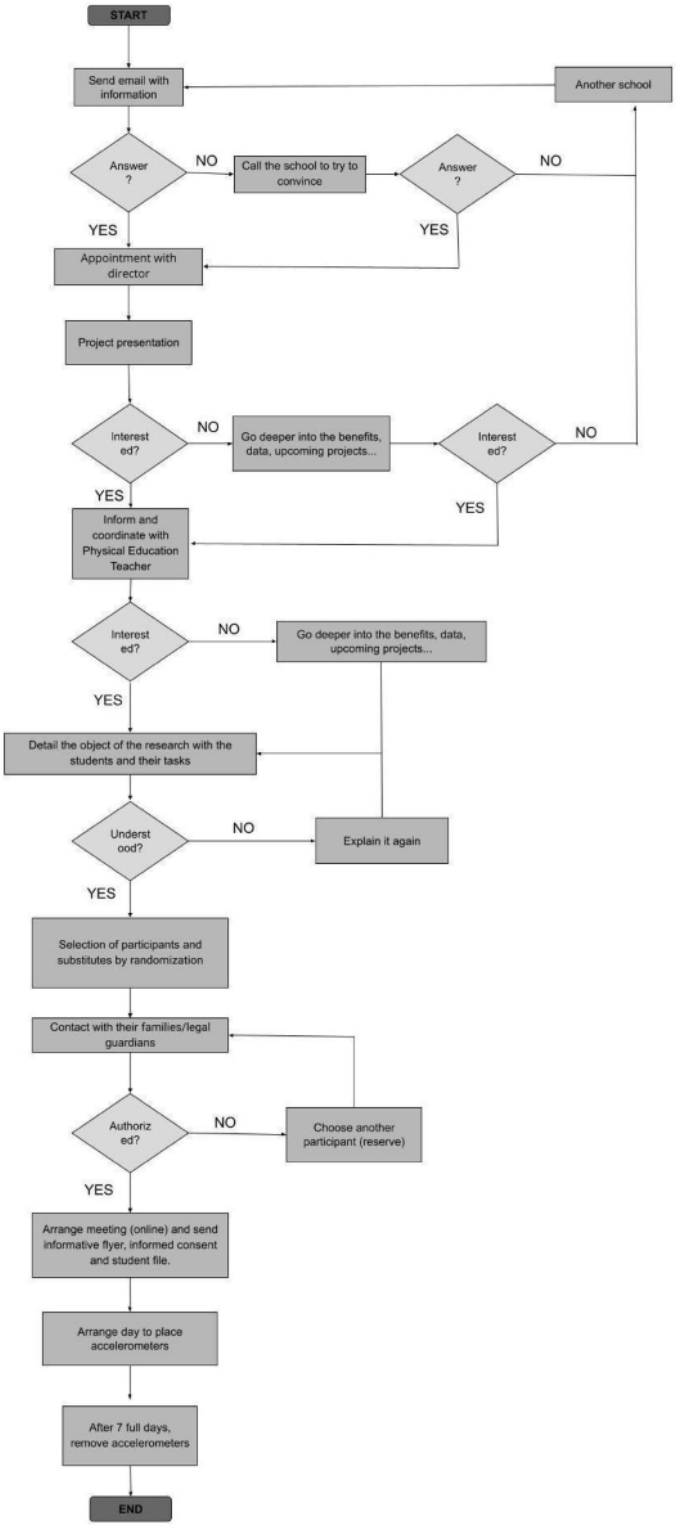
Exclusion criteria:

- Nonconsent or refusal by the child or adolescent to complete the physical activity diary or use the accelerometer, even with signed informed consent by their parents or legal guardians
- Physical or intellectual disability that prevents completing the daily physical activity or use of the accelerometer in accordance with the defined protocol. Each case will be assessed with each school's teaching team and with the parents or legal guardians of the minor

To reach 1111 individuals, the sample unit selection will be conducted in multiple stages, so that a random sample of all age strata will be obtained, and a subsequent selection (by randomization) will be made to meet the needs of age, gender, and type of established center for the project. The recruitment process can be divided into three phases: contact with the center, contact with the families, and contact with the students. The contact with the centers is initiated via email or a telephone call.

If the school shows interest, we arrange a meeting with the school's principal. The principal selects a person from the school who then becomes the person in charge of the selection process and sending the information to the parents, usually the physical education teacher. Depending on the requirements of the total sample, the age ranges that need to be covered in each school are selected. During this phase, the physical education teachers of each center will be contacted, as well as the counselors for each grade level to request their collaboration and thus carry out an initial preselection of students to participate in the study (Figure 1). The information and the documents to be filled out are sent to the parents. A lottery is held among the participants who meet the requirements (age, gender, and others) and whose parents have given their consent. If the principal, the physical education teacher, the participants, or their parents have queries about the study, they are clarified. A day is set for placing the accelerometers on the selected participants, and after 7 full days they are removed.

Figure 1. Flow chart of the study.



Instrument

ActiGraph WGT3X-BT accelerometers were chosen to measure MVPA and sedentary time behavior. ActiGraph sensors have been widely studied and have shown adequate reproducibility, validity, and feasibility for children and adolescents [21].

Participants will wear the ActiGraph WGT3X-BT accelerometer for 7 full, consecutive days, including weekdays and weekends. It will be worn on the wrist of the nondominant hand. This technique is the most reliable method to record and store the amount of PA quantitatively, allowing us to determine and compare levels of PA demonstrated by each individual in a given period [4,15]. In addition, like other authors, diary recordings will be collected with data such as sleep time, mode of transportation to school or other places, and the kind of PA they engage in [22-24].

To conduct this protocol, 50 ActiGraph WGT3X-BT accelerometers were purchased along with the ActiLife 6 ActiGraph PA software. All the accessories necessary for its use were also purchased, including Velcro wrist straps and cables for data transfer. The accelerometer will be worn on the nondominant wrist for 7 full, consecutive days (worn on day 1 and removed on day 8) including weekdays and weekends. The accelerometer consists of a triaxial activity monitor to capture and record continuous, high-resolution PA and sleep and wake information that provides the following measurements: acceleration, energy expenditure, steps, PA intensity, body position, and lighting ambient.

Data Collection

Each participant will be identified by a unique and confidential numerical code, which is assigned after informed consent is obtained from the legal guardians. Those responsible for the database must save the files on servers with a password. To determine the codes, the following process is followed:

- Province: 1=Araba, 2=Bizkaia, and 3=Gipuzkoa.
- Town (000)—each town is assigned a 3-digit number (eg, 001, 002, etc)
- Education center is assigned a 3-digit number (eg, 001, 002, etc).
- The academic stage between primary, secondary, or high school or secondary vocational training (1, 2, and 3)
- The grade
- The classroom in letters A, B, C, and D
- The student number (00) considering that each classroom will have a maximum of 30 students.

Consequently, each participant will have a 12-digit code; for example, 2 022 060 11 D 07.

Prior to data collection, the study presentation brochure was emailed to the centers inviting them to participate, along with the informed consent document to authorize their children to participate in the study. Each center shares information about the study with the families through the established communication channel in each case (web platform or email), sending the legal guardians the information received at the center.

If families are interested in participating, they must send their informed consent to the head of the education center, and depending on the number of volunteers, a random draw will be held until the sample requirements established in each area and center are reached.

Programming and Downloading Accelerometry Data

The computer operating system, compatible with ActiGraph software, has the software license to download the data. Before starting data logging with the accelerometer, each device will be turned on or programmed for data logging.

To program the accelerometer, the participant's date of birth, gender, height (cm), body mass (kg), and dominant hand must first be recorded. These data will be requested after receiving informed consent. It will be important to inform the participant how to wear the accelerometer before putting it on. The accelerometer can be placed either in contact with the skin or over a sleeve in case of allergies by touch with the parts of the device [25].

The accelerometer is placed on the nondominant wrist without being too tight so that it is not uncomfortable, nor so loose that it moves around the wrist. The part of the button to close the band should be facing toward the fingers of the hand (as if it were a watch). In addition, instructions with photographs will be sent to explain the placement of the accelerometers. The accelerometer should be worn all day and night and should only be removed during activities where the accelerometer could be submerged in water (bathing, showering, water activities, etc). In these situations, it must be removed and put back on the same wrist after the activity is finished.

During the week that the accelerometer is worn, participants must complete a diary to record their hours of daily activity: the time they wake up, nap, sleep, mode, and duration of travel to class, mode and duration of travel to other places, physical education class, scheduled and unscheduled PA, the time when the accelerometer was removed, and circumstances that may have affected their sleep/wake pattern (travel, illness, etc).

With the youngest children, it is important to understand that the diary will often be completed by the parents, guardians or even their teachers. Similarly, from the age of 8 years, children should be encouraged to help their parents or guardians to complete the diary. However, adolescents will be encouraged to complete their diaries themselves, as their parents are often unaware of their sleep routines [25].

Once the full 7 days have elapsed, on the agreed upon date, the accelerometer will be delivered to the person in charge (the teacher or researcher) along with the daily activity log sheet, only including the participant´s code.

Once the accelerometer and the daily activity record sheet have been collected, the data will be downloaded [26]. As this study is conducted with children, the originals will be converted to 60 seconds to calculate sleep [27] and to 5 seconds for daily PA [27,28].

Study Variables

The following study variables will be recorded: (1) active time and sedentary behavior (in the overall school schedule, physical education classes, school recess time, activities related to school sports and other organized and unorganized exercise, the rest of the time during the day while awake, school days, nonschool days, bouts of light PA, moderate PA, vigorous PA, and MVPA); (2) time asleep (school days, nonschool days, and bouts of sleeping time); (3) type of transportation to school (bus, car, motorcycle, or any other means that do not involve active travel, and time or day of using transportation); (4) type of transportation to places other than school (bus, car, motorcycle, or any other means that do not involve active travel, walking, cycling, or skating); and (5) grade students are attending, BMI, type of centers, location, and socioeconomic status.

Ethical Considerations

This study was approved by the Euskadi Drug Research Ethics Committee (Basque Government Department of Health) in accordance with law 14/2007 on biomedical research, ethical principles of the Declaration of Helsinki, and other ethical principles and applicable legislation with internal code PI2020011. In turn, current regulations on the protection of personal data will be followed: Regulation (EU) 2016/679 of April 27, 2016 (GDPR), Organic Law 3/2018, of December 5 on Personal Data Protection and guaranteeing digital rights (ES), and Royal Decree (ES) 1720/2007 of December 21.

Statistical Analysis

The statistical analysis will be performed using SPSS (version 27.0.1.0; IBM Corp). Parametric tests will be performed after all assumptions for each test are met. For comparison between groups (girls vs boys, age groups, school types, and regions) 2-tailed *t* tests, 1-way analysis of variance, or the nonparametric method of Kruskal-Wallis and chi-square tests will be used for the main primary outcomes as dependent variables.

Descriptive statistics will be used for all outcome variables, and the effect size and the level of significance corresponding to the main group (between participants) will be reported. To prevent a type I error, post hoc comparison will be performed when a significant interaction effect is present. Values will be expressed as mean (SD). The significance level will be set at 95% (α=.05).

Results

The study began in September 2020, and the data collection will take place for the next 2 academic years (2021-2022). To obtain the total study sample (N=1111), 175 participants will be required in the Araba province, 554 in the Bizkaia province, and 382 in the Gipuzkoa province. In addition, these variables are subject to gender, age, center (public or private), location, and socioeconomic status.

There is currently a total of 341 samples among the three historical territories (Table 1). It is estimated that by the end of December 2021, up to 75% of the sample size will have been reached, with 100% potentially reached by April 2022. The impact of the current COVID-19 pandemic has resulted in schedule delays.

Discussion

Expected Findings

This study aims to provide current estimates of the levels of PA, sedentary behavior, and sleep among children and adolescents in the BAC (6-17 years old). Sedentary behavior is fundamentally important in the health of children and adolescents [29], and it is highly influenced by the social environment in which they live [30]. Recent statistics show that 28% of children in Spain between the ages of 3 and 8 years are overweight, and 9% are obese [31]. It is clear that increased PA and reducing sedentary behavior play a leading role in preventing overweight and excess adiposity, even when people present an adverse genetic condition that predisposes them to be obese [32]. Some studies seem to have shown an inverse relationship between MVPA with adiposity and cardiometabolic risk, and a positive relationship between MVPA with cardiorespiratory fitness and total body bone mineral density among children and adolescents [33]. Other studies maintain that screen time has a significant and inverse relationship with hours of sleep, unregulated activity (games and other activities), and physical exercise [34].

The results of this study intend to demonstrate significant differences in PA levels in different age groups since the volume of school PA in the BAC decreases as the age of the schoolchildren increases. In terms of gender, it is higher among boys than in girls, reaching the same conclusions as other authors [35,36].

The results will be used to analyze and assess the relevance of creating a battery of assessment of physical condition in the Basque Government (interdepartmental) as an improvement in the diagnosis of the state of physical condition and to be able to intervene from the department of education and the department of health to address, together with the Directorate of Physical Activity and Sport the concrete actions to be developed as process improvement and move in an improvement of physical conditions in future evaluations.

While there are some limitations, such as the current COVID-19 situation affecting the ability to recruit schools, this study has the potential to determine current levels of PA, sedentary behavior, and sleep of children and adolescents in the BAC. Comparing PA levels extracted in situations with and those without COVID-19 restrictions (with or without structured school sports, with or without the geographical limitation of municipal mobility, and with or without structured federated competition among other states that will affect the measurements gathered) can be established.

This study will make it possible to obtain data, analyze them, and discuss various options to promote cross-sectional policies so that children and adolescents increase their levels of PA, improving both the school environment and positive behavior. In turn, this can improve students' academic results [37]. The results of this study will provide a realistic insight into PA levels among children and adolescents in the BAC. In addition, our

results aim to offer a scientific contribution to the positive relationship between PA levels and sleep quality in this population.

Future Prospects

This study will enable further research on mobility and PA levels among different ages, examine, and analyze intervention protocols to improve health at education centers, and continue studying the relationship between PA levels and sleep in school-age children, in addition to looking for differences between PA levels and race.

Conclusions

This study might highlight the need for the promotion of cross-sectional policies so that children and adolescents may increase their levels of PA, improving both the school environment and positive healthy behavior.

Acknowledgments

This study has been performed with the support of the Basque government and all the centers. In addition, we express our gratitude to all the children, adolescents, and their families who have and will participate in the study.

Authors' Contributions

ALU, XR, GAG, and AC contributed to study conceptualization. ALU, XR, GAG, AC, SMM, JAA, IGA, NA, and AMAB contributed to the methodology and investigation. ALU, XR, and GAG contributed to the writing of the original draft. MB, NA, JRSI, and AC contributed to the project administration. GMLA, GAG, MRAL, XGS, and IEA contributed to the data curation. ALU, XR, GAG, MGB, and AC contributed to study supervision. All authors have read, provided feedback, and agreed to the final version of the manuscript for publication.

Conflicts of Interest

None declared.

References

1. Arocha Rodulfo JI. Sedentary lifestyle a disease from xxi century. Clin Investig Arterioscler 2019;31(5):233-240. [doi: 10.1016/j.arteri.2019.04.004] [Medline: 31221536]
2. Bull F, Al-Ansari S, Biddle S, Borodulin K, Buman M, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
3. Rey-López JP, Vicente-Rodríguez G, Biosca M, Moreno L. Sedentary behaviour and obesity development in children and adolescents. Nutr Metab Cardiovasc Dis 2008 Mar;18(3):242-251. [doi: 10.1016/j.numecd.2007.07.008] [Medline: 18083016]
4. Jiménez-Pavón D, Fernández-Vázquez A, Alexy U, Pedrero R, Cuenca-García M, Polito A, HELENA Study Group. Association of objectively measured physical activity with body components in European adolescents. BMC Public Health 2013 Jul 18;13:667 [FREE Full text] [doi: 10.1186/1471-2458-13-667] [Medline: 23866681]
5. Wu X, Han L, Zhang J, Luo S, Hu J, Sun K. The influence of physical activity, sedentary behavior on health-related quality of life among the general population of children and adolescents: A systematic review. PLoS One 2017;12(11):e0187668 [FREE Full text] [doi: 10.1371/journal.pone.0187668] [Medline: 29121640]
6. Global recommendations on physical activity for health. World Health Organization. 2010 Jan 01. URL: https://www.who.int/publications/i/item/9789241599979 [accessed 2022-02-22]
7. Hallal P, Andersen L, Bull F, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
8. Physical activity. World Health Organization. 2020 Nov 26. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2022-02-22]
9. Telama R, Yang X, Viikari J, Välimäki I, Wanne O, Raitakari O. Physical activity from childhood to adulthood: a 21-year tracking study. Am J Prev Med 2005 Apr;28(3):267-273. [doi: 10.1016/j.amepre.2004.12.003] [Medline: 15766614]
10. Ding D, Lawson K, Kolbe-Alexander T, Finkelstein E, Katzmarzyk P, van Mechelen W, Lancet Physical Activity Series 2 Executive Committee. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. Lancet 2016 Sep 24;388(10051):1311-1324. [doi: 10.1016/S0140-6736(16)30383-X] [Medline: 27475266]
11. Capuzzi D, Stauffer M. Human Growth and Development Across the Lifespan: Applications for Counselors. Nashville, TN: John Wiley & Sons; 2016.
12. Ávila-García M, Femia-Marzo P, Huertas-Delgado F, Tercedor P. Bidirectional Associations between Objective Physical Activity and Sleep Patterns in Spanish School Children. Int J Environ Res Public Health 2020 Jan 22;17(3):710 [FREE Full text] [doi: 10.3390/ijerph17030710] [Medline: 31979039]

13. Janssen X, Martin A, Hughes A, Hill C, Kotronoulas G, Hesketh K. Associations of screen time, sedentary time and physical activity with sleep in under 5s: A systematic review and meta-analysis. Sleep Med Rev 2020 Feb;49:101226 [FREE Full text] [doi: 10.1016/j.smrv.2019.101226] [Medline: 31778942]
14. Kredlow M, Capozzoli M, Hearon B, Calkins A, Otto M. The effects of physical activity on sleep: a meta-analytic review. J Behav Med 2015 Jun;38(3):427-449. [doi: 10.1007/s10865-015-9617-6] [Medline: 25596964]
15. Cain K, Sallis J, Conway T, Van Dyck D, Calhoon L. Using accelerometers in youth physical activity studies: a review of methods. J Phys Act Health 2013 Mar;10(3):437-450 [FREE Full text] [doi: 10.1123/jpah.10.3.437] [Medline: 23620392]
16. Active Healthy Kids Basque Country 2021 Report Card. Active Health Kids Global Alliance. URL: https://bit.ly/3pf9F7A[accessed [accessed 2021-12-01]
17. Sallis J, McKenzie T, Alcaraz J, Kolody B, Faucette N, Hovell M. The effects of a 2-year physical education program (SPARK) on physical activity and fitness in elementary school students. Sports, Play and Active Recreation for Kids. Am J Public Health 1997 Aug;87(8):1328-1334. [doi: 10.2105/ajph.87.8.1328] [Medline: 9279269]
18. Cooper A, Goodman A, Page A, Sherar L, Esliger D, van Sluijs EMF, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). Int J Behav Nutr Phys Act 2015 Sep 17;12:113 [FREE Full text] [doi: 10.1186/s12966-015-0274-5] [Medline: 26377803]
19. Población de la C.A. de Euskadi por año de nacimiento, según el territorio histórico y el sexo. 01/01/2021. Eustat. URL: https://www.eustat.eus/elementos/ele0011400/Poblacion_de_la_CA_de_Euskadi_por_ano_de_nacimiento_segun_el_territorio_historico_y_el_sexo/tbl0011424_c.html [accessed 2021-03-25]
20. González RA. Review of Metodología Cuantitativa. Estrategias y técnicas de investigación social, by Ma. A. C. D'Ancona. Reis 1997;80:240-424. [doi: 10.2307/40183928]
21. Sherar L, Griew P, Esliger D, Cooper A, Ekelund U, Judge K, et al. International children's accelerometry database (ICAD): design and methods. BMC Public Health 2011 Jun 21;11:485 [FREE Full text] [doi: 10.1186/1471-2458-11-485] [Medline: 21693008]
22. Laíño FA, Santa María CJ, Bazán NE, Salvia HA, Tuñón I. Validación de un cuestionario de actividad física en niños y adolescentes de distintos estratos socioeconómicos. Apunts 2017 Mar 30(127):35-43. [doi: 10.5672/apunts.2014-0983.es.(2017/1).127.03]
23. Laíño FA, Santa María CJ, Bazán NE, Salvia HA, Tuñón I. Validació d'un qüestionari d'activitat física en nens i adolescents de diferents estrats socioeconòmics. Apunts 2017 Mar 30(127):35-43. [doi: 10.5672/apunts.2014-0983.cat.(2017/1).127.03]
24. Manchola-González J, Bagur-Calafat C, Girabent-Farrés M. Fiabilidad de la versión española del Cuestionario de actividad física PAQ-C / Reliability of the Spanish Version of Questionnaire of Physical Activity PAQ-C. rimcafd 2017;65(2017). [doi: 10.15366/rimcafd2017.65.010]
25. Ancoli-Israel S, Martin JL, Blackwell T, Buenaver L, Liu L, Meltzer LJ, et al. The SBSM Guide to Actigraphy Monitoring: Clinical and Research Applications. Behav Sleep Med 2015;13 Suppl 1:S4-S38. [doi: 10.1080/15402002.2015.1046356] [Medline: 26273913]
26. Martinez Aguirre-Betolaza A, Mujika I, Loprinzi P, Corres P, Gorostegi-Anduaga I, Maldonado-Martín S. Physical Activity, Sedentary Behavior, and Sleep Quality in Adults with Primary Hypertension and Obesity before and after an Aerobic Exercise Program: EXERDIET-HTA Study. Life (Basel) 2020 Aug 17;10(8):153 [FREE Full text] [doi: 10.3390/life10080153] [Medline: 32824416]
27. McClain J, Abraham T, Brusseau TJ, Tudor-Locke C. Epoch length and accelerometer outputs in children: comparison to direct observation. Med Sci Sports Exerc 2008 Dec;40(12):2080-2087. [doi: 10.1249/MSS.0b013e3181824d98] [Medline: 18981941]
28. Dobell A, Eyre E, Tallis J, Chinapaw M, Altenburg T, Duncan M. Examining accelerometer validity for estimating physical activity in pre-schoolers during free-living activity. Scand J Med Sci Sports 2019 Oct;29(10):1618-1628. [doi: 10.1111/sms.13496] [Medline: 31206785]
29. van Ekris E, Altenburg T, Singh A, Proper K, Heymans M, Chinapaw M. An evidence-update on the prospective relationship between childhood sedentary behaviour and biomedical health indicators: a systematic review and meta-analysis. Obes Rev 2016 Sep;17(9):833-849. [doi: 10.1111/obr.12426] [Medline: 27256486]
30. Cantallops Ramón J, Ponseti Verdaguer FJ, Vidal Conti J, Borràs Rotger PA, Palou Sampol P. Adolescencia, sedentarismo y sobrepeso: análisis en función de variables sociopersonales de los padres y del tipo de deporte practicado por los hijos (Adolescence, physical inactivity and overweight: analysis based on socio-personal variables of the paren. Retos 2015 Mar 07(21):5-8. [doi: 10.47197/retos.v0i21.34595]
31. Gómez SF, Homs C, Wärnberg J, Medrano M, Gonzalez-Gross M, Gusi N, et al. Study protocol of a population-based cohort investigating Physical Activity, Sedentarism, lifestyles and Obesity in Spanish youth: the PASOS study. BMJ Open 2020 Sep 23;10(9):e036210 [FREE Full text] [doi: 10.1136/bmjopen-2019-036210] [Medline: 32967871]
32. Marrodán MD, Martínez-Álvarez JR, González-Montero De Espinosa M, López-Ejeda N, Cabañas MD, Prado C. [Diagnostic accuracy of waist to height ratio in screening of overweight and infant obesity]. Med Clin (Barc) 2013 Apr 15;140(7):296-301. [doi: 10.1016/j.medcli.2012.01.032] [Medline: 22726769]

33. García-Hermoso A, Ezzatvar Y, Ramírez-Vélez R, Olloquequi J, Izquierdo M. Is device-measured vigorous physical activity associated with health-related outcomes in children and adolescents? A systematic review and meta-analysis. J Sport Health Sci 2021 May;10(3):296-307 [FREE Full text] [doi: 10.1016/j.jshs.2020.12.001] [Medline: 33285309]
34. Amigo Vázquez I, Busto Zapico R, Herrero Díez J, Fernández Rodríguez C. [Physical activity, sedentary leisure, short sleeping and childhood overweight]. Psicothema 2008 Nov;20(4):516-520. [Medline: 18940044]
35. Rodríguiez-Fernandez J, Pereira B, Navarro-Patón R. Estudio del nivel de actividad físico-deportiva en escolares de educación primaria [Study of the level of physical-sports activity in primary school children]. In: Serrano J, Petrica M, Kerkoski M, Pereira B, editors. Educação física, saúde e bem-estar - novos caminhos. Castelo Branco, Portugal: Instituto Politécnico de Castelo Branco (IPCB); 2020:75-83.
36. Colley R, Garriguet D, Janssen I, Craig C, Clarke J, Tremblay M. Physical activity of Canadian children and youth: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. Health Rep 2011 Mar;22(1):15-23 [FREE Full text] [Medline: 21510586]
37. Caldarella P, Shatzer RH, Gray KM, Young KR, Young EL. The Effects of School-wide Positive Behavior Support on Middle School Climate and Student Outcomes. RMLE Online 2015 Aug 25;35(4):1-14. [doi: 10.1080/19404476.2011.11462087]

Abbreviations

BAC: Basque Autonomous Community
MVPA: moderate to vigorous physical activity
PA: physical activity
PI: physical inactivity
WHO: World Health Organization

Edited by G Eysenbach; submitted 17.06.21; peer-reviewed by C Carrion, D Filos, WF Khaw; comments to author 13.10.21; revised version received 03.12.21; accepted 22.12.21; published 11.03.22.

Please cite as:
Larrinaga-Undabarrena A, Albisua N, Río X, Angulo-Garay G, González-Santamaria X, Etxeberria Atxa I, Martínez de Lahidalga Aguirre G, Ruiz de Azua Larrinaga M, Martínez Aguirre-Betolaza A, Gorostegi-Anduaga I, Maldonado-Martín S, Aldaz Arregui J, Guerra-Balic M, Bringas M, Sánchez Isla JR, Coca A
Level of Physical Activity, Sedentary Behavior, and Sleep in the Child and Adolescent Population in the Autonomous Community of the Basque Country (6-17 Years Old): Protocol for the Mugikertu Study
JMIR Res Protoc 2022;11(3):e31325
URL: https://www.researchprotocols.org/2022/3/e31325
doi:10.2196/31325
PMID:35275088

Protocol

Rehabilitation Supported by Technology: Protocol for an International Cocreation and User Experience Study

Sylvie Bernaerts[1,2], PhD; Nele A J De Witte[1,2], PhD; Vicky Van der Auwera[1], MSc; Bert Bonroy[3], PhD; Luiza Muraru[3], PhD; Panagiotis Bamidis[4], PhD; Christos Frantzidis[4], PhD; Chrysoula Kourtidou-Papadeli[4,5], PhD; Nancy Azevedo[6], PhD; Jokin Garatea[7], MSc; Idoia Muñoz[7], MBA; Rosa Almeida[8], MSc; Raquel Losada[8], MSc; Joyce Fung[6,9], PhD; Eva Kehayia[6,9], PhD; Anouk Lamontagne[6,9], PhD; Elaine de Guise[6,10], PhD; Cyril Duclos[6,10], PhD; Johanne Higgins[6,10], PhD; Sylvie Nadeau[6,10], PhD; Lucie Beaudry[6,11], PhD; Evdokimos Konstantinidis[4,12], PhD

[1]LiCalab, Thomas More University of Applied Sciences, Geel, Belgium

[2]Expertise Unit Psychology, Technology & Society, Thomas More University of Applied Sciences, Antwerpen, Belgium

[3]Mobilab & Care, Thomas More University of Applied Sciences, Geel, Belgium

[4]Medical Physics and Digital Innovation Laboratory, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

[5]Greek Aerospace Medical Association and Space Research, Thessaloniki, Greece

[6]Centre de recherche interdisciplinaire en réadaptation du Montréal métropolitain, Institut Universitaire sur la réadaptation en déficience physique de Montréal, Montreal, QC, Canada

[7]GAIA, Asociación de Industrias de Conocimiento y Tecnologías Aplicadas, San Sebastian, Spain

[8]Fundación INTRAS, Research, Development and Innovation Department, Valladolid, Spain

[9]School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada

[10]Université de Montreal, Montreal, QC, Canada

[11]Dance Department, Université du Québec à Montréal, Montreal, QC, Canada

[12]European Network of Living Labs, Brussels, Belgium

Corresponding Author:
Sylvie Bernaerts, PhD
LiCalab
Thomas More University of Applied Sciences
Kleinhoefstraat 4
Geel, 2440
Belgium
Phone: 32 034324050
Email: sylvie.bernaerts@thomasmore.be

Abstract

Background: Living labs in the health and well-being domain have become increasingly common over the past decade but vary in available infrastructure, implemented study designs, and outcome measures. The Horizon 2020 Project *Virtual Health and Wellbeing Living Lab Infrastructure* aims to harmonize living lab procedures and open living lab infrastructures to facilitate and promote research activities in the health and well-being domain in Europe and beyond. This protocol will describe the design of a joint research activity, focusing on the use of innovative technology for both rehabilitation interventions and data collection in a rehabilitation context.

Objective: With this joint research activity, this study primarily aims to gain insight into each living lab's infrastructure and procedures to harmonize health and well-being living lab procedures and infrastructures in Europe and beyond, particularly in the context of rehabilitation. Secondarily, this study aims to investigate the potential of innovative technologies for rehabilitation through living lab methodologies.

Methods: This study has a mixed methods design comprising multiple phases. There are two main phases of data collection: cocreation (phase 1) and small-scale pilot studies (phase 2), which are preceded by a preliminary harmonization of procedures among the different international living labs. An intermediate phase further allows the implementation of minor adjustments to the intervention or protocol depending on the input that was obtained in the cocreation phase. A total of 6 small-scale pilot studies using innovative technologies for intervention or data collection will be performed across 4 countries. The target study sample

comprises patients with stroke and older adults with mild cognitive impairment. The third and final phases involve Delphi procedures to reach a consensus on harmonized procedures and protocols.

Results: Phase 1 data collection will begin in March 2022, and phase 2 data collection will begin in June 2022. Results will include the output of the cocreation sessions, small-scale pilot studies, and advice on harmonizing procedures and protocols for health and well-being living labs focusing on rehabilitation.

Conclusions: The knowledge gained by the execution of this research will lead to harmonized procedures and protocols in a rehabilitation context for health and well-being living labs in Europe and beyond. In addition to the harmonized procedures and protocols in rehabilitation, we will also be able to provide new insights for improving the implementation of innovative technologies in rehabilitation.

International Registered Report Identifier (IRRID): PRR1-10.2196/34537

(JMIR Res Protoc 2022;11(3):e34537) doi:10.2196/34537

KEYWORDS

cocreation; harmonization; living lab; rehabilitation; small-scale real-life testing; technology

Introduction

Background

Virtual Health and Wellbeing Living Lab Infrastructure

Living labs in the health and well-being domains have become increasingly common over the past decade but vary in available infrastructure, implemented study designs, and outcome measures. Therefore, increased transnational collaboration and harmonization can further improve research quality. The Horizon 2020 Project *Virtual Health and Wellbeing Living Lab Infrastructure* (VITALISE), funded by the European Union (under grant agreement 101007990; April 2021 to March 2024), unites 19 partners (AIT, AUTH, AV, CERTH, ENoLL, GAIA, INTRAS, LAUREA, LiCalab, LLSA, McGILL, SLIMMER, SIT, TREBAG, UdeM, UPM, VICOM, VILABS, and WITA) across 11 countries (Table 1). VITALISE aims to harmonize living lab procedures and open living lab infrastructures as a means to facilitate and promote research activities in the health and well-being domain in Europe and beyond. To do so, the VITALISE consortium will conduct joint research activities (JRAs) in the fields included in the care pathway of patients, namely, rehabilitation, transitional care, and everyday living environments for the older adults. These JRAs combine and capitalize on research experience and expertise from the different living labs in the consortium and create innovation test beds for the harmonized procedures and infrastructures in the context of health and well-being research. This protocol will describe the design of the rehabilitation JRA, focusing on the use of innovative technology for both rehabilitation interventions and data collection in a rehabilitation context. More information about the project can be found on the project website [1].

Table 1. List of partners in the VITALISE[a] project.

Abbreviation	Full (legal) name	Location
AIT	Austrian Institute of Technology GmbH	Austria
AUTH[b]	Aristotelio Panepistimio Thessalonikis	Greece
AV	Anthology Venture Ad	Bulgaria
CERTH	Center for Research and Technology Hellas (Ethniko Kentro Erevnas Kai Technologikis Anaptyxis)	Greece
CRIR[c]	Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal	Canada
ENoLL	European Network of Living Labs	Belgium
GAIA[b]	Association of Electronic and Information Technologies in the Basque Country	Spain
INTRAS[b]	Fundación INTRAS	Spain
LAUREA	Laurea-ammatikorkeakoulu Oy	Finland
LiCalab[b]	Living & Care Lab (Thomas More Kempen vzw)	Belgium
LLSA	Le Forum des Living Labs en Santé et Autonomie	France
McGILL[b]	McGill University	Canada
SLIMMER	Coöperatie Slimmer Leven 2020 U.A.	The Netherlands
SIT	Social IT Software & Consulting srl	Italy
TREBAG[b]	TREBAG Intellectual Property and Project Manager Ltd	Hungary
UdeM[b]	Université de Montreal	Canada
UQAM[c]	Université de Québec à Montréal	Canada
UPM	Universidad Politécnica de Madrid	Spain
VICOM	Fundación Centro De Tecnologías De Interacción Visual y Comunicaciones VICOMTECH	Spain
VILABS	VILABS OE	Greece
WITA	WITA S.r.l.	Italy

[a]VITALISE: Virtual Health and Wellbeing Living Lab Infrastructure.

[b]The living labs involved in the joint research activity.

[c]The living labs and organizations that participate as VITALISE consortium external partners.

Living Lab Methodology

Living labs are defined as user-centered, open innovation ecosystems based on a systematic user cocreation approach integrating research and innovation processes in real-life communities and settings. In practice, living labs are research infrastructures that place the citizen at the center of innovation and have thus shown the ability to better exploit the opportunities offered by new information and communication technology concepts and solutions to the specific needs and aspirations of local contexts, cultures, and creativity potentials [2]. In the context of rehabilitation, these infrastructures can range from research labs in universities (of applied sciences), over rehabilitation centers or rehabilitation departments of (health) care facilities, to smart living homes and individual's home environments. Living labs can thus be defined as open innovation systems that aim to enable innovation through research activities in realistic circumstances with multiple stakeholders from the quadruple helix (university, industry, government, and public actors) [3,4]. They not only provide access to research infrastructures and end user populations to help tailor innovations to the needs of the local context but also facilitate international upscaling of innovations through cross-border research and collaboration within international networks such as the ENoLL. The living lab methodology is grounded in an iterative, agile, and multimethod research approach including activities for exploration, cocreation, and testing and evaluation of innovations [5,6]. Cocreation activities are central to living lab research. Cocreation or co-design sessions are aimed at creating ideas and concepts together in an interactive manner by, for example, using prototypes, wireframes, or other verbal and visual generative tools [7]. Information from cocreation activities can be implemented to improve innovations or their implementation processes. A subsequent step in the iterative living lab approach can consist of lab-based or real-life testing of innovations with end users. This evaluation activity provides further information about the usability, acceptability, and impact of interventions.

Technological Innovations in a Rehabilitation Context

For this study (JRA), we will focus on the use of technology for rehabilitation in patients with stroke and people living with mild cognitive impairment (MCI; in accordance with the expertise and available infrastructures of the involved living

labs). Stroke is a leading health problem, and its burden is increasing worldwide [8]. Similarly, the prevalence of dementia, a neurodegenerative condition with a high physical, psychological, social, and economic impact for both patients and their carers, is rising sharply [9]. MCI is an intermediate state between normal cognition and dementia with preserved functional abilities [10]. Accumulating evidence supports the use of technology during the rehabilitation process both on a cognitive level and on a physical level. Concerning cognitive rehabilitation, García-Casal et al [11] have shown that computer-based cognitive interventions for people living with dementia lead to (moderate) beneficial effects on cognition, depression, and anxiety. For example, Gradior cognitive rehabilitation is a computer-based program for neuropsychological assessment and cognitive stimulation in healthy individuals as well as for neuropsychological rehabilitation in people with one or more cognitive disorders [12]. Concerning physical rehabilitation, a wide variety of technologies have been proposed, for example, virtual reality (VR), technologies for remote rehabilitation (telerehabilitation), technologies for augmenting gait training, and a short-arm human centrifuge (SAHC) for gravity therapy. A growing body of evidence supports the use of VR for rehabilitation of different pathologies. The benefits of using VR during the rehabilitation process of patients with stroke, patients with cerebral palsy, patients with spinal cord injuries, and other pathologies include improvements in balance and gait [13] and motor functions, greater community participation, and improved psychological and cognitive function [14]. With advances in information and communication technologies, telerehabilitation has gained increased research interest, as it allows rehabilitation services to be provided to patients remotely in their homes or elsewhere [15]. Research has shown that the effectiveness of telerehabilitation for patients with stroke is similar to that of traditional face-to-face rehabilitation but that different stakeholders have different opinions on the subject [16]. Patients are satisfied with telerehabilitation provided that it is appropriate and some social interaction occurs, whereas clinicians prefer face-to-face interactions and will only use telerehabilitation when face-to-face interactions are not feasible [16]. Another example of a promising technology for physical rehabilitation is the use of a split-belt treadmill for balance and gait training in patients with stroke [17,18]. Finally, gravity therapy by means of a SAHC, an integrated multisystem countermeasure to provide artificial gravity training for rehabilitation purposes, has been proposed to have beneficial effects for individuals with neuromuscular disorders, balance disorders, stroke, and sports injuries. By simulating natural gravity, it targets physiological deconditioning imposed by inactivity or a lack of gravitational force. It functions by exerting a centrifugal force on a body accelerated centripetally in a rotating device [15]. However, more research is needed to be able to provide a personalized therapy [19,20].

In addition, technology is used not only during interventions but also for diagnostics, evaluation of intervention effects, and remote monitoring. For example, in the context of the evaluation of physical activity (physical rehabilitation) and stress (mental rehabilitation), wearables are becoming increasingly common. Wearables are a specific type of mobile health application comprising sensors and devices intended to be worn on the body while collecting longitudinal and continuous data on cardiac cycles, electrodermal activity, skin temperature, acceleration, and so on, on a reliable and noninvasive manner outside of lab settings [21]. Also, a variety of wearable technologies for stroke rehabilitation have been studied to improve the diagnosis and treatment of upper-limb impairment (for a review, see the study by Maceira-Elvira et al [22]).

Despite proven effectiveness in research, however, the implementation of these technologies in practice is slow. For example, in the case of VR, barriers to successful adoption in rehabilitation practice involve three categories: technology development (eg, the degree of match or mismatch between the system and the client's goals or needs), competency development for end users (eg, perceived ease of use and utility), and facilitated clinical implementation (eg, access to technology and support for setup) [23]. Thus, there is a clear need not only to assess the effectiveness of these technology-based interventions but also to involve stakeholders from the design process to the evaluation of these innovations and to include assessments of feasibility, usability, user experience, and user acceptance. Therefore, living lab research methodologies are needed to implement these innovations and technologies in practice more successfully.

Study Aims and Objectives

VITALISE aims to harmonize living lab procedures and open living lab infrastructures as a means to facilitate and promote research activities in the health and well-being domain in Europe and beyond. With this JRA, we primarily aim to gain insight into each living lab's infrastructure and procedures to harmonize health and well-being living lab procedures and infrastructures, particularly in the context of rehabilitation. Secondarily, we aim to investigate the potential of innovative technologies for rehabilitation through living lab methodologies. To do so, multiple international living labs will organize cocreation sessions with patients and care professionals to capture stakeholder perspectives on rehabilitation technology (phase 1) and organize separate small-scale pilot studies to test 6 innovative intervention or data collection technologies in practice using preliminary harmonized procedures (phase 2). The living lab research activities have the following objectives: (1) to assess how different stakeholders across multiple countries view the role of technology in rehabilitation; (2) to collect suggestions regarding specific rehabilitation innovations (phase 1); (3) to perform small-scale pilot studies providing insights into end user experiences and usability of 6 selected cognitive and motor rehabilitation interventions; and (4) to explore the effects of these pilot interventions through self-report, physiological, and motor data of small samples (phase 2).

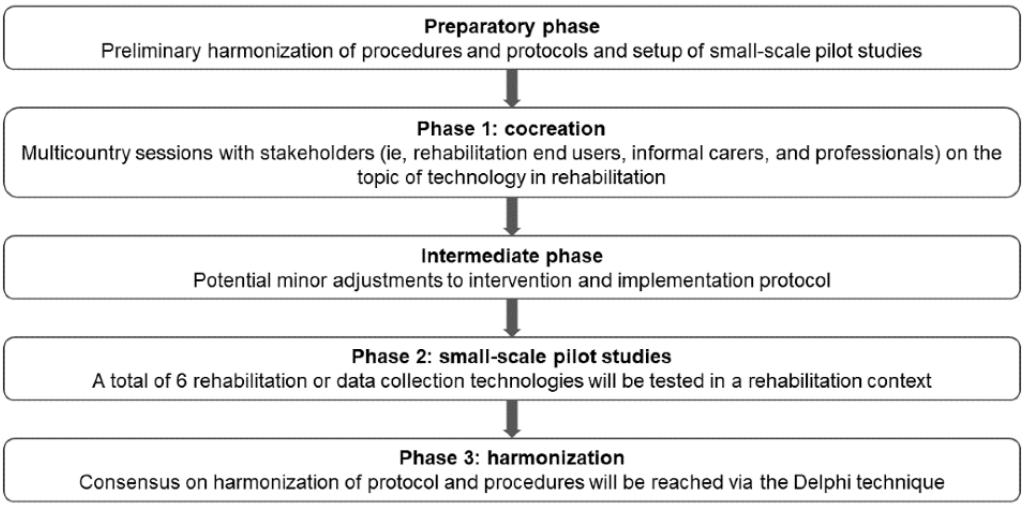
Methods

Design

This study has a mixed methods design comprising multiple phases (Figure 1) [24]. There are two main phases: phase 1 (cocreation) and phase 2 (small-scale pilot studies), which are preceded by a preliminary harmonization of procedures and protocols among the different international living labs. During

this preparatory phase, common interests concerning outcomes, outcome measures, and available infrastructures are discussed among the involved living labs to reach a preliminary harmonization and set up the small-scale pilot studies. In addition, ethical committee applications will be performed by the respective living labs. An intermediate phase (between phases 1 and 2) further allows the implementation of minor adjustments to the intervention or protocol depending on the input that was obtained in cocreation and discussions among living labs. This is in line with an iterative and agile design cycle during which innovations can have different maturity levels and undergo improvements during the research activities. However, adjustments should remain within the limits of the protocol approved by the ethical committee or be submitted as an amendment to this protocol. Note that the intermediate phase might differ among study partners depending on the results of the separate cocreation sessions. The preparatory phase is currently ongoing, and preliminary protocols to implement harmonized living lab procedures and infrastructures are being created.

Figure 1. Overview of the protocol design.



Site Selection, Sample, and Recruitment

Study phases 1 and 2 will be conducted by each living lab; therefore, recruitment will be carried out at each living lab's partner rehabilitation centers, rehabilitation departments of hospitals, or transitional living facilities in the context of rehabilitation across Europe and Canada. Participants will include patients with stroke or individuals with MCI (depending on the technology of interest and the planned small-scale pilot study).

In phase 1 (cocreation), approximately 4 to 12 participants will be included per session in accordance with guidelines for group sessions [25,26]. In phase 2 (small-scale pilot studies), each pilot aims to include approximately 15 participants, as previous research has suggested that this sample size is sufficient to detect most usability problems (up to 90%) [27,28]. Note, however, that the exact numbers can differ among the test sites depending on partner-specific objectives and methods.

Data Collection

Cocreation (Phase 1)

In cocreation, we will bring together rehabilitation end users, informal caregivers, and professionals in sessions to explore their views on technology in rehabilitation in general as well as on the specific technologies used in the small-scale pilot studies. End users are older adults with physical or (mild) cognitive disabilities. Informal caregivers are individuals who provide physical and emotional help to the end users, often close family members or friends, but who are not paid or trained by statutory bodies to do so. Professionals are individuals who work in a rehabilitation setting and can, for example, be clinicians (eg, physicians, physiotherapists, occupational therapists, psychologists, and speech therapists), directors, innovation managers, or other key staff members. Sessions will be organized at six test sites in four countries (Belgium, Canada, Greece, and Spain). Recruitment will be conducted by each regional living lab. The VITALISE project website [29] provides links to living lab hosting organizations home pages where more detailed information about each living lab and their infrastructure is available. Each session will include approximately 8 end users or professionals in accordance with the guidelines for group sessions [25,26]. The cocreation phase will consist of two parts: (1) a general part addressing views on technology in rehabilitation in general and (2) a specific part addressing the technology that will be studied in the small-scale pilot study. Both parts may be addressed in 1 to 4 sessions, depending on the availability (time) of the participants. The cocreation workshop scenario will be provided by LiCalab to direct the cocreation sessions across the different test sites, but the actual sessions will be supervised by each study site independently. This supports a harmonized study design with sufficient flexibility to take cross-cultural differences into account, which appears to be relevant for international living lab research [4]. Each session will adopt the following format: introduction, mapping challenges, demonstration, and reflection and conclusion. The main research topics for cocreation include acceptability, feasibility, usability, facilitators and barriers, advantages and limitations, and implementation, from the perspective of both end users (and informal caregivers) and

professionals. These topics will be studied via questions and exercises. The duration of a session is approximately 1.5 to 2 hours. Following cocreation, an intermediate phase (between phases 1 and 2) allows the implementation of minor adjustments to the intervention (eg, duration or frequency of the intervention and content of video vignettes) or protocol (eg, patient-reported outcome measures [PROMs] and patient-reported experience measures [PREMs]) depending on the input that was obtained from professionals, patients, or informal carers during cocreation and discussions among living labs. Note, however, that the intermediate phase might differ among study partners depending on the results of the separate cocreation sessions.

Small-Scale Pilot Studies (Phase 2)

General Overview

A mixed methods study design will be adopted in which 6 (technological) interventions will be implemented in rehabilitation practice across four countries (Greece, Spain, Belgium, and Canada) to assess user experience and explore the (preliminary) effects of the implemented intervention. Note, however, that not all interventions are technological in nature but that technology will also be used for data collection purposes. The research protocol comprises a variety of study designs, including pre-post designs, controlled studies, or a collaborative participatory research design. However, similar patient samples and data collection instruments and technologies will be adopted across small-scale pilot studies (Textbox 1). For example, general PROMs and PREMs concerning health and well-being include measures of quality of life (eg, 36-Item Short-Form Health Survey [30], World Health Organization Quality of Life Brief Version [31], Patient-Reported Outcomes Measurement Information System–Short Form [version 1.1]) and well-being (eg, 36-Item Short-Form Health Survey [30] and Patient-Reported Outcomes Measurement Information System–Short Form [version 1.1]). In addition, common living lab outcomes include measures of user experience (eg, Weiner scales [32] and System Usability Scale [33]). Note, however, that these PROMs and PREMs might undergo minor adjustments as a consequence of the cocreation process in phase 1 of the study. In addition to these PROMs, study-specific outcome measures are presented in Table 2. The pilot studies not only implement traditional living lab methodology, that is, experience-based (self-report) measures focusing on usability and user experience, but also demonstrate living lab infrastructure, such as wearables and other pieces of measurement equipment. This combination of different data collection methods provides a rich data set to evaluate rehabilitation interventions and share expertise regarding data collection and analysis among living labs.

Textbox 1. Overview of data collection methodology: general data collection.

Overview of data collection methodology

- Demographics (eg, gender, age, marital status, employment status, educational level, and stroke information)
- Anthropometry (eg, weight, height, body mass index, and neck and waist circumference)
- Sleep behavior (eg, sleep hours)
- Smoking behavior (eg, pack years)
- Quality of life (eg, World Health Organization Quality of Life–Age [34], 36-Item Short-Form Health Survey [30], Patient-Reported Outcomes Measurement Information System [35], and the 5-level classification system of the EQ-5D [36])
- Well-being (eg, 36-Item Short-Form Health Survey [30] or Warwick-Edinburgh Mental Well-being Scale [37])
- User experience of patients and clinicians (eg, System Usability Scale [33], User Experience Questionnaire [38], Weiner measures [32], or Usability Metric for User Experience (UMUX) [39])
- Interviews of patients and clinicians

Table 2. Overview of data collection methodology: additional pilot study–specific data collection.

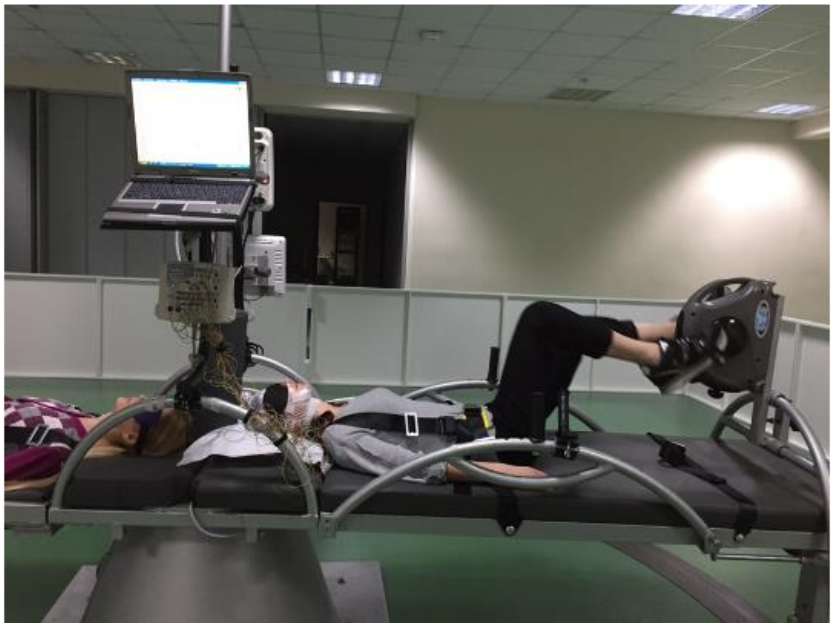
Living Lab and intervention and outcomes	Instruments
AUTH	
Short-arm human centrifuge	
Physical activity	Smartwatch (Polar Electro Oy, Kempele, Finland)
Cardiovascular activity	ClearSight (Edwards Lifesciences Corporation) noninvasive monitor, CNOGA (CNOGA Tensor tip MTX)
Muscle oxygen saturation	Moxy (Fortiori Design LLC)
Electroencephalographic data	Neurofax EEG-1200 32-channel device (Nihon Kohden)
Neurologic impairment	Expanded Disability Status Scale [40]
Mobility	6-minute walk test [41], timed up and go test [42], five times sit-to-stand test [43]
Balance	Berg Balance Scale [44], KFORCE (KINVENT), posturography, Dynamic Gait Index [45], backwards walking
Cognitive assessment	Symbol Digit Modalities Test [46]
Dual tasking	Walking while talking [47]
GAIA-Ocean Living Lab	
Art therapy	
Physical activity	International Physical Activity Questionnaire for elderly [48] or Global Physical Activity Questionnaire [49]
Balance	Berg Balance Scale [44], Activities-Specific Balance Confidence Scale [50], Smart Balance Board (Smartifier Oy)
Physiology	Smart devices (Xiaomi Mi Band 5, Samsung Galaxy Watch 3, and CAPTAIN eCoach [51])
Stress	Perceived Stress Scale [52]
INTRAS	
Gradior Cognitive	
Neuropsychological or cognitive assessment	Screening: Montreal Cognitive Assessment [53], Gradior Cognitive
Performance characteristics	Gradior Cognitive: % successes, graded cognitive performance scores, mistakes and successes by commission, mistakes by omission, number of performed sessions
LiCalab-Mobilab & Care	
Virtual reality mirror therapy	
Arm functionality	Fugl-Meyer assessment [54]
Pain	Visual analogue scale
Range of motion	Wearable wireless T-Sens Motion sensors of the Captiv L7000 (TEA Ergo)
UQAM-CRIR	
Teledance	
Participation and performance characteristics	Frequency and time of use and active practice time or motor engagement (via OpenTera platform or video recordings)
User experience	Short punctual or spot-check surveys (audio and visual or via OpenTera platform)
McGILL-UdeM-CRIR	
WALKAGAIN (augmented gait training)	
Walking in the community (eg, shopping and visiting a museum)	Qualitative (open-ended question)
Single and dual tasking	Assessment of walking (single), walking and reading, and walking in the presence of distractors
Movement characteristics (eg, tone, loss of sensation, sensorimotor deficits, and muscle strength)	Inertial measurement units or actigraph units

Living Lab and intervention and outcomes	Instruments
Mobility (eg, balance, gait impairments, and endurance)	Berg Balance Scale [44], 6-minute walk test [41], Mini-Balance Evaluation Systems Test [55], Activities-Specific Balance Confidence Scale [50]
Pain	Visual analogue scale
Participation	Community Healthy Activities Model Program for Seniors [56]

Short-Arm Human Centrifuge

This small-scale pilot study will be conducted by AUTH in Greece. In this pilot study, the SAHC is used for physical and cognitive rehabilitation to mitigate the detrimental effects of bed rest [20,57] (Figure 2). The objective is to assess the combined effects of artificial gravity and physical activity compared with standard of care in patients with stroke (acute or chronic) and healthy older adults over a period of 3 months (1 hour per session, 3 sessions per week). Participants will be assigned randomly to the SAHC training, standard of care training, or passive control. Data will be collected across the domains of body structure and function, activity, and participation as classified by the World Health Organization International Classification of Functioning, disability, and health, at six time points: at baseline, 4 weeks, 8 weeks, 3 months, and 6- and 12-month follow-ups.

Figure 2. Short-arm human centrifuge.



Participants will include 30 patients with acute stroke (15 SAHC and 15 passive control) and 30 patients with chronic stroke (15 SAHC and 15 passive control). They will be recruited from the Euromedica-Arogi Rehabilitation Center in Thessaloniki, Greece. The inclusion criteria include the following: individuals aged 18 to 70 years, without psychiatric disorder, without vertigo, without nausea or chronic pain, and without a history of syncope. The exclusion criteria include the following: individuals with a height >2 m, elite athletes, individuals with chronic use of substances (drugs or alcohol), individuals who have had a recent surgery, individuals with current arrhythmias, individuals with severe migraines, pregnant individuals, individuals with epilepsy, individuals with cholelithiasis or kidney stones, individuals with recent wounds from surgery, individuals with recent fractures, individuals with acute inflammation or pain, individuals with newly inserted metal pins or plates, and individuals with newly implanted stents.

The centrifugation on the SAHC will be combined with mild-intensity exercise based on the maximum heart rate. The SAHC intervention consists of 3 sessions per week, each with a duration of 1 hour, for 3 months. The participant is positioned in a supine and horizontal position on the rotation bed, with the head located toward the center. The beds with the patients turn around the axis of rotation with a force that is the product of body mass, distance from the axis of rotation, and angular velocity squared. Initially, there will be 1 session to familiarize participants with the SAHC group and to individually assess the optimal *g* load according to the participant's cardiovascular functioning with cardiac output, stroke volume, mean arterial pressure, diastolic blood pressure, systolic blood pressure, and heart rate. These criteria are monitored at each training session and are used to dynamically adapt the intervention intensity. More specifically, after 6 training sessions (2 weeks), the centrifugation load will be increased, and centrifugation will be combined with either aerobic exercise (through an ergometer) or resistance training through elastic training bands during centrifugation (depending on cardiovascular criteria). Further verification of the dynamic configuration of the intervention will be provided by the electroencephalographic assessment. Functional connectivity and cortical-network features derived from graph theory will be used by deep learning algorithms

(convolutional neural networks) to define the optimal centrifuge training.

Art Therapy

This small-scale pilot study will be conducted by GAIA-Ocean Living Lab in Spain. The objective of this small-scale pilot study is to assess the performing art methodologies (music and dance therapy) in different green and public spaces in the area of Gernika-Lumo (Spain) by assessing their feasibility and to explore their potential impact for individuals with stress-related conditions to support the therapists as well as formal and informal caregivers. Data will be collected at two time points: at baseline and after the end of the intervention.

A sample of 20 healthy men and women will be recruited from Gernika-Lumo (Biscay, Spain). Participants will be included if they are aged 18 to 85 years and will be allocated randomly to the testing or control group. Individuals with a medical history of severe cognitive impairment are not eligible to participate.

The performing arts intervention is a form of physical and mental rehabilitation offered to volunteers in public spaces. Each intervention will be conducted by experts on the methodologies for 2 months, twice a week. Participants will be instructed to follow the different activities, and their biometrics will be monitored during the sessions (smartwatch, smart band, and balance board) to assess their performance and compare their progress. At the end of the pilot period, participants in the control group will also have the opportunity to try out the intervention.

Gradior Cognitive

This small-scale pilot study will be conducted by INTRAS in Valladolid, Spain. The objective of this small-scale pilot study is to assess the most recent functionalities of Gradior Cognitive (Figure 3) by exploring the usability and preliminary effectiveness of the new features oriented to support the therapists and by assessing user experience and satisfaction in older adults with MCI. Data will be collected at two time points: at baseline and immediately after the end of the intervention.

Figure 3. Gradior Cognitive screenshots.

Outcome measures and general feedback will be collected from therapists ([clinical] psychologists, neuropsychologists, and gerontologists) using the new version of Gradior Cognitive (N=6 to 10). Furthermore, approximately 30 older adults with MCI will be recruited from the INTRAS Memory Clinic and Neuropsychological Rehabilitation Center in Valladolid, Spain. Potential older participants will be screened (cognitive abilities) before study participation. The inclusion criteria include the following: patients aged ≥55 years; patients with willingness and ability to collaborate in the study; and patients with either (1) a Mini-Mental State Examination [58] score of ≥21 and ≤27 indicating MCI according to the Petersen criteria for MCI-amnestic [59] and the international working group on MCI-amnestic or (2) initial dementia as categorized by the Diagnostic and Statistical Manual of Mental Disorders, fifth edition, and a Clinical Dementia Rating of ≥1 and ≤2 and a Geriatric Depression Scale score of <5 [60]. The exclusion criteria include the following: hearing or visual impairments (inability to use the devices); psychiatric, neurological, or nutritional condition preventing the individual from participating in the study; having a history of substance abuse (alcoholism or alcoholic-type dementia); and being on antipsychotic medication.

Gradior Cognitive is a neuropsychological evaluation and rehabilitation software for the implementation of higher cognitive function training programs for people with cognitive deficit or impairment. Gradior Cognitive allows working on attention, perception, orientation, memory, calculation, executive function, language, and reasoning in adults. The system consists of a website (intended for professionals) and an app for users to conduct intervention sessions. It analyzes and updates the results obtained reporting on users' performance in different cognitive areas, proposing changes in the session plan. This software supports therapists in the early detection and monitoring of cognitive impairment, also enabling a personalized cognitive rehabilitation program for improvement or maintenance of cognitive skills. Participant therapists included in the study will be using the new Gradior Cognitive features on a daily basis for at least two months. Individual interviews will be conducted to gather details on the therapist's experience and assess usability and perceived effectiveness. For

the older adult participants, the adopted intervention consists of 2-weekly sessions of 30 minutes using Gradior Cognitive for a period of 4 to 6 months. Sessions will be supervised by experienced Gradior professionals.

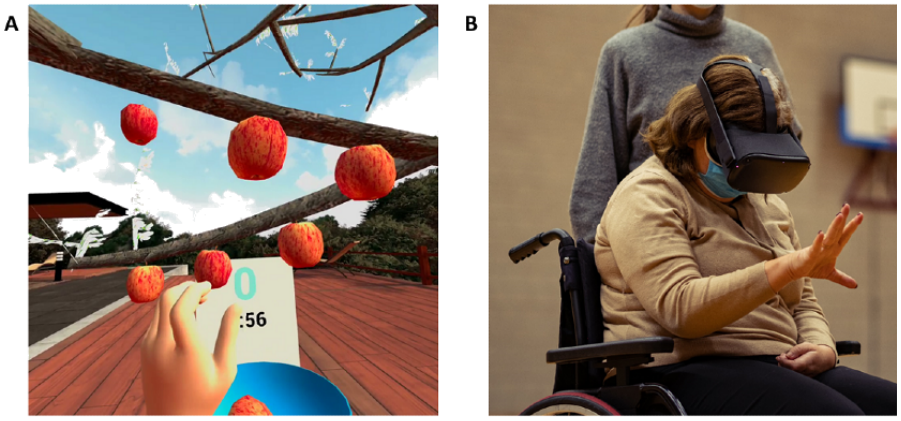
VR Mirror Therapy

This small-scale pilot study will be conducted by LiCalab-Mobilab & Care in Belgium. Mirror therapy is a rehabilitation method during which a mirror is placed between the arms or legs of a patient to create the illusion that a patient's affected limb is moving when the patient sees his or her unaffected limb moving in the mirror. The objective is to assess the noninferiority of VR mirror therapy as compared with standard mirror therapy in patients with subacute and chronic stroke, as well as to assess user experience of the VR mirror therapy from the perspective of both patients and clinicians. To do so, a two-arm, unblinded, randomized, controlled, parallel design is adopted to compare VR mirror therapy to regular mirror therapy (treatment as usual) in patients with stroke. A total of 20 patients with stroke will be randomly assigned to either the VR group or the control group receiving mirror therapy as usual. Data will be collected at four time points: at baseline, immediately after the first intervention session, before the last intervention session, and immediately after the last intervention session.

A total of 20 patients with stroke will be recruited by clinicians in partner rehabilitation centers or rehabilitation departments of general hospitals in Belgium and will be randomized to VR mirror therapy or treatment as usual. The clinicians themselves will also participate in the study by contributing data on usability and feasibility. The inclusion criteria for patients include the following: patients with stroke in the subacute or chronic phase, aged 18 to 85 years, with normally functioning upper limb on the nonaffected side and impaired functioning of the upper limb on the affected side, with the ability to sit independently on a chair or in a wheelchair in order to freely move the (unaffected) upper limb, and with the ability to follow verbal instructions. The exclusion criteria for patients include the following: patients with acute stroke, medically unstable patients, patients with visual deficits interfering with the execution of activities of daily living, patients with allergies for materials used on the VR headset (eg, silicone), patients with epilepsy, patients who have extreme sensitivity to motion sickness, and patients with facial wounds or lesions impairing the use of the VR headset.

Using an Oculus Quest VR headset (Meta) with built-in hand tracking, participants will be immersed in a VR kitchen and garden, where they will perform various actions in order to successfully complete a cooking program (Figure 4). These actions must be performed (in reality) by the side of the body that is unaffected by the stroke and include only upper-limb activities of daily living such as grasping ingredients or kitchen utensils, cutting vegetables, putting ingredients in a cooking pot, and stirring in the cooking pot. The movements of the unaffected side are projected to the affected side (mirrored), so that the patient appears to be performing the actions with the affected side. Participants in the control group will perform standard mirror therapy. Both groups will be administered the assigned therapy once per week over a period of 8 weeks.

Figure 4. Virtual reality (VR) mirror therapy: (A) screenshot of the apple-picking task from the perspective of a participant and (B) a patient performing the VR mirror therapy.



Teledance

This small-scale pilot study will be conducted by UQAM-CRIR in Canada. The objective is to test the teledance intervention, including a series of video vignettes adapted for stroke rehabilitation and follow-up modalities with users (asynchronous and synchronous). To do so, the intervention will be subject to a formative evaluation by the users to identify and resolve problems that may influence their experience [61] and to explore its potential impacts. By testing the intervention within an intensive functional rehabilitation setting, the objective is to identify the characteristics of users to whom such an intervention can be applied, evaluate its usability and safety, and explore its potential impacts. Data will be collected during each intervention session and at the end of the study.

Participants will include 9 clinicians (eg, physiotherapists, occupational therapists, and speech language therapists), 15 patients with stroke, and their relatives (N=15, comprising spouses, adult family members, or caregivers) recruited from 3 rehabilitation sites in Montreal, Canada. Included clinicians are required to have at least one year of work experience in a stroke rehabilitation program. The inclusion criteria for patients include the following: new admission to the facility's stroke program (≤14 days) and ability to participate safely and independently following an initial assisted teledance session with a clinician

(or with the sustained assistance of a relative). Patients who are aggressive or have known symptoms of dementia and Alzheimer disease before the stroke and who do not have sufficient cognitive abilities to follow instructions and give informed consent will be excluded. Relatives will include spouses, adult family members, or caregivers who are in close contact with the patient during rehabilitation. All participants must be able to understand and express themselves in French.

The teledance intervention will be offered to patients during their stay in the hospital and will consist of video vignettes delivered via the OpenTera software platform [62] on a digital tablet (eg, iPad), as well as follow-up modalities offered via OpenTera or other means. The intervention will be conducted by 2 dance educators for 5 to 6 weeks. Video vignettes will be +20 to −20 minutes long and offer thematic content, which will be performed in a sitting or standing position and either alone or assisted. Patients will be instructed to perform teledance at least 6 times per week for approximately 2 hours of dance per week. Follow-up sessions (patient-clinician, patient-dance educator, or clinician-dance educator) will also be included.

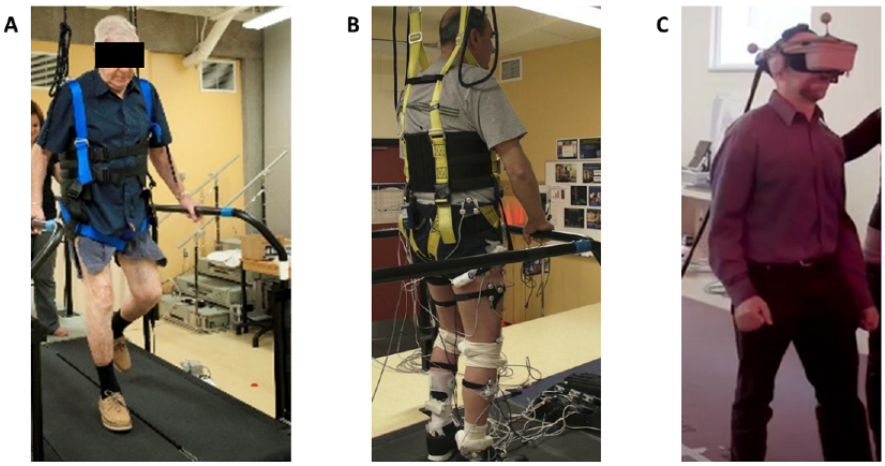
Augmented Gait Training

This small-scale pilot study (WALKAGAIN) will be conducted by McGILL-UdeM-CRIR in Canada. The main objective of this pilot study is to assess the effect of adding augmented gait training to usual rehabilitation and technology targeting the walking abilities of individuals admitted to rehabilitation after stroke in single-tasking, dual-tasking, and multitasking conditions. A secondary objective is to establish an international common data set for standing mobility assessment and conceptualize the common elements (dose, progression, etc) of the training protocols to be able to compare data among sites or analyze them together. To do so, a before-after pragmatic design with a blind outcome assessment will be set up to assess the performance of participants using augmented training protocols as compared with a control group or previously collected data (if available). The data from two Montreal Metropolitan (Canada) rehabilitation sites (Jewish Rehabilitation Hospital and Institut de réadaptation Gingras-Lindsay-de-Montréal) will be pooled together to assess the overall effect of the concepts of augmented gait training. Standardized outcomes will be collected in the week before and week after the locomotor training. We will also implement a follow-up assessment at 6 months (telephone call).

A total of 12 to 15 patients with stroke will be recruited from the stroke unit of the 2 sites for each locomotor training protocol. Participants will be selected for their potential to improve gait capacities and according to the requirements of their individual goals for improvement. The inclusion criteria will be the following: (1) having a first unilateral stroke and being in rehabilitation, (2) having initiated the gait training with or without assistive devices, (3) tolerating 1 hour of exercise with breaks, (4) in the presence of aphasia showing no more than mild or mild to moderate impairment, (5) having a reliable yes and no. The exclusion criteria will include cerebellar lesions, major pain, hemineglect or hemianopsia, signs of major depression quantified with a score of ≥10 out of 15 on the Geriatric Depression Scale [63], severe cognitive deficits defined with a score of <25 out of 30 on the Mini-Mental State Examination developed by Folstein et al [64], cardiorespiratory problems, and other medical and cognitive conditions that could affect the ability to understand instructions, as verified by medical records. We will also collect data on usability, feasibility, and opinions from clinicians about the training interventions through questionnaires or interviews.

Three major technologies (either individually or combined) will be used for locomotor training: biomechanical (eg, split belt training, walking with loaded segment, ascending slope), haptic (eg, vibrations), and VR (virtual environment), as shown in Figure 5. The precise locomotor training protocol will be decided by the clinicians in consultation with the researchers and will take into consideration the preferences of the participants (after phase 1). All locomotor training will include cognitive or language tasks as well attention-demanding tasks, implemented gradually from simple to more complex (single tasking, dual tasking, and multitasking) while the gait training will also be progressed. The training will consist of, for example, 30 minutes of walking or doing a number of steps, using the selected protocol with 5 minutes of warming-up and 5 minutes of recovery. Participants will be allowed to take breaks if their heart rate exceeds 60% of their maximal heart rate or they reported an exertion of 5 on the Borg rating scale, which represents severe exertion [65]. The training will be restarted when the participant's heart rate will lower back to its level at rest. The augmented gait training will be given twice a week for 10 sessions (5 weeks). Before each training session, the participant's heart rate (eg, Polar Electro Oy) and blood pressure will be taken at rest. The age-related maximum heart rate will be calculated using the Karvonen formula [66].

Figure 5. Examples of technologies for augmented gait training: (A) split-belt training (biomechanics), (B) vibrations (haptic), and (C) virtual reality.



Harmonization (Phase 3)

To define harmonized protocols and procedures (eg, for patient screening=outcomes and outcome measures) for living lab research in the context of technology in rehabilitation, the consensus process will involve a modified Delphi methodology (web-based, 3-round Delphi survey) followed by a consensus meeting of the living labs. The Delphi technique is a method for reaching consensus comprised of sequential questionnaires answered anonymously by a panel of experts with feedback provided after each questionnaire round [67,68]. A preliminary list of items for the Delphi survey will be based on literature research and findings of phases 1 and 2. For each living lab, 2 researchers participated (N=12).

Data Analysis

Cocreation (Phase 1)

Best practices for content analysis will be exchanged among partners to ensure a harmonized approach and the generalizability of the findings. Subsequently, for each test site, one or more researchers trained in qualitative methodology will provide a written report of each cocreation workshop and will use content analysis to analyze the data using main themes and subthemes.

Small-Scale Pilot Studies (Phase 2)

Best practices for analyses will be exchanged among partners to ensure a harmonized approach and the generalizability of the findings. However, for each small-scale pilot study, the statistical approach might vary according to the needs of the respective study. In general, descriptive statistics will be presented for demographic information. For quantitative data, normal distribution will be tested. The effects of an intervention on dependent variables will be assessed over time (eg, before-after, follow-up sessions) using repeated-measures parametric or nonparametric tests. The effect sizes of changes found among evaluations will be computed. The significance level is set at $P<.05$.

Harmonization (Phase 3)

Via the Delphi method, the living labs will reach consensus on which procedures (ie, outcomes, outcome measures, and devices) to use in a rehabilitation context. Each outcome, outcome measure, and device will be represented as an item in a list. In Delphi survey round 1, the international panel of living lab experts will be instructed to rate the importance of each item and asked to provide new items or suggest revisions of existing items. Consensus will be reached when at least 70% of the panel strongly agree or disagree on inclusion or exclusion of an item. Items that did not reach agreement and newly suggested items will be taken to round 2, together with feedback of round 1 responses. After the second round, items will be ranked by importance, and a list of items will be created. This list will be distributed among the panel members to discuss during a final consensus meeting.

Ethical and Legal Considerations

On a lower level, each partner institution involved in this study will submit an institutional review board application or ethical committee application according to the respective national regulations at the latest in February 2022. Informed consent will be obtained from all participants before data collection. For the pilot study involving individuals with (mild) cognitive disabilities, a team of trained psychologists will conduct recruitment of participants to ensure that potential participants' cognitive functions and comprehension abilities allow them to provide written informed consent. Only pseudonymized data will be shared among the living labs in the context of harmonization of procedures and protocols.

On a higher level, with the VITALISE project, we also aim to collect information on differences and similarities concerning legal frameworks, data management, and privacy. To date, the following project deliverables have addressed the issues of legal frameworks, privacy, and data management: Deliverable *D1.2 First Version of Ethics and Safety Manual* (submitted to European Commission on September 29, 2021) reports on the VITALISE project's international, European, and national ethical regulations, device standards, and certifications as well as accepted data management procedures; presents a first plan for the VITALISE scheduled data collection to be compliant with the reported regulations and guidelines; and evaluates potential concerns. In addition, project deliverable *D1.4 Data Management Plan (first version)* (submitted to European Commission on September 29, 2021) provides provisory information regarding the data to be collected during the project,

addresses the issue of data reuse for further exploitation by expanding on the manner in which the data will be made accessible, and gives information on the data's curation and preservation. Deliverables will be available to the general public via the VITALISE project website [1] after completion of the European Commission review process (approximately October 2022). Concerning informed consent procedures, identification and recruitment of participants, security measures, and anonymization or pseudonymization procedures of personal data and data transfer to non-EU countries, two project deliverables *D14.1 H—Requirement No. 1* and *14.2 Protection of Personal Data—Requirement No. 3* have been submitted to the European Commission on November 29, 2021. These deliverables will, however, only be available to the consortium members and Commission services.

Results

Research activities (eg, participant enrollment, data collection, and data analysis) for this project will start in March 2022 (phase 1) and June 2022 (phase 2) and end in March 2024. Phase 1 (cocreation) will provide insights into different stakeholders' views (ie, end users, clinicians, and informal caregivers) on the role of technology in rehabilitation on a European level (and beyond), as well as specific suggestions regarding rehabilitation innovations. Phase 2 (small-scale pilot studies) will provide insights into end user experiences, and usability and preliminary effectiveness of 6 selected neuropsychological or physical rehabilitation interventions for patients with stroke and older individuals. Phase 3 (harmonization via the Delphi methodology) will provide a consensus on harmonized screening, outcomes, and outcome measures for living lab research in the context of technology in rehabilitation. Together, these research activities will provide an opportunity for all involved living labs (and beyond) to gain knowledge on specific technologies intended for rehabilitation and to discuss and adopt each other's technologies or procedures, to harmonize living lab procedures and infrastructures concerning the use of innovative technology in rehabilitation on a European level.

Discussion

Study Significance

This JRA protocol needs to be considered in light of the European Commission's Horizon 2020 program called *Integrating Activities for Starting Communities* (INFRAIA-02-2020). This JRA is part of three JRAs included in the VITALISE project that will investigate three domains of health and well-being research, namely, rehabilitation, transitional care, and everyday living environments. With these JRAs, the VITALISE project aims to improve, in quality or quantity, the integrated services provided at a European level by serving as test beds for harmonizing procedures and services. This JRA, in particular, focuses on the use of technology in rehabilitation.

The primary outcome of this study (JRA) allows the involved living labs to exploit similarities in existing services and infrastructures, such as devices, hardware, and software being used for either intervention or data collection purposes, and harmonize (and potentially eliminate) identified differences. The secondary outcome of this JRA is an increased knowledge and improvement of technological innovations for rehabilitation of patients with stroke and individuals living with MCI or dementia, on an international level.

The study design involves a collaboration of multiple international living labs (1) to gain insight into stakeholders' views on the use of innovative technologies for rehabilitation on an international level, (2) to explore the (preliminary) effects of the pilot intervention, and (3) to explore and learn from each living lab by using each other's infrastructures, to harmonize procedures and infrastructures across living labs.

Conclusions

The JRA described in this research protocol demonstrates cocreation activities and implements a variety of study designs that illustrate living lab practices and infrastructure. The knowledge gained by the execution of this research will lead to harmonized procedures and protocols in a rehabilitation context for living labs in Europe and beyond. In addition, we will also be able to provide new insights for improving the implementation of innovative technologies in rehabilitation.

Acknowledgments

The Virtual Health and Wellbeing Living Lab Infrastructure is funded by the Horizon 2020 Framework Programme of the European Union for Research Innovation (grant agreement 101007990).

Conflicts of Interest

None declared.

References

1. VITALISE. 2021. URL: https://vitalise-project.eu/ [accessed 2022-02-09]
2. ENOLL-Frequently asked questions. 2021. URL: https://enoll.org/about-us/ [accessed 2022-02-09]
3. Konstantinidis EI, Petsani D, Bamidis PD. Teaching university students co-creation and living lab methodologies through experiential learning activities and preparing them for RRI. Health Informatics J 2021;27(1):1460458221991204 [FREE Full text] [doi: 10.1177/1460458221991204] [Medline: 33535859]

4. De Witte NA, Adriaensen I, Broeckx L, Van Der Auwera V, Van Daele T. Cross-cultural differences in user-centred research: an international living lab survey. Health Informatics J 2021;27(3):14604582211038268 [FREE Full text] [doi: 10.1177/14604582211038268] [Medline: 34424056]
5. Ballon P, Van Hoed M, Schuurman D. The effectiveness of involving users in digital innovation: measuring the impact of living labs. Telemat Informat 2018 Aug;35(5):1201-1214. [doi: 10.1016/j.tele.2018.02.003]
6. van den Kieboom RC, Bongers IM, Mark RE, Snaphaan LJ. User-driven living lab for assistive technology to support people with dementia living at home: protocol for developing co-creation-based innovations. JMIR Res Protoc 2019 Jan 28;8(1):e10952 [FREE Full text] [doi: 10.2196/10952] [Medline: 30688653]
7. Mukhtar M, Ismail MN, Yahya Y. A hierarchical classification of co-creation models and techniques to aid in product or service design. Comput Indust 2012 May;63(4):289-297. [doi: 10.1016/j.compind.2012.02.012]
8. Feigin V, Krishnamurthi RV, Parmar P, Norrving BO, Mensah GA, Bennett DA, GBD 2013 Writing Group, GBD 2013 Stroke Panel Experts Group. Update on the global burden of ischemic and hemorrhagic stroke in 1990-2013: the GBD 2013 study. Neuroepidemiology 2015;45(3):161-176 [FREE Full text] [doi: 10.1159/000441085] [Medline: 26505981]
9. Prince M, Wimo A, Guerchet M, Ali G, Wu Y, Prina M. World Alzheimer Report 2015 - The global impact of dementia: an analysis of prevalence, incidence, cost and trends. Alzheimer's Disease International, London. 2015. URL: http://www.alz.co.uk/research/world-report-2015 [accessed 2022-02-09]
10. Hugo J, Ganguli M. Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. Clin Geriatr Med 2014 Aug;30(3):421-442 [FREE Full text] [doi: 10.1016/j.cger.2014.04.001] [Medline: 25037289]
11. García-Casal JA, Loizeau A, Csipke E, Franco-Martín M, Perea-Bartolomé MV, Orrell M. Computer-based cognitive interventions for people living with dementia: a systematic literature review and meta-analysis. Aging Ment Health 2017 Dec;21(5):454-467. [doi: 10.1080/13607863.2015.1132677] [Medline: 26806365]
12. Martín F, Villameriel O, Aguado B, Díez C, Jimenez A. Programa Gradior. Rehabilitacion cognitiva por ordenador. Issuu. 2013. URL: https://issuu.com/fundacion_intras/docs/gradior-iberdiscap-2000/1 [accessed 2022-02-10]
13. Porras DC, Siemonsma P, Inzelberg R, Zeilig G, Plotnik M. Advantages of virtual reality in the rehabilitation of balance and gait: systematic review. Neurology 2018 May 29;90(22):1017-1025. [doi: 10.1212/WNL.0000000000005603] [Medline: 29720544]
14. Massetti T, da Silva TD, Crocetta T, Guarnieri R, de Freitas BL, Lopes P, et al. The clinical utility of virtual reality in neurorehabilitation: a systematic review. J Cent Nerv Syst Dis 2018;10:1179573518813541 [FREE Full text] [doi: 10.1177/1179573518813541] [Medline: 30515028]
15. Brennan D, Mawson S, Brownsell S. Telerehabilitation: enabling the remote delivery of healthcare, rehabilitation, and self management. In: Studies in Health Technology and Informatics. Amsterdam, Netherlands: IOS Press; 2009:231-248.
16. Caughlin S, Mehta S, Corriveau H, Eng J, Eskes G, Kairy D, et al. Implementing telerehabilitation after stroke: lessons learned from Canadian trials. Telemed J E Health 2020 Jun;26(6):710-719 [FREE Full text] [doi: 10.1089/tmj.2019.0097] [Medline: 31633454]
17. Esmaeili V, Juneau A, Dyer J, Lamontagne A, Kairy D, Bouyer L, et al. Intense and unpredictable perturbations during gait training improve dynamic balance abilities in chronic hemiparetic individuals: a randomized controlled pilot trial. J Neuroeng Rehabil 2020 Jun 17;17(1):79 [FREE Full text] [doi: 10.1186/s12984-020-00707-0] [Medline: 32552850]
18. Betschart M, McFayden B, Nadeau S. Lower limb joint moments on the fast belt contribute to a reduction of step length asymmetry over ground after split-belt treadmill training in stroke: a pilot study. Physiother Theory Pract 2020 Sep;36(9):989-999 [FREE Full text] [doi: 10.1080/09593985.2018.1530708] [Medline: 30358510]
19. Yang C, Zhang S, Zhang Y, Wang B, Yao Y, Wang Y, et al. Combined short-arm centrifuge and aerobic exercise training improves cardiovascular function and physical working capacity in humans. Med Sci Monit 2010 Dec;16(12):575-583. [Medline: 21119574]
20. Kourtidou-Papadeli C, Frantzidis C, Gilou S, Plomariti C, Nday C, Karnaras D, et al. Gravity threshold and dose response relationships: health benefits using a short arm human centrifuge. Front Physiol 2021;12:644661 [FREE Full text] [doi: 10.3389/fphys.2021.644661] [Medline: 34045973]
21. Debard G, De Witte N, Sels R, Mertens M, Van Daele T, Bonroy B. Making wearable technology available for mental healthcare through an online platform with stress detection algorithms: the Carewear project. J Sensors 2020 Nov 25;2020:1-15. [doi: 10.1155/2020/8846077]
22. Maceira-Elvira P, Popa T, Schmid A, Hummel F. Wearable technology in stroke rehabilitation: towards improved diagnosis and treatment of upper-limb motor impairment. J Neuroeng Rehabil 2019 Nov 19;16(1):142 [FREE Full text] [doi: 10.1186/s12984-019-0612-y] [Medline: 31744553]
23. Glegg S, Levac D. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237-1251 [FREE Full text] [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]
24. O'cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2017 Sep 21;13(2):92-98 [FREE Full text] [doi: 10.1258/JHSRP.2007.007074]
25. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. Field Methods 2016 Jul 24;29(1):3-22. [doi: 10.1177/1525822x16639015]

26. Carlsen B, Glenton C. What about N? A methodological study of sample-size reporting in focus group studies. BMC Med Res Methodol 2011 Mar 11;11(1):26 [FREE Full text] [doi: 10.1186/1471-2288-11-26] [Medline: 21396104]
27. Cazañas A, de San Miguel A, Parra E. Estimating sample size for usability testing (Estimación del tamaño de la muestra para pruebas de usabilidad). Enfoque UTE. 2017. URL: https://ingenieria.ute.edu.ec/enfoqueute/index.php/revista/article/view/126 [accessed 2022-02-09]
28. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
29. VITALISE Consortium. VITALISE. URL: https://vitalise-project.eu/consortium/ [accessed 2022-02-22]
30. Ware JE, Sherbourne CD. The MOS 36-ltem Short-Form Health Survey (SF-36). Med Care 1992;30(6):473-483. [doi: 10.1097/00005650-199206000-00002]
31. The Whoqol Group. The World Health Organization quality of life assessment (WHOQOL): development and general psychometric properties. Soc Sci Med 1998 Jun;46(12):1569-1585. [doi: 10.1016/s0277-9536(98)00009-4]
32. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
33. Brooke J. SUS: A 'Quick and Dirty' usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. Usability Evaluation In Industry. London: Taylor and Francis; 1996.
34. Caballero F, Miret M, Power M, Chatterji S, Tobiasz-Adamczyk B, Koskinen S, et al. Validation of an instrument to evaluate quality of life in the aging population: WHOQOL-AGE. Health Qual Life Outcomes 2013 Oct 23;11:177 [FREE Full text] [doi: 10.1186/1477-7525-11-177] [Medline: 24152691]
35. Katzan I, Lapin B. PROMIS GH (Patient-Reported Outcomes Measurement Information System Global Health) scale in stroke. Stroke 2018 Jan;49(1):147-154. [doi: 10.1161/strokeaha.117.018766]
36. Janssen M, Pickard A, 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]
37. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes 2007 Nov 27;5:63 [FREE Full text] [doi: 10.1186/1477-7525-5-63] [Medline: 18042300]
38. Schrepp M, Hinderks A, Thomaschewski J. Construction of a Benchmark for the User Experience Questionnaire (UEQ). IJIMAI 2017;4(4):40. [doi: 10.9781/ijimai.2017.445]
39. Lewis J, Utesch B, Maher D. UMUX-LITE: when there's no time for the SUS. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; April 27 - May 2, 2013; Paris France p. 2099-2102. [doi: 10.1145/2470654.2481287]
40. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983 Nov 01;33(11):1444-1452. [doi: 10.1212/wnl.33.11.1444] [Medline: 6685237]
41. Butland RJ, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, six-, and 12-minute walking tests in respiratory disease. Br Med J (Clin Res Ed) 1982 May 29;284(6329):1607-1608 [FREE Full text] [doi: 10.1136/bmj.284.6329.1607] [Medline: 6805625]
42. Podsiadlo D, Richardson S. The timed "Up and Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
43. Csuka M, McCarty DJ. Simple method for measurement of lower extremity muscle strength. Am J Med 1985 Jan;78(1):77-81. [doi: 10.1016/0002-9343(85)90465-6]
44. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83 Suppl 2:S7-11. [Medline: 1468055]
45. Shumway-Cook A, Woollacott M. Motor Control : Theory and Practical Applications. Baltimore, MD: Williams and Wilkins; 1995.
46. Smith A. Symbol Digit Modalities Test (SDMT). Los Angeles: Western Psychological Services; 1982.
47. Verghese J, Buschke H, Viola L, Katz M, Hall C, Kuslansky G, et al. Validity of divided attention tasks in predicting falls in older individuals: a preliminary study. J Am Geriatr Soc 2002 Sep;50(9):1572-1576. [doi: 10.1046/j.1532-5415.2002.50415.x] [Medline: 12383157]
48. Hurtig-Wennlöf A, Hagströmer M, Olsson LA. The International Physical Activity Questionnaire modified for the elderly: aspects of validity and feasibility. Public Health Nutr 2010 Mar 03;13(11):1847-1854. [doi: 10.1017/s1368980010000157]
49. Cleland CL, Hunter RF, Kee F, Cupples ME, Sallis JF, Tully MA. Validity of the global physical activity questionnaire (GPAQ) in assessing levels and change in moderate-vigorous physical activity and sedentary behaviour. BMC Public Health 2014 Dec 10;14(1):1255 [FREE Full text] [doi: 10.1186/1471-2458-14-1255] [Medline: 25492375]
50. Powell LE, Myers AM. The Activities-specific Balance Confidence (ABC) scale. J Gerontol A Biol Sci Med Sci 1995 Jan 01;50A(1):28-34. [doi: 10.1093/gerona/50a.1.m28] [Medline: 7814786]
51. CAPTAIN. URL: https://www.captain-eu.org/ [accessed 2022-02-22]

52. Chan SF, La GA. Perceived Stress Scale (PSS). In: Gellman MD, Turner JR, editors. Encyclopedia of Behavioral Medicine. New York: Springer; 2013.
53. Nasreddine Z, Phillips N, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
54. Gladstone D, Danells C, Black S. The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair 2002 Sep;16(3):232-240. [doi: 10.1177/154596802401105171] [Medline: 12234086]
55. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BESTest. J Rehabil Med 2010 Apr;42(4):323-331 [FREE Full text] [doi: 10.2340/16501977-0537] [Medline: 20461334]
56. Stewart AL, Mills KM, King AC, Haskell WL, Gillis D, Ritter PL. CHAMPS physical activity questionnaire for older adults: outcomes for interventions. Med Sci Sports Exerc 2001 Jul;33(7):1126-1141. [doi: 10.1097/00005768-200107000-00010] [Medline: 11445760]
57. Hötting K, Röder B. Beneficial effects of physical exercise on neuroplasticity and cognition. Neurosci Biobehav Rev 2013 Nov;37(9 Pt B):2243-2257. [doi: 10.1016/j.neubiorev.2013.04.005] [Medline: 23623982]
58. Arevalo-Rodriguez I, Smailagic N, Figuls MR, Ciapponi A, Sanchez-Perez E, Giannakou A, et al. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev 2015 Mar 05(3):CD010783 [FREE Full text] [doi: 10.1002/14651858.CD010783.pub2] [Medline: 25740785]
59. Petersen R, Smith G, Waring S, Ivnik R, Tangalos E, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999 Mar;56(3):303-308. [doi: 10.1001/archneur.56.3.303] [Medline: 10190820]
60. Morris J. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993 Nov;43(11):2412-2414. [doi: 10.1212/wnl.43.11.2412-a] [Medline: 8232972]
61. Campeau-Vallerand C, Michaud F, Routhier F, Archambault P, Létourneau D, Gélinas-Bronsard D, et al. Development of a web-based monitoring system for power tilt-in-space wheelchairs: formative evaluation. JMIR Rehabil Assist Technol 2019 Oct 26;6(2):e13560 [FREE Full text] [doi: 10.2196/13560] [Medline: 31674918]
62. Panchea AM, Brière S, Hamel M, Maheux M, Godin C, Tousignant M, et al. OpenTera: a microservice architecture solution for rapid prototyping of robotic solutions to COVID-19 challenges in care facilities Internet. arXiv 2021 [FREE Full text]
63. Yesavage JA, Brink T, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiat Res 1982 Jan;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4]
64. Burns A, Brayne C, Folstein M. Key Papers in Geriatric Psychiatry: mini-mental state: a practical method for grading the cognitive state of patients for the clinician. M. Folstein, S. Folstein and P. McHugh,Journal of Psychiatric Research, 1975,12, 189-198. Int J Geriat Psychiatry 1998 May;13(5):285-294. [doi: 10.1002/(sici)1099-1166(199805)13:5<285::aid-gps753>3.0.co;2-v]
65. Borg G. Psychophysical scaling with applications in physical work and the perception of exertion. Scand J Work Environ Health 1990;16 Suppl 1:55-58 [FREE Full text] [doi: 10.5271/sjweh.1815] [Medline: 2345867]
66. Pescatello L, Arena R, Riebe D, Thompson P. ACSM's Guidelines for Exercise Testing and Prescription. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins; 2014.
67. Dalkey NC. The Delphi Method : an experimental study of group opinion. RAND Corporation. 1969. URL: https://www.rand.org/pubs/research_memoranda/RM5888.html [accessed 2022-02-09]
68. Sinha I, Smyth R, Williamson P. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. PLoS Med 2011 Jan 25;8(1):e1000393 [FREE Full text] [doi: 10.1371/journal.pmed.1000393] [Medline: 21283604]

Abbreviations

JRA: joint research activity
MCI: mild cognitive impairment
PREM: patient-reported experience measure
PROM: patient-reported outcome measure
SAHC: short-arm human centrifuge
VITALISE: Virtual Health and Wellbeing Living Lab Infrastructure
VR: virtual reality

Edited by T Derrick; submitted 28.10.21; this is a non–peer-reviewed article;accepted 22.12.21; published 10.03.22.

Please cite as:
Bernaerts S, De Witte NAJ, Van der Auwera V, Bonroy B, Muraru L, Bamidis P, Frantzidis C, Kourtidou-Papadeli C, Azevedo N, Garatea J, Muñoz I, Almeida R, Losada R, Fung J, Kehayia E, Lamontagne A, de Guise E, Duclos C, Higgins J, Nadeau S, Beaudry L, Konstantinidis E
Rehabilitation Supported by Technology: Protocol for an International Cocreation and User Experience Study
JMIR Res Protoc 2022;11(3):e34537
URL: https://www.researchprotocols.org/2022/3/e34537
doi:10.2196/34537
PMID:35266874

Protocol

Island Study Linking Aging and Neurodegenerative Disease (ISLAND) Targeting Dementia Risk Reduction: Protocol for a Prospective Web-Based Cohort Study

Larissa Bartlett[1], PhD; Kathleen Doherty[1], PhD; Maree Farrow[1], PhD; Sarang Kim[1], PhD; Edward Hill[1], PhD; Anna King[1], PhD; Jane Alty[1], MD; Claire Eccleston[1], PhD; Alex Kitsos[1], MSc; Aidan Bindoff[1], BSc (Hons); James C Vickers[1], PhD

Wicking Dementia Research and Education Centre, University of Tasmania, Hobart, Australia

Corresponding Author:
Larissa Bartlett, PhD
Wicking Dementia Research and Education Centre
University of Tasmania
Medical Sciences Precinct
17 Liverpool Street
Hobart, 7000
Australia
Phone: 1800 982 600
Email: larissa.bartlett@utas.edu.au

Abstract

Background: Up to 40% of incident dementia is considered attributable to behavioral and lifestyle factors. Given the current lack of medical treatments and the projected increase in dementia prevalence, a focus on prevention through risk reduction is needed.

Objective: We aim to increase dementia risk knowledge and promote changes in dementia risk behaviors at individual and population levels.

Methods: The Island Study Linking Aging and Neurodegenerative Disease (ISLAND) is a long-term prospective, web-based cohort study with nested interventions that will be conducted over a 10-year period. Target participants (n=10,000) reside in Tasmania and are aged 50 years or over. Survey data on knowledge, attitudes, and behaviors related to modifiable dementia risk factors will be collected annually. After each survey wave, participants will be provided with a personalized dementia risk profile containing guidelines for reducing risk across 9 behavioral and lifestyle domains and with opportunities to engage in educational and behavioral interventions targeting risk reduction. Survey data will be modeled longitudinally with intervention engagement indices, cognitive function indices, and blood-based biomarkers, to measure change in risk over time.

Results: In the initial 12 months (October 2019 to October 2020), 6410 participants have provided baseline data. The study is ongoing.

Conclusions: Recruitment targets are feasible and efforts are ongoing to achieve a representative sample. Findings will inform future public health dementia risk reduction initiatives by showing whether, when, and how dementia risk can be lowered through educational and behavioral interventions, delivered in an uncontrolled real-world context.

International Registered Report Identifier (IRRID): DERR1-10.2196/34688

(JMIR Res Protoc 2022;11(3):e34688) doi:10.2196/34688

KEYWORDS

public health; online; prospective research cohort; dementia; aging; older adult; neurodegenerative; modifiable risk factors; risk reduction; prevention; lifestyle and behaviors; lifestyle; behavior change; intervention; risk; cognition; blood-based dementia biomarkers; research translation

Introduction

Dementia is one of the most prevalent and burdensome conditions that impact the community, and projections suggest that 152 million people will be affected worldwide by 2050 [1]. In the absence of medical treatments that halt or slow disease progression, dementia risk reduction has been identified as a priority for public health research and policy development [2-6].

Recent estimates suggest that between 40% and 48% of dementia could be prevented or delayed through dementia risk reduction interventions [3,7]. Modifiable dementia risk factors include diabetes, hypertension, high cholesterol, obesity, hearing loss, physical inactivity, smoking, excessive alcohol consumption, low educational attainment, lack of cognitive stimulation, social isolation, brain injury, pollution exposure, stress, and depression [7,8]. The population attributable fraction of risk (which combines the relative risk with prevalence) for independent risk factors ranges from 1% (obesity, diabetes) to 8% (hearing loss) [7]. Behavioral and lifestyle factors can co-occur and the interplay between multiple risk factors is not yet well understood [9,10]. Models with Australian data that account for nonindependence suggest that reducing the rate of each risk factor by 20% could reduce the prevalence of dementia by up to 30% by 2050 [3].

The effects of behavioral interventions for reducing dementia risk—a group-based intervention targeting diet and exercise, with cognitive training and vascular risk factor management [11]; dietary advice [12,13]; changes in physical activity [14,15]; cognitive training [16,17]; undertaking formal education in later life [18]; and a multicomponent program targeting cognitive activity, exercise, diet, sleep, and depression management [19]—are actively being investigated. These trials [11-19] have a necessary focus on internal validity and are subject to time constraints; therefore, they largely target at-risk older individuals to maximize the chance of observing change in dementia-related outcomes within the study period. Although trial results can provide evidence of cause and effect, external validity (including the extent to which findings can be replicated in real-world settings and generalized to public health or long-term benefit) needs to be verified [11].

There has also been a concerted effort in dementia research focused on understanding the biological and functional changes that precede clinical symptoms of neurodegeneration. Promising new molecular markers that support the potential for individual dementia risk profiling have been identified [9,20]. For example, apolipoprotein epsilon-E and polygenic risk can now be identified within families, and blood-based biomarkers (eg, beta-amyloid, phosphorylated tau, and neurofilament light) can indicate neuropathology and neurodegeneration up to 20 years before functional symptoms occur [21,22]. Similarly, new web-based tools are emerging that use computer-based audiovisual techniques and artificial intelligence to remotely measure cognitive, visual, speech, and motor functions [23]. Subtle changes in thinking and memory and in lexicosyntactic and motor domains are understood to be associated with early-stage dementia [24-26]. Using these emerging techniques, it is possible to detect and track preclinical signals of neuropathology.

Life-course modeling suggests that the future incidence of dementia could be reduced by modifying dementia risk-related behaviors in midlife [7,27], and it is now important to test these models in real-world contexts. Long-term large-sample prospective cohort studies present a viable approach for tracking changes in dementia risk factors over time at the individual and population level [2,5,28]. Linking biological and functional markers of dementia pathology with information about dementia risk behaviors and demographic data provides the opportunity for detailed cohort characterization and enables precision in selecting and targeting risk reduction interventions [9]. It is then feasible to investigate patterns of change in relation to intervention engagement for symptomatic, high-risk, and asymptomatic participants. A large prospective research cohort can thus provide a framework for investigating the modification potential of individual and population-level dementia risk profiles in relation to intervention exposures and can help determine the conditions in which public health dementia risk reduction interventions could be most beneficial [29,30].

Tasmania has a population of 0.5 million that has an aging profile and higher incidences of many dementia risk factors than the rest of Australia. Furthermore, while there is a public appetite for dementia prevention, only 30% of Australians reportedly recognize that dementia risk is modifiable [31,32]. Emerging evidence indicates help-seeking older adults find that knowing their own risk profile information improves their understanding of how they can reduce their risk of developing dementia [33]. This is encouraging and offers support for providing personalized feedback to help increase the personal resonance of risk reduction messages and to stimulate health behavior change [34,35]. This knowledge, understanding that it is possible to change one's health status, having access to the opportunity and support required to take action, and being motivated to do so are all considered to be important contributing factors for achieving effective health-related behavior change [36,37].

The primary hypothesis for this study is that increasing knowledge and motivations related to modifiable dementia risk, and engaging in behavioral and lifestyle-related risk reduction activities, will be protective over time against declining cognition and neurobiological markers of dementia, at the individual level and the population level. There are 3 key aims: (1) characterize dementia risk knowledge and behaviors, current and historical health status, cognition, and dementia-related biomarkers; (2) deliver health-related community-based and educational dementia risk reduction interventions; and (3) determine the real-world conditions in which dementia risk reduction interventions yield the greatest benefit to inform future dementia prevention research, policy, and practice.

Methods

Overview

The Island Study Linking Aging and Neurodegenerative Disease (ISLAND) is a large-sample prospective public health cohort

study with nested interventions. Targeting dementia risk-related behaviors and modifiable dementia risk factors, ISLAND is being conducted in Tasmania, Australia.

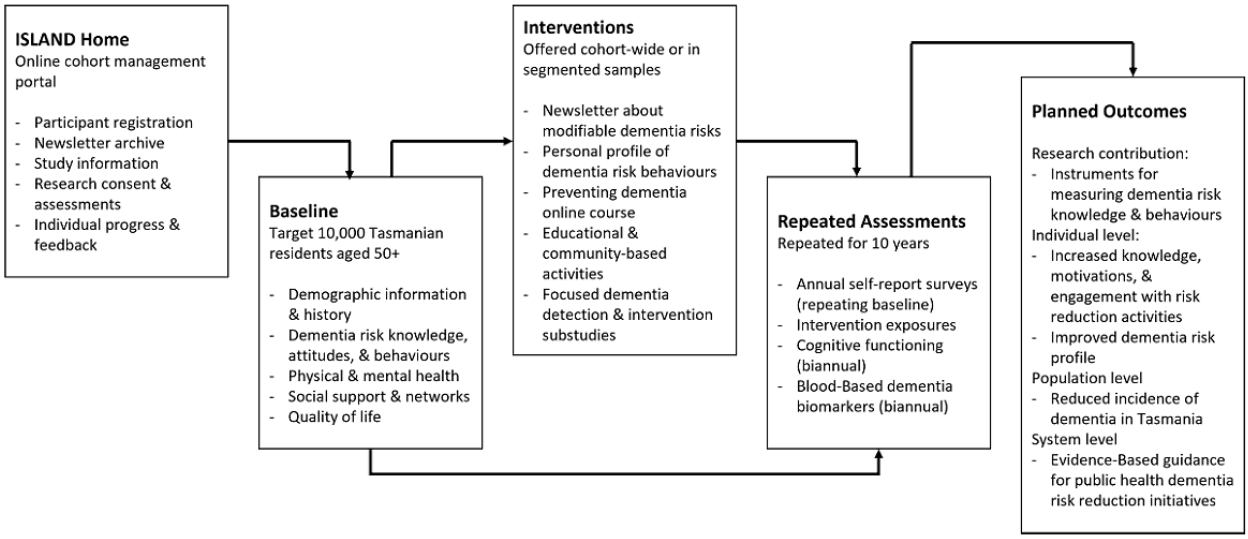
Ethics

The study has been approved by University of Tasmania's Health and Medical Human Research Ethics Committee (HREC H001864) and will be conducted in accordance with the National Health and Medical Research Council's National Statement on Ethical Conduct in Human Research [38]. Potential participants will be given up-to-date information and asked to provide consent to the research conditions at baseline; enrolled participants will be provided updated study materials and asked to reconsent at all subsequent assessment waves.

Study Design

This study will be conducted over a period of 10 years (2019 to 2029), with repeated measurements and a range of risk reduction interventions (Figure 1). We anticipate that this timeframe will be sufficient for changes in cognitive and biological markers to occur. Participants are considered active agents who are provided the opportunity to engage in managing their own dementia risk based on their knowledge, personal motivations, and circumstances, and to encourage dementia risk reduction in their communities. Longitudinal patterns of knowledge about modifiable dementia risks, motivations to engage with risk-reduction interventions, and changes in dementia risk behaviors, cognitive function, and blood-based biomarkers will be investigated. Data will be modeled to test whether, when, and how dementia risk reduction interventions are likely to reduce future dementia incidence in the long term.

Figure 1. Study design and planned outcomes. ISLAND: Island Study Linking Aging and Neurodegenerative Disease.



The study will primarily be web-based to maximize access for Tasmanian residents regardless of their location. A dedicated, secure web-based environment has been designed to manage participant engagement and to remotely administer questionnaires and cognitive assessments. In-person attendance is only required for blood sample collection, which will take place in community-based clinics.

Measures

Participants' knowledge of and motivations to change behaviors related to dementia risk will be collected annually with surveys, in conjunction with current behaviors and engagement indices (Table 1). The Checklist for Reporting the Results of Internet e-Surveys [39] is available in Multimedia Appendix 1. Data from the Knowledge of Dementia Risk Reduction instrument, the Motivations to Change Lifestyle and Health Behaviours for Dementia Risk Reduction scale [40], and dementia risk profile will be used to support behavior change modeling [41]. A portion of the cohort will also complete the New General Self-Efficacy scale [42], the All Aspects of Health Literacy Scale [43], and the Perceived Stress Scale [44].

The Cambridge Neuropsychological Test Automated Battery [45,46] will be used to measure changes in thinking and memory, and Talk2Me [23], an automated linguistic data collection instrument, will be used to detect lexicosyntactic indicators of cognitive function. Motor–cognition will be measured using a screening instrument for preclinical Alzheimer disease (TAS Test) [47] that applies computer vision to capture the speed and rhythm of hand movements, which are understood to decline with the onset of dementia pathology [26]. The Cambridge Neuropsychological Test Automated Battery, Talk2Me, and TAS Test instruments are designed for web-based administration, allowing a wider reach than that which could be achieved with in-person assessment. Multisource cognitive data provide the opportunity to understand performance changes in multiple domains and can be used to triangulate, and cross-validate changes observed in other data. Data on nonmodifiable risk factors, such as age, sex, and genetics (from blood samples), and procedural factors on intervention implementation, engagement, and outcomes will also collected [29,30].

Table 1. Measurement instruments.

Instrument or assessment	Outcome
Background and Health Survey	Detailed demographic, health, and lifestyle characteristics
Intervention engagement	System reports of course enrolment and progression and newsletter engagement; self-report intervention and community activity engagement data
Knowledge of Dementia Risk Reduction Survey	Knowledge of dementia risk; recall and recognition of modifiable and nonmodifiable dementia risk factors
Motivation to Change Lifestyle and Health Behaviours for Dementia Risk Reduction Scale [40]	Beliefs and attitudes toward lifestyle and behavioral changes for dementia risk reduction
All Aspects of Health Literacy Scale[a] [43]	Functional, communicative, and critical health literacy
New General Self-Efficacy Scale[a] [42]	Perceived ability to achieve a range of different types of task.
Dementia Risk Profile	Behaviors affecting dementia risk factors: diagnosis, regular checks and management of cardiometabolic health; BMI; physical and cognitive activity; diet, alcohol consumption, and smoking
Hospital Anxiety and Depression Scale [48]	Symptom severity for 2 dimensions (anxiety and depression); pooled score indicates psychological distress
Perceived Stress Scale[a] [44]	Extent to which daily life is perceived as stressful
Lubben Social Network Scale [49]	Extent of social networks; 3 dimensions (family, neighbors and friends)
Assessment of Quality of Life [50]	Multiattribute utility instrument that generates psychometric and Quality of Life Years scores based on 8 dimensions of health-related quality of life: physical health (independent living, pain, and senses) and psychosocial health (mental health, happiness, coping, relationships, and self-worth)
Written reflection task [51]	Cognitive performance: idea density, grammar, and sentence construction
Talk2Me Online [23]	Cognitive performance: image naming, picture description, and audio files providing approximately 2000 lexicosyntactic, acoustic, and semantic features for analysis
Cambridge Neuropsychological Test Automated Battery Online [45]	Cognitive performance: Paired Associates Learning captures learning and recall of visual information over successive trials and is sensitive to cognitive decline in early Alzheimer disease and mild cognitive impairment [52] Spatial Working Memory assesses executive function, retention, and manipulation of visuospatial information
TAS Test [47]	Motor and cognitive performance using keyboard tapping, visuomotor reaction tests, and visuospatial working memory tests providing approximately 1000 motor-cognitive features for analysis
Blood samples	Biomarker levels indicative of dementia pathology (eg, beta-amyloid, phosphorylated tau, and neurofilament light) measured in plasma or serum using enzyme-linked immunoassay, SIMOA,[b] and mass spectrometry
Genetics	Candidate gene markers related to Alzheimer disease such as apolipoprotein epsilon-E and the brain-derived neurotrophic factor polymorphism measured via blood samples

[a]Instruments only administered to some participants.

[b]SIMOA: Single Molecule Array.

The dementia risk profile questionnaire and scoring are based on risk criteria outlined by the World Health Organization [6] and the Australian National University's Alzheimer Risk Index [53]. For cognitive activity, the frequencies of 11 different cognitive and cultural activities are recorded (low risk: 33 or higher; high risk: less than 33). For physical activity, minutes per week of light, moderate, and vigorous activity are recorded, with each minute assigned 3.3 MET, 4 MET, and 8 MET and summed for the total score (low risk: 600 or more; high risk: less than 600). For alcohol consumption, the number of standard drinks per week is recorded (low risk: less than 14 drinks per week, and no more than 2 standard drinks per day; medium risk: 14 or less standard drinks per week and either more than 2 standard drinks per occasion or binge drinking; high risk: more than 14 standard drinks per week). Risk levels associated with blood pressure management are based on whether hypertension has been diagnosed, check-up frequency, and treatment (low risk: no diagnosis and regular check-ups, or diagnosis under medical management; medium risk: regular check-ups but with either unsure diagnosis or with hypertension and being treated; high risk: hypertension diagnosis and no regular check-ups or medical treatment, or no diagnosis and no regular check-ups). Risk levels for cholesterol management and diabetes management are scored in the same way as hypertension. BMI is calculated as weight divided by height squared (low risk: 18 to 24.9 kg/m^2; medium risk: 25 to 29.9 kg/m^2 or less than 18 kg/m^2; high risk: 30 kg/m^2 or over). Adherence to the Mediterranean-DASH Intervention for Neurodegenerative Delay

diet is calculated using the total score, which ranges from 0 to 14 (low risk: 12 or higher; medium risk: 7.5 to 11.9; high risk: less than 7.5). For smoking, low risk is assigned when no smoking is reported, medium risk is assigned when occasional smoking is reported, and high risk is assigned when participants smoke daily, almost daily, or weekly. In addition to dementia risk profile scores, risk levels for depression are computed using data from the Hospital Anxiety and Depression Scale [48]. Normative cut points are applied to assign risk levels: low (normal, 0 to 7), medium (borderline, 8 to 10) and high risk (11 to 21).

Setting and Participants

Tasmania's population is distributed across urban, rural, and remote settings. A larger percentage of residents (19.4%) are over the age of 65 years (compared to 14.8% nationally), and Tasmania has the highest median age of all Australian states (Tasmania: 42 years; Australia: 37 years) [54]. Relative to the Australian population, Tasmania has high rates of smoking (Tasmania: 16.4%; Australia: 13.8%); obesity (Tasmania: 70.9%; Australia: 67.0%); heart, stroke, or vascular disease (Tasmania: 6.0%, Australia: 4.8%); diabetes (Tasmania: 5.5%, Australia: 4.9%); and poor mental health (Tasmania: 21.7%, Australia: 20.1%). Educational attainment is also low in Tasmania, with 27.5% of the population aged 15 to 75 years having no educational qualification above year 10 (or 16 years of age) compared with the national average of 19.3% [55]. These characteristics mean the Tasmanian population is at higher risk of dementia than the broader Australian population and is ideal for a public health research initiative implementing strategies to achieve dementia risk reduction.

Recruitment and Retention

We aim to recruit 10,000 participants, which is 5% of the Tasmanian population aged 50 or over. This sample offers a meaningful slice of the population for the project's public health objectives. Eligibility criteria are age 50 years or older (based on the likelihood of higher personal health motivations and to provide maximum opportunity for observing longitudinal changes from midlife in behavioral, cognitive, and biological markers [35]), resident of Tasmania, having an email address, and having access to the internet. Invitations to join the study are regularly and widely distributed using web-based, print, and broadcast media; community talks; information booths; social media; flyers; and posters. Potential participants register interest via ISLAND Home, a secure web-based cohort management portal [56]. On their personal ISLAND Home page, participants can monitor their study engagement history and progress, update their contact information, access newsletter archives, and respond to research and intervention invitations. ISLAND project staff liaise with participants by email, through social media, and in person at the community-level to support recruitment, engagement, and retention. Participants who self-identify as having deteriorating cognitive functioning in their annual survey responses will be advised to seek medical referral to the affiliated ISLAND Clinic for clinical assessment, diagnosis, and treatment.

Interventions

ISLAND Newsletter

Upon registration, participants are sent an email newsletter every 3 to 4 weeks. The newsletters provide a digest of evidence-based information, project findings, information about modifiable dementia risks, and invitations to join research and community activities.

Dementia Risk Profile

Participants complete a dementia risk profile questionnaire at each timepoint. Upon completion, they are provided with a personalized report of their individual risk level (low, medium or high) for 9 behavioral and lifestyle domains, and includes advice, based on World Health Organization guidelines [6], for shifting to a lower risk level. The dementia risk profile is intended to be used by participants to monitor changes in their own dementia risk-related behaviors over time, as a source of motivation for addressing risk, or to share with their primary health provider for guidance in risk factor management.

Preventing Dementia Massive Open Online Course

The Preventing Dementia Massive Open Online Course presents contemporary evidence-based information on dementia risk and guidance on risk reduction behaviors [57]. The free 4-week course is offered twice each year. Quizzes provide participant feedback, and social engagement is supported by monitored discussion forums. In 2019, the Preventing Dementia Massive Open Online Course was ranked by Class Central as the all-time third highest among health-related Massive Open Online Courses [58]. Since its launch in 2016, the course has attracted 154,000 enrolments from 175 countries, of whom 43% have completed the course. Based on pilot evidence [59], exposure to the course is expected to increase knowledge and motivations related to modifiable dementia risks, and thus support changes in dementia risk behaviors.

ISLAND Campus

Evidence from the Tasmanian Healthy Brain Project [60,61] showed later life university education may protect against cognitive decline. Between June and September 2020, full fee-waiver scholarships were made available by the University of Tasmania to study participants to undertake a full university course of their choosing. To be eligible, participants needed to provide additional survey data (Perceived Stress Scale, New General Self-Efficacy Scale, All Aspects of Health Literacy Scale). These data will be used to measure the influence of perceived stress, self-efficacy, and health literacy on student engagement, and the combined long-term effect of university study in later life on dementia risk knowledge, attitudes, behaviors, and cognitive functioning will be examined.

We have also put in place partnerships with established government and nongovernment health promotion organizations (eg, Tasmanian Health Service, Libraries Tasmania, municipal offices within the Local Government Association of Tasmania, Quit Tasmania, The Heart Foundation, Diabetes Tasmania). These partnerships will be leveraged to provide study participants opportunities to connect with local

community-based public health activities that target dementia risk factors.

Future Interventions

This study is a valuable opportunity to conduct substudies focusing on specific research questions. For example, cohort segments can be targeted with particular interventions, or recruited into studies of factors that influence or detect change in dementia trajectories. Participant invitations to join future studies, health promotion programs, and community-initiated events (walking groups, information sessions, social groups) will be promulgated via notices on ISLAND Home, in newsletters, and on social media. Biological samples collected from participants, such as blood, will support further investigation of links between proteins, genetic markers, neurodegenerative disease, and brain function.

Data Management and Storage

Survey, cognition, and biomarker data will be deidentified and stored on secure databases in accordance with institutional ethics and privacy policies. Blood samples will be stored indefinitely in secure freezers (at –80 °C) for future analyses.

Data is collated using the secure web-based portal. Survey, cognitive assessment, biomarker, and intervention data are linked by a unique identifier, forming a comprehensive database of participant demographics, risks, exposures, functioning, neuropathology, and potential confounders at participant level.

Planned Analysis

Relevant literature will be monitored to inform covariate selection and exclusion criteria. Missing data mechanisms will be assessed and accounted for using standard techniques.

Multilevel regression models will be used to estimate changes in dementia risk knowledge, attitudes and behaviors over time with respect to exposures. These models allow control of type 1 errors where data do not meet assumptions of independence. Multinomial state-space models will be used to model changes in risk behaviors over time by estimating the expected probability of transition from state *i* at time *t*–1 to state *j* at time *t*, where states *i* and *j* represent discrete risk (eg, high risk and low risk). Longitudinal profiles of cognitive assessment and biomarker trajectories will be modeled using multilevel regression models, adjusting for covariates including age, sex, years of education, and intervention exposure. Multivariate relationships between risk knowledge, attitudes, and behavior states with dementia risk reduction, cognitive assessments, biomarkers, measured exposures, other demographic data, and survey data will be explored using clustering or ordination techniques such as *t*-distributed stochastic neighbor embedding.

Results

Recruitment in the first 12 months (October 2019-October 2020) indicates the target baseline sample is feasible. As of October 1, 2020, a total 12,706 people had registered interest in the study, of whom 6410 provided complete data in the baseline demographic and dementia risk profile surveys. The cohort is distributed across the state's populated regions, with greater density in the capital city (south) and regional urban centers (north and northwest) (Figure 2). The characteristics of the ISLAND cohort compared with the Tasmanian population are presented in Table 2. Study participants' median age is 63 years, and there is a nearly even split between retired (3002/6410, 46.8%) and employed (2924/6410, 45.6%) participants. Nearly three-quarters (4630/6410, 72.2%) are female, and a higher proportion have university degrees (3107/6410, 50.0%) than population norms (36,337/206,421, 17.5%). There is also a higher proportion of people living in areas of socioeconomic advantage (857/6410, 13.4%) than in the general population (9495/206,421, 4.6%). Of 6410 participants, 1731 (27.0%) have completed the Preventing Dementia Massive Open Online Course. Recruitment is complete for ISLAND Campus; 1322 participants are enrolled, of whom 784 (59.3%) have commenced a new university course.

Figure 2. Geographical distribution of the Tasmanian population (N=542,000) and study participants (per 100, n=6410). The map was traced from Google Maps [62], and population [63] and participant data were overlaid using Google's JavaScript and Geocoding APIs.

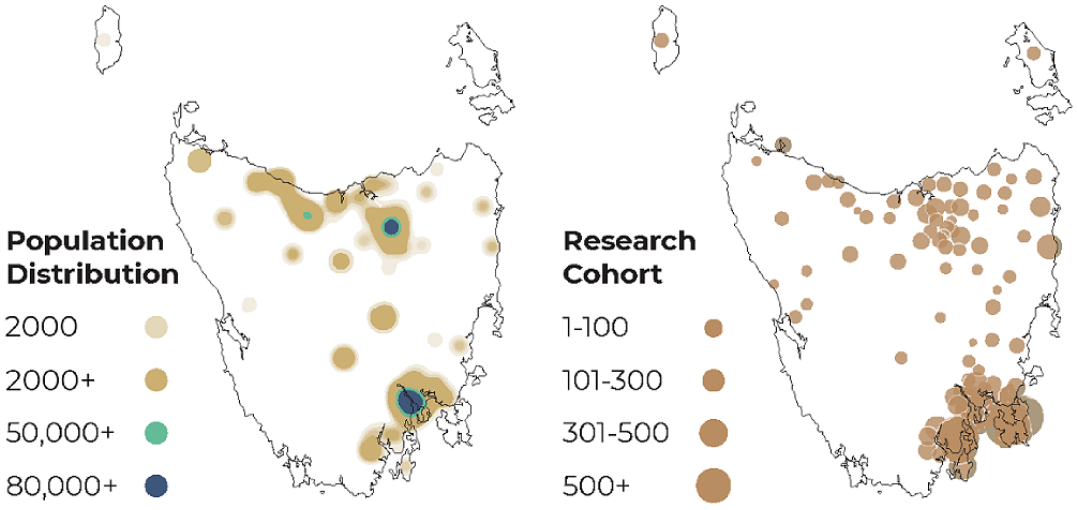


Table 2. Study participants (October 2019-October 2020) compared with Tasmanian residents 50 years and older. Tasmanian population data drawn from the Australian Bureau of Statistics [63].

Demographic variables	Participants (n=6410), n (%)	Tasmanian population >50 years of age (n=206,421), n (%)
Age (years)		
Mean (SD)	63.1 (7.5)	65.3 (10.6)
Median (range)	63 (50, 94)	63 (50, 105)
Age categories (%)		
50-59 years	2306 (36.0)	72,912 (35.3)
60-69 years	2724 (42.5)	67,723 (32.8)
70-79 years	1237 (19.3)	42,044 (20.4)
80+ years	143 (2.2)	23,742 (11.5)
Gender (%)		
Female	4630 (72.2)	108,014 (52.3)
Male	1771 (27.6)	98,418 (47.7)
Other	9 (0.01)	N/A[a]
Work Status (%)		
Retired	3002 (46.8)	118,221 (49.1)
Employed/Work-ready	2924 (45.6)	105,931 (41.9)
Missing	78 (0.01)	16,816 (0.07)
Education level (%)		
Postgraduate degree	1832 (28.6)	12,344 (5.9)
Bachelor's degree	1275 (21.5)	23,993 (11.6)
Diploma/trade	1929 (30.1)	62,095 (34.4)
High school	1018 (15.9)	52,192 (28.9)
Primary school	4 (0.01)	29,756 (16.5)
Missing	252 (3.9)	N/A
Regional distribution (%)		
North and northeast	1452 (22.7)	58,094 (28.2)
West and northwest	919 (14.3)	46,597 (22.6)
South and southeast	4006 (62.5)	101,456 (49.2)
Missing	27 (0.4)	0 (0)
Socioeconomic status[b] (%)		
Quintile 1	1822 (28.4)	76,788 (37.2)
Quintile 2	1380 (21.5)	53,876 (26.1)
Quintile 3	914 (14.3)	37,775 (18.3)
Quintile 4	1404 (21.9)	28,486 (13.8)
Quintile 5	857 (13.4)	9495 (4.6)
Missing	33 (0.5)	0 (0)

[a]N/A: not available.

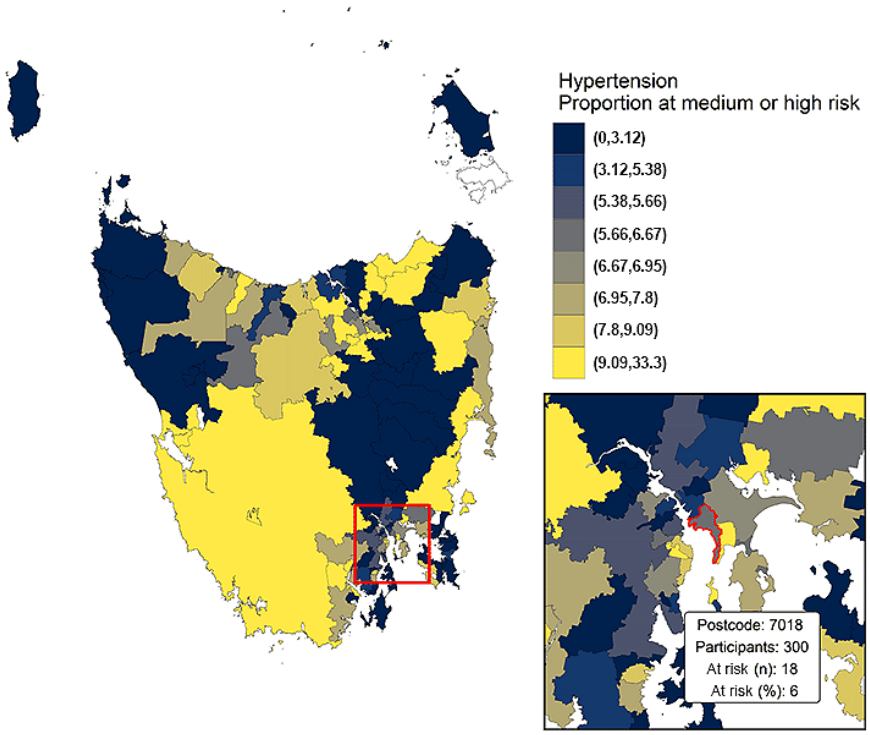
[b]Quintile 1 is the most relatively disadvantaged areas, and Quintile 5 is the most relatively advantaged areas.

Risk status for 10 of the established modifiable dementia risk factors [7,8] (Table 3) and the geographical distribution of participants scoring high risk for unmanaged hypertension (Figure 3) are shown. Figure 3 is presented to show how the combination of dementia risk profile data and geographical location enable the identification of areas where community-based risk reduction interventions might be most warranted.

Table 3. Risk status of participants for 10 modifiable dementia risk factors.

Risk factor and level	Participants (n=6410), n (%)
Cognitive activity	
Low	3519 (54.9)
High	2800 (43.7)
Physical activity	
Low	5429 (84.7)
High	741 (11.6)
Alcohol consumption	
Low	3207 (50.0)
Medium	1934 (30.2)
High	779 (12.1)
Blood pressure	
Low	5925 (92.4)
Medium	175 (2.7)
High	258 (4.0)
BMI	
Low	2293 (35.8)
Medium	2197 (34.3)
High	1616 (25.2)
Cholesterol level	
Low	5209 (81.3)
Medium	230 (3.6)
High	922 (14.4)
Diabetes	
Low	5270 (82.2)
Medium	113 (1.8)
High	963 (15.0)
Mediterranean diet adherence	
Low	1284 (20.0)
Medium	4646 (72.5)
High risk	435 (6.8)
Smoking	
Low	6109 (95.3)
Medium	80 (1.2)
High	205 (3.2)
Depression	
Low	5784 (91.1)
Medium	407 (6.4)
High	155 (2.4)

Figure 3. Participants at high or medium risk due to unmanaged hypertension, by postcode. The map was created using ggplot2 in R [64] overlaid with 2016 postcode boundaries from the Australian Bureau of Statistics [63].



Discussion

This project presents a unique opportunity to observe, at individual and population levels, the long-term natural life-course trajectories related to dementia indicators and risk behaviors and responds to calls for public health action on dementia prevention [2]. The overall objective of the project is to extend and translate epidemiological and experimental intervention evidence regarding modifiable dementia risk factors, integrate life-course perspectives and inform policy development for public health dementia risk reduction initiatives [9,10]. By including 10,000 participants (or 5% of the target population aged 50 years or over), characterizing dementia risk at individual and aggregate levels, and tracking changes in relation to engagement with risk reduction interventions, this study will contribute valuable knowledge about the long-term effects of adjusting health behaviors in middle and later life on dementia risk [6,28]. This is a valuable opportunity to find out whether, when, and how dementia risk can be lowered through risk reduction interventions delivered in an uncontrolled population-based public health context.

We will apply and develop innovative web-based techniques for remotely assessing cognitive function and at scale. These functional data will be combined with detailed demographic data and data on established and emerging biomarker and genetic indicators, to characterize dementia risk in this population cohort. The research sample can thus be segmented for interventions and research questions (eg, ISLAND Campus) as well as broad-scale offerings (eg, Preventing Dementia Massive Open Online Course). Even with anticipated attrition that is common in longitudinal cohorts [65], the large sample size provides the opportunity to identify patterns of change in dementia risk over time, accommodating a wide range of behavioral and health-related antecedents, procedural factors, and demographic covariates.

While our target population sample is an ambitious goal, our preliminary results indicate that it is achievable. The early responder sample has a high proportion of women and people with university qualifications, and the proportions at medium or high risk of dementia due to obesity, smoking, and depression are lower than those in the Tasmanian population [54]. This means our sample is not yet reflective of Tasmanian norms. Because educational attainment is an established determinant of socioeconomic inequalities in both physical health and cognitive functioning [66], we expect people with lower levels of education and living in areas of relative socioeconomic disadvantage may have poorer health profiles and stand to benefit more from engaging with a dementia risk reduction public health program. Ideally, we aim to achieve a representative research sample, to maximize the reach of risk reduction information and opportunities and to support the future generalization of findings. The demographic profile of the current sample will guide targeted recruitment calls to increase the number of men and people with low formal education. Ongoing efforts are focused on recruitment through community-based networks in lower socioeconomic status areas, facilitated by partner organizations such as information technology access centers, neighborhood houses, municipal offices, and technical or trade-dominant workplaces. We anticipate that this targeted and supported recruitment strategy will increase the representation of people who may not otherwise join a web-based research program and whose data will

contribute to more population-equivalent distributions across the risk domains (eg, smoking, BMI, diet). A related study [67] is examining the drivers of reach and resonance of risk reduction messages at community and population level and will inform the ongoing development and dissemination of actionable information about dementia risk reduction. Cultural and linguistic diversity was not included in the initial baseline data but is increasingly recognized as an important determinant of health and health behaviors. These data are being collected from the sample in follow-up surveys.

Because we use a participant-centered design that is informed by behavior change theory and public health research methods, participants are not required to join predetermined research interventions, they are invited to choose from a range of opportunities. The choices that participants make are based on their own knowledge, beliefs, attitudes, and circumstances. In the absence of existing measures that can be used to model the role of knowledge in behavior changes, data from this project will support the validation of new research instruments that assess dementia risk reduction knowledge (Knowledge of Dementia Risk Reduction instrument) and dementia risk behaviors. The constructs measured by these instruments, along with dementia risk reduction motivations and attitudes [40], health literacy [43], and self-efficacy [42], are central to understanding behavior change in relation to dementia risk. Initial development of these instruments was conducted using data samples separate from ISLAND data. The TAS Test motor–cognitive screening test is currently being validated with drawn from within the ISLAND sample. It is our hope that these 3 new tools will strengthen research into dementia risk reduction within and beyond the ISLAND Project. Using multisource data from the Cambridge Neuropsychological Test Automated Battery, Talk2Me, and TAS Test cognitive assessments in conjunction with blood-based biomarkers and self-report surveys will enable triangulation of objective and self-report data. This will add weight to the validity of ISLAND findings, and providing valuable insights into the potential for altering pathological dementia trajectories through behavioral intervention.

The web-based portal is intended to support the project's reach across Tasmania's distributed population. The framework is scalable and can be replicated for deployment in other populations, and thus offers multiple opportunities for research collaborations in the field of dementia risk reduction. Independent of their contribution to dementia, many of the modifiable dementia risk factors—particularly obesity, smoking, hypertension, depression, isolation, and diabetes—often co-occur and are directly associated with poor quality of life and mortality [68]. Collaboration at the community level with established health promotion organizations provides a sustainable and cost-effective way for participants to be alerted to risk reduction activities occurring in their area. Linking participants' geographical location with risk profile information (Figure 3) provides the opportunity to identify areas with a high proportion of risk behaviors. This enables a data-driven rationale for partner organizations (eg, Diabetes Tasmania, Quit Tasmania) to direct community-based interventions into these areas. Thus, our approach has the potential to incidentally address other chronic health conditions and could feasibly lead to considerable public health benefits. Indeed, during the initial COVID-19 lockdown period in Tasmania, in contrast to global public narratives, ISLAND participants did not report detrimental changes in their mental or physical health [69]. The design of this study may prove to be a model that is particularly valuable for low- to middle-income countries where both dementia risk and multimorbidity are high. Ultimately, it is our intention that this project will advance what is known about behavioral and lifestyle dementia risk reduction and contribute to global public health research and practice focused on preventing dementia.

Acknowledgments

This work was supported by the Australian Government's Medical Research Future Fund, the National Health and Medical Research Council (2004051), the J.O. and J.R. Wicking Trust (Equity Trustees), the University of Tasmania, St. Luke's Health, and the Masonic Centenary Medical Research Foundation. The funding bodies have no direct role in the study design, data collection, analysis, and interpretation or manuscript preparation. The authors are grateful for contributions made by Helen Douglas, Dr Adam Kane, and Justine Keay (the Island Study Linking Aging and Neurodegenerative Disease [ISLAND] Project team); Timothy Saunder, Alex Kitsos, Chris Parker, and Joshua Eastgate (technical and data management); and Oliver Freeman and Jay Borchard (graphics). We are indebted to all the participants of the ISLAND Project, who make this work possible.

Authors' Contributions

KD, MF, SK, CE, JA, AK, AB, and JCV conceived and designed the study and secured the initial ethics approval. LB directed the design of the Island Study Linking Aging and Neurodegenerative Disease Home portal, oversaw data collection, prepared the initial manuscript, and coordinated coauthor contributions. AB and LB conducted analyses and directed the development of figures. All authors contributed to the development and refinement of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Checklist for Reporting Results of Internet E-Surveys.

[PDF File (Adobe PDF File), 162 KB - resprot_v11i2e34688_app1.pdf]

References

1. 2021 Dementia statistics. Alzheimer's Disease International. URL: https://www.alzint.org/about/dementia-facts-figures/dementia-statistics/ [accessed 2022-03-08]
2. Chong TW, Macpherson H, Schaumberg MA, Brown BM, Naismith SL, Steiner GZ. Dementia prevention: the time to act is now. Med J Aust 2021 Mar 14;214(7):302. [doi: 10.5694/mja2.50972] [Medline: 33715184]
3. Ashby-Mitchell K, Burns R, Shaw J, Anstey KJ. Proportion of dementia in Australia explained by common modifiable risk factors. Alzheimers Res Ther 2017 Feb 17;9(1):11 [FREE Full text] [doi: 10.1186/s13195-017-0238-x] [Medline: 28212674]
4. Frankish H, Horton R. Prevention and management of dementia: a priority for public health. Lancet 2017 Dec 16;390(10113):2614-2615. [doi: 10.1016/S0140-6736(17)31756-7] [Medline: 28735854]
5. Brayne CE, Barnes LE, Breteler MMB, Brooks RL, Dufouil C, Fox C, et al. Dementia research fit for the planet: reflections on population studies of dementia for researchers and policy makers alike. Neuroepidemiology 2020;54(2):157-170 [FREE Full text] [doi: 10.1159/000505626] [Medline: 32018263]
6. Risk reduction of cognitive decline and dementia: WHO guidelines. World Health Organization. 2019 Apr 20. URL: https://www.who.int/mental_health/neurology/dementia/guidelines_risk_reduction/en/ [accessed 2021-04-30]
7. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec 16;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6] [Medline: 28735855]
8. Yu J, Xu W, Tan C, Andrieu S, Suckling J, Evangelou E, et al. Evidence-based prevention of Alzheimer's disease: systematic review and meta-analysis of 243 observational prospective studies and 153 randomised controlled trials. J Neurol Neurosurg Psychiatry 2020 Nov;91(11):1201-1209 [FREE Full text] [doi: 10.1136/jnnp-2019-321913] [Medline: 32690803]
9. Frisoni GB, Molinuevo JL, Altomare D, Carrera E, Barkhof F, Berkhof J, et al. Precision prevention of Alzheimer's and other dementias: anticipating future needs in the control of risk factors and implementation of disease-modifying therapies. Alzheimers Dement 2020 Oct;16(10):1457-1468. [doi: 10.1002/alz.12132] [Medline: 32815289]
10. Anstey KJ, Peters R, Zheng L, Barnes DE, Brayne C, Brodaty H, et al. Future directions for dementia risk reduction and prevention research: an international research network on dementia prevention consensus. J Alzheimers Dis 2020;78(1):3-12 [FREE Full text] [doi: 10.3233/JAD-200674] [Medline: 32925063]
11. Kivipelto M, Mangialasche F, Snyder HM, Allegri R, Andrieu S, Arai H, et al. World-wide FINGERS network: a global approach to risk reduction and prevention of dementia. Alzheimers Dement 2020 Jul;16(7):1078-1094. [doi: 10.1002/alz.12123] [Medline: 32627328]
12. MINDSpeed food and brain training RCT. ClinicalTrials.gov. URL: http://clinicaltrials.gov/ct2/show/NCT03419052 [accessed 2020-03-02]
13. Early intervention in cognitive aging. ClinicalTrials.gov. URL: http://clinicaltrials.gov/ct2/show/NCT02751866 [accessed 2020-03-02]
14. Investigating gains in neurocognition in an intervention trial of exercise. ClinicalTrials.gov. URL: http://clinicaltrials.gov/ct2/show/NCT02875301 [accessed 2020-03-02]
15. Exercise and intensive vascular risk reduction in preventing dementia. ClinicalTrials.gov. URL: http://clinicaltrials.gov/ct2/show/NCT02913664 [accessed 2020-03-02]
16. Effects of an 8-weeks mindfulness-based intervention in individuals with subjective cognitive decline. ClinicalTrials.gov. URL: http://clinicaltrials.gov/ct2/show/NCT03005652 [accessed 2020-03-02]
17. Application of the modified story memory technique to aging. ClinicalTrials.gov. URL: http://clinicaltrials.gov/ct2/show/NCT03370224 [accessed 2020-03-02]
18. Summers MJ, Saunders NL, Valenzuela MJ, Summers JJ, Ritchie K, Robinson A, et al. The Tasmanian healthy brain project (THBP): a prospective longitudinal examination of the effect of university-level education in older adults in preventing age-related cognitive decline and reducing the risk of dementia. Int Psychogeriatr 2013 Jul;25(7):1145-1155. [doi: 10.1017/S1041610213000380] [Medline: 23522602]
19. Kim S, McMaster M, Torres S, Cox KL, Lautenschlager N, Rebok GW, et al. Protocol for a pragmatic randomised controlled trial of body brain life-general practice and a lifestyle modification programme to decrease dementia risk exposure in a primary care setting. BMJ Open 2018 Mar 17;8(3):e019329 [FREE Full text] [doi: 10.1136/bmjopen-2017-019329] [Medline: 29550779]
20. Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological alterations in Alzheimer disease. Cold Spring Harb Perspect Med 2011 Sep;1(1):a006189 [FREE Full text] [doi: 10.1101/cshperspect.a006189] [Medline: 22229116]
21. Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, Proceedings of the Meeting of the International Working Group (IWG) and the American Alzheimer's Association on “The Preclinical State of AD”; July 23, 2015; Washington DC, USA. Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. Alzheimers Dement 2016 Mar;12(3):292-323 [FREE Full text] [doi: 10.1016/j.jalz.2016.02.002] [Medline: 27012484]
22. Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol 2013 Feb;12(2):207-216 [FREE Full text] [doi: 10.1016/S1474-4422(12)70291-0] [Medline: 23332364]

23. Komeili M, Pou-Prom C, Liaqat D, Fraser KC, Yancheva M, Rudzicz F. Talk2Me: automated linguistic data collection for personal assessment. PLoS One 2019;14(3):e0212342 [FREE Full text] [doi: 10.1371/journal.pone.0212342] [Medline: 30917120]
24. Buchman AS, Bennett DA. Loss of motor function in preclinical Alzheimer's disease. Expert Rev Neurother 2011 May;11(5):665-676 [FREE Full text] [doi: 10.1586/ern.11.57] [Medline: 21539487]
25. Albers MW, Gilmore GC, Kaye J, Murphy C, Wingfield A, Bennett DA, et al. At the interface of sensory and motor dysfunctions and Alzheimer's disease. Alzheimers Dement 2015 Jan;11(1):70-98 [FREE Full text] [doi: 10.1016/j.jalz.2014.04.514] [Medline: 25022540]
26. Mollica MA, Tort-Merino A, Navarra J, Fernández-Prieto I, Valech N, Olives J, et al. Early detection of subtle motor dysfunction in cognitively normal subjects with amyloid-β positivity. Cortex 2019 Dec;121:117-124. [doi: 10.1016/j.cortex.2019.07.021] [Medline: 31561128]
27. Cadar D. A life course approach to dementia prevention. J Aging Geriatr Med 2017;1(2):1-6 [FREE Full text]
28. Prince M, Wimo A, Guerchet M, Ali G, Wu Y, Prina M. World Alzheimer report 2019: The global impact of dementia. an analysis of prevalence, incidence, cost and trends. Alzheimer's Disease International. 2015. URL: https://www.alzint.org/u/WorldAlzheimerReport2015.pdf [accessed 2019-10-01]
29. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
30. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci 2012;7:50 [FREE Full text] [doi: 10.1186/1748-5908-7-50] [Medline: 22651257]
31. Smith BJ, Ali S, Quach H. Public knowledge and beliefs about dementia risk reduction: a national survey of Australians. BMC Public Health 2014 Jun 28;14:661 [FREE Full text] [doi: 10.1186/1471-2458-14-661] [Medline: 24972448]
32. Kim S, Sargent-Cox KA, Anstey KJ. A qualitative study of older and middle-aged adults' perception and attitudes towards dementia and dementia risk reduction. J Adv Nurs 2015 Jul;71(7):1694-1703. [doi: 10.1111/jan.12641] [Medline: 25702823]
33. Choi I, La Monica H, Naismith S, Rahmanovic A, Mowszowski L, Glozier N. Communicating personal risk profiles of Alzheimer's disease to older adults: a pilot trial. J Prev Alzheimers Dis 2022;9(1):144-150. [doi: 10.14283/jpad.2021.34] [Medline: 35098985]
34. Michie S, Carey RN, Johnston M, Rothman AJ, de Bruin M, Kelly MP, et al. From theory-inspired to theory-based interventions: a protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action. Ann Behav Med 2016 Jul 11;52:501-512. [doi: 10.1007/s12160-016-9816-6] [Medline: 27401001]
35. Curran E, Chong TWH, Godbee K, Abraham C, Lautenschlager NT, Palmer VJ. General population perspectives of dementia risk reduction and the implications for intervention: a systematic review and thematic synthesis of qualitative evidence. PLoS One 2021;16(9):e0257540 [FREE Full text] [doi: 10.1371/journal.pone.0257540] [Medline: 34534250]
36. Michie S. Designing and implementing behaviour change interventions to improve population health. J Health Serv Res Policy 2008 Oct;13(3):64-69. [doi: 10.1258/jhsrp.2008.008014] [Medline: 18806194]
37. Paterson M, Lagosky S, Mason R. Health promotion and knowledge translation: two roads to the same destination? Glob Health Promot 2018 Sep;25(3):65-69. [doi: 10.1177/1757975916665340] [Medline: 30246629]
38. National statement on ethical conduct in human research. National Health and Medical Research Council. URL: http://www.nhmrc.gov.au/guidelines/publications/e72 [accessed 2022-02-14]
39. 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]
40. Kim S, Sargent-Cox K, Cherbuin N, Anstey KJ. Development of the motivation to change lifestyle and health behaviours for dementia risk reduction scale. Dement Geriatr Cogn Dis Extra 2014 May;4(2):172-183 [FREE Full text] [doi: 10.1159/000362228] [Medline: 25028583]
41. Glanz K, Rimer B, Viswanath K, editors. Health Behavior: Theory, Research and Practice. 5th ed. San Francisco, USA: Josey-Blass; 2015:41-279.
42. Chen G, Gully SM, Eden D. Validation of a new general self-efficacy scale. Organ Res Methods 2016 Jun 29;4(1):62-83. [doi: 10.1177/109442810141004]
43. Chinn D, McCarthy C. All aspects of health literacy scale (AAHLS): developing a tool to measure functional, communicative and critical health literacy in primary healthcare settings. Patient Educ Couns 2013 Feb;90(2):247-253. [doi: 10.1016/j.pec.2012.10.019] [Medline: 23206659]
44. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
45. Robbins TW, James M, Owen AM, Sahakian BJ, McInnes L, Rabbitt P. Cambridge neuropsychological test automated battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. Dementia 1994;5(5):266-281. [Medline: 7951684]
46. Backx R, Skirrow C, Dente P, Barnett JH, Cormack FK. Comparing web-based and lab-based cognitive assessment using the Cambridge neuropsychological test automated battery: a within-subjects counterbalanced study. J Med Internet Res 2020 Aug 04;22(8):e16792 [FREE Full text] [doi: 10.2196/16792] [Medline: 32749999]

47. Alty J, Bai Q, George RJS, Bindoff A, Li R, Lawler K, et al. TasTest: Moving towards a digital screening test for pre‐clinical Alzheimer's disease. Alzheimer's & Dementia 2021 Dec 31;17(S5):e058732. [doi: 10.1002/alz.058732]
48. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [Medline: 6880820]
49. Lubben JE. Assessing social networks among elderly populations. Fam Community Health 1988;11(3):42-52. [doi: 10.1097/00003727-198811000-00008]
50. Hawthorne G, Richardson J, Osborne R. The assessment of quality of life (AQoL) instrument: a psychometric measure of health-related quality of life. Qual Life Res 1999;8(3):209-224. [doi: 10.1023/a:1008815005736]
51. Kemper S, Marquis J, Thompson M. Longitudinal change in language production: effects of aging and dementia on grammatical complexity and propositional content. Psychol Aging 2001 Dec;16(4):600-614. [doi: 10.1037//0882-7974.16.4.600] [Medline: 11766915]
52. Swainson R, Hodges JR, Galton CJ, Semple J, Michael A, Dunn BD, et al. Early detection and differential diagnosis of Alzheimer's disease and depression with neuropsychological tasks. Dement Geriatr Cogn Disord 2001;12(4):265-280. [doi: 10.1159/000051269] [Medline: 11351138]
53. Anstey KJ, Cherbuin N, Herath PM, Qiu C, Kuller LH, Lopez OL, et al. A self-report risk index to predict occurrence of dementia in three independent cohorts of older adults: the ANU-ADRI. PLoS One 2014;9(1):e86141 [FREE Full text] [doi: 10.1371/journal.pone.0086141] [Medline: 24465922]
54. National health survey report no. 4364.0.55.001. Australian Bureau of Statistics. 2017. URL: https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey-state-and-territory-findings/latest-release [accessed 2019-10-30]
55. Eslake S. Education, productivity and economic performance: Tasmania, then, now and tomorrow. 2017 Presented at: 29th John West Memorial Lecture; March 16, 2016; Launceston, Australia URL: http://launcestonhistory.org.au/wp-content/uploads/2017/03/Saul-Eslake-2017.pdf
56. Wicking Dementia Research and Education Centre. The ISLAND project. University of Tasmania. URL: https://island.mooc.utas.edu.au/ [accessed 2022-02-08]
57. Wicking Dementia Research and Education Centre. Preventing dementia massive open online course. University of Tasmania. URL: https://www.utas.edu.au/wicking/preventing-dementia [accessed 2022-02-08]
58. Shah D. Class Central. The top 100 courses of all time. 2019. URL: https://www.classcentral.com/report/top-moocs-2019-edition [accessed 2021-03-24]
59. Farrow M, Doherty K, McInerney F, Klekociuk SZ, Bindoff A, Vickers JC. Improving knowledge and practice through massive open online dementia education: the understanding dementia and preventing dementia MOOCs. 2018 Jul Presented at: Alzheimer's Association International Conference; July 22-26, 2018; Chicago, Illinois. [doi: 10.1016/j.jalz.2018.06.2368]
60. Thow ME, Summers MJ, Saunders NL, Summers JJ, Ritchie K, Vickers JC. Further education improves cognitive reserve and triggers improvement in selective cognitive functions in older adults: the Tasmanian healthy brain project. Alzheimers Dement (Amst) 2018;10:22-30 [FREE Full text] [doi: 10.1016/j.dadm.2017.08.004] [Medline: 29034310]
61. Bindoff AD, Summers MJ, Hill E, Alty J, Vickers JC. Studying at university in later life slows cognitive decline: a long-term prospective study. Alzheimers Dement (N Y) 2021;7(1):e12207 [FREE Full text] [doi: 10.1002/trc2.12207] [Medline: 34522765]
62. Google Maps. 2022. URL: https://www.google.com/maps/place/Tasmania [accessed 2022-01-05]
63. Census data. Australian Bureau of Statistics. 2016. URL: http://stat.data.abs.gov.au/Index.aspx [accessed 2019-11-24]
64. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.
65. Ioannidis JPA. Evolution and translation of research findings: from bench to where? PLoS Clin Trials 2006 Nov 17;1(7):e36 [FREE Full text] [doi: 10.1371/journal.pctr.0010036] [Medline: 17111044]
66. Stefler D, Prina M, Wu Y, Sánchez-Niubò A, Lu W, Haro JM, et al. Socioeconomic inequalities in physical and cognitive functioning: cross-sectional evidence from 37 cohorts across 28 countries in the ATHLOS project. J Epidemiol Community Health 2021 Oct;75(10):980-986 [FREE Full text] [doi: 10.1136/jech-2020-214714] [Medline: 33649052]
67. Fair H, Eccleston C, Doherty K, Klekociuk SZ, Farrow M. The diffusion of dementia risk education. 2020 Dec 07 Presented at: Alzheimer's Association International Conference; July 31-Aug 4, 2020; San Diego, California. [doi: 10.1002/alz.046618]
68. Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, Poblador-Plou B, van den Akker M. Multimorbidity patterns: a systematic review. J Clin Epidemiol 2014 Mar;67(3):254-266. [doi: 10.1016/j.jclinepi.2013.09.021] [Medline: 24472295]
69. Bartlett L, Brady JJR, Farrow M, Kim S, Bindoff A, Fair H, et al. Change in modifiable dementia risk factors during COVID-19 lockdown: the experience of over 50s in Tasmania, Australia. Alzheimers Dement (N Y) 2021;7(1):e12169 [FREE Full text] [doi: 10.1002/trc2.12169] [Medline: 34027023]

Abbreviations

ISLAND: Island Study Linking Aging and Neurodegenerative Disease

Edited by G Eysenbach; submitted 04.11.21; peer-reviewed by T Chong, R Rastmanesh; comments to author 16.12.21; revised version received 22.12.21; accepted 23.12.21; published 01.03.22.

Please cite as:
Bartlett L, Doherty K, Farrow M, Kim S, Hill E, King A, Alty J, Eccleston C, Kitsos A, Bindoff A, Vickers JC
Island Study Linking Aging and Neurodegenerative Disease (ISLAND) Targeting Dementia Risk Reduction: Protocol for a Prospective Web-Based Cohort Study
JMIR Res Protoc 2022;11(3):e34688
URL: https://www.researchprotocols.org/2022/3/e34688
doi:10.2196/34688
PMID:35230251

Protocol

Leveraging Artificial Intelligence to Improve the Diversity of Dermatological Skin Color Pathology: Protocol for an Algorithm Development and Validation Study

Eman Rezk[1*], MSc; Mohamed Eltorki[2*], MBChB; Wael El-Dakhakhni[1*], PhD

[1]School of Computational Science and Engineering, McMaster University, Hamilton, ON, Canada

[2]Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

[*]all authors contributed equally

Corresponding Author:
Eman Rezk, MSc
School of Computational Science and Engineering
McMaster University
1280 Main St W
Hamilton, ON, L8S 4L8
Canada
Phone: 1 9055259140
Email: rezke@mcmaster.ca

Abstract

Background: The paucity of dark skin images in dermatological textbooks and atlases is a reflection of racial injustice in medicine. The underrepresentation of dark skin images makes diagnosing skin pathology in people of color challenging. For conditions such as skin cancer, in which early diagnosis makes a difference between life and death, people of color have worse prognoses and lower survival rates than people with lighter skin tones as a result of delayed or incorrect diagnoses. Recent advances in artificial intelligence, such as deep learning, offer a potential solution that can be achieved by diversifying the mostly light-skin image repositories through generating images for darker skin tones. Thus, facilitating the development of inclusive cancer early diagnosis systems that are trained and tested on diverse images that truly represent human skin tones.

Objective: We aim to develop and evaluate an artificial intelligence–based skin cancer early detection system for all skin tones using clinical images.

Methods: This study consists of four phases: (1) Publicly available skin image repositories will be analyzed to quantify the underrepresentation of darker skin tones, (2) Images will be generated for the underrepresented skin tones, (3) Generated images will be extensively evaluated for realism and disease presentation with quantitative image quality assessment as well as qualitative human expert and nonexpert ratings, and (4) The images will be utilized with available light-skin images to develop a robust skin cancer early detection model.

Results: This study started in September 2020. The first phase of quantifying the underrepresentation of darker skin tones was completed in March 2021. The second phase of generating the images is in progress and will be completed by March 2022. The third phase is expected to be completed by May 2022, and the final phase is expected to be completed by September 2022.

Conclusions: This work is the first step toward expanding skin tone diversity in existing image databases to address the current gap in the underrepresentation of darker skin tones. Once validated, the image bank will be a valuable resource that can potentially be utilized in physician education and in research applications. Furthermore, generated images are expected to improve the generalizability of skin cancer detection. When completed, the model will assist family physicians and general practitioners in evaluating skin lesion severity and in efficient triaging for referral to expert dermatologists. In addition, the model can assist dermatologists in diagnosing skin lesions.

International Registered Report Identifier (IRRID): DERR1-10.2196/34896

(JMIR Res Protoc 2022;11(3):e34896) doi:10.2196/34896

KEYWORDS

artificial intelligence; skin cancer; skin tone diversity; people of color; image blending; deep learning; classification; early diagnosis

Introduction

Background

Dermatology textbooks and atlases lack diversity in skin tones, which propagates structural racism in the health care system [1]. Descriptions and image documentations of differing dermatologic conditions have been largely based on light skin, posing challenges for dermatologists to promptly recognize conditions in darker skins [2]. These challenges may result in serious negative consequences when early diagnosis is crucial, such as in skin cancer. Prior work [3,4] has demonstrated that people of color have worse prognoses and lower survival rates attributed to delayed or incorrect diagnoses. In people of color, squamous cell carcinoma is the most common type of cancer, and delayed diagnosis is linked to higher rates of metastasis and a decrease in the 10-year survival rate to 20% [4]. The second most common cancer is basal cell carcinoma, in which 50% of the cases are pigmented and often misdiagnosed as melanoma, seborrheic keratosis or nevus sebaceous [4]. Finally, people of color are typically diagnosed at more advanced stages of melanoma, which is responsible for 75% of mortality from all skin cancers [5], and this late diagnosis causes a dramatic decrease in the 5-year survival rate to 66.7% compared to 92.5% in individuals with light skin [6].

Artificial Intelligence in Dermatology

Artificial intelligence refers to techniques that allow machines to mimic human behavior to analyze complex data [7]. Deep learning is a leading technology in artificial intelligence that leverages the capabilities of neural networks to analyze complex system structures independently from human intervention. As a result, deep learning has led to breakthroughs in the development of intelligent medical image analysis and diagnosis with performance comparable to that of health care providers [8].

In dermatology, deep learning models have been performing on par with dermatologists in diagnosing skin cancer. In melanoma classification, a convolutional neural network trained on 12,378 dermoscopic images and tested on 100 clinical images performed on par with 145 dermatologists [9]. The convolutional neural network and dermatologists achieved a mean sensitivity of 89.4%, while the convolutional neural network had a specificity of 68.2% compared with a sensitivity of 64.4% for the dermatologists. Furthermore, a convolutional neural network was utilized to assist 12 board-certified dermatologists in melanoma diagnosis [10]. With the support of the convolutional neural network, the mean sensitivity of the dermatologists significantly (P=.003) improved from 59.4% to 74.6%, and the mean accuracy significantly (P=.002) increased from 65.0% to 73.6%.

Haenssle et al [11] compared the performance of a deep learning–based model to that of 96 dermatologists with varying levels of experience in classifying skin lesion dermoscopic images as malignant or benign. The model obtained a sensitivity of 95.0% (95% CI 83.5%-98.6%) and a specificity of 76.7% (95% CI 64.6%-85.6%); however, the dermatologists had a mean sensitivity of 89.0% (95% CI 87.4%-90.6%) and specificity of 80.7% (95% CI 78.8%-82.6%). In another experiment [11] that involved diagnosis using a combination of dermoscopic images and clinical vignettes, dermatologists had a sensitivity of 94.1% (95% CI 93.1%-95.1%) and a specificity of 80.4% (95% CI 78.4%-82.4%). At the same specificity level as the dermatologists, the deep learning–based model had a sensitivity of 95% (95% CI 83.5%-98.6%) [11].

In diagnosing nonpigmented skin cancer [12], the performance of a convolutional neural network trained on 13,724 images (7895 dermoscopic and 5829 clinical) and tested on 2072 images was compared with that of beginner, intermediate, and expert dermatologists (95 categorized by years of experience) and achieved an area under the curve (AUC) of 0.742 (95% CI 0.729-0.755) compared with an AUC of 0.695 (95% CI 0.676-0.713) for the dermatologists. For particularly challenging conditions, 37.6% (95% CI 36.6%-38.4%) of the network's diagnoses were correct, which was higher than the corresponding percentages for beginner and intermediate dermatologists but less than that of expert dermatologists (accuracy 40.0%, 95% CI 37.0%-43.0%). Given the low number of expert dermatologists in many health care jurisdictions [13], this technology would be very useful in assisting nonexpert dermatologists (eg, general practitioners) in triaging skin lesions that need a referral to an expert dermatologist for advanced assessment.

Unfortunately, despite these advancements, the ability to take full advantage of deep learning capabilities is limited due to the lack of adequate quantity and quality of real-world data. Furthermore, and notwithstanding its advantages, the use of deep learning has further put people with darker skin tones at a disadvantage, because training data used in the development of published models lack the true breadth of human skin tones [14,15]. The paucity of nonwhite skin tones in training data limits the generalizability of developed models to nonwhite skin tones. For example, a deep learning–based classifier to diagnose 12 malignant and benign skin lesions using clinical images [16] trained on Asian skin images and validated with Caucasian images classified basal cell carcinoma with an AUC of 0.78, SD 0.02. However, when the training data was augmented by including Caucasian images, the AUC improved to 0.90, SD 0.01. Despite of the excellent performance metrics during development and validation, the model's generalizability was deficient when tested on a different patient population [17].

The lack of skin tone diversity also limits the study of the relationship between skin tone and diagnostic accuracy. Analysis performed on 2 publicly available data sets—the international skin imaging collaboration (ISIC), which consists of 10,015 dermoscopic images from 7 skin diseases data set, and SD-198, which consists of 6548 clinical images from 198 skin diseases [18]—showed no measurable correlation between classification accuracy and skin tone, which was attributed to the fact that the images in the data sets consisted mainly of light skin tones [18].

Study Goals

Several promising initiatives are being employed to address the lack of diversity, such as a deliberate focus on increasing the diversity of students in medical schools and postgraduate dermatology training programs [19,20], emerging textbooks and literature targeting darker skin pathology [21], and new

websites to build a database of skin pathology in darker skin tones [1]; however, these efforts are complex, challenging, and take time, thus, implementing a rapid yet effective solution to address this gap is imperative.

We aim to (1) identify the underrepresented tones in publicly available dermatology clinical image atlases; (2) generate realistic images for darker skin with closely related malignant and benign conditions; (3) extensively evaluate the images, using quantitative ratings as well as qualitative human expert and nonexpert ratings; and (4) develop a classification model using several deep learning networks to detect malignancy on all skin tones.

Methods

Overview

This work has 4 main phases (Figure 1). The focus of this work is on common malignant skin pathology and closely related disorders that resemble those malignant lesions and form part of their differential diagnosis; therefore, we collected all clinical images representing those conditions such as basal cell carcinoma, squamous cell carcinoma, melanoma, and nevus from DermNet NZ (994 images) [22], ISIC 2018 (100 images) [23], and a dermatology atlas (607 images) [24]. Images from ISIC and DermNet NZ (Set A) were utilized in training, finetuning, and internal validation. The dermatology atlas images (Set B) will be utilized only for testing the classification model in phase 4 (Table 1).

Figure 1. Phases of the proposed work.

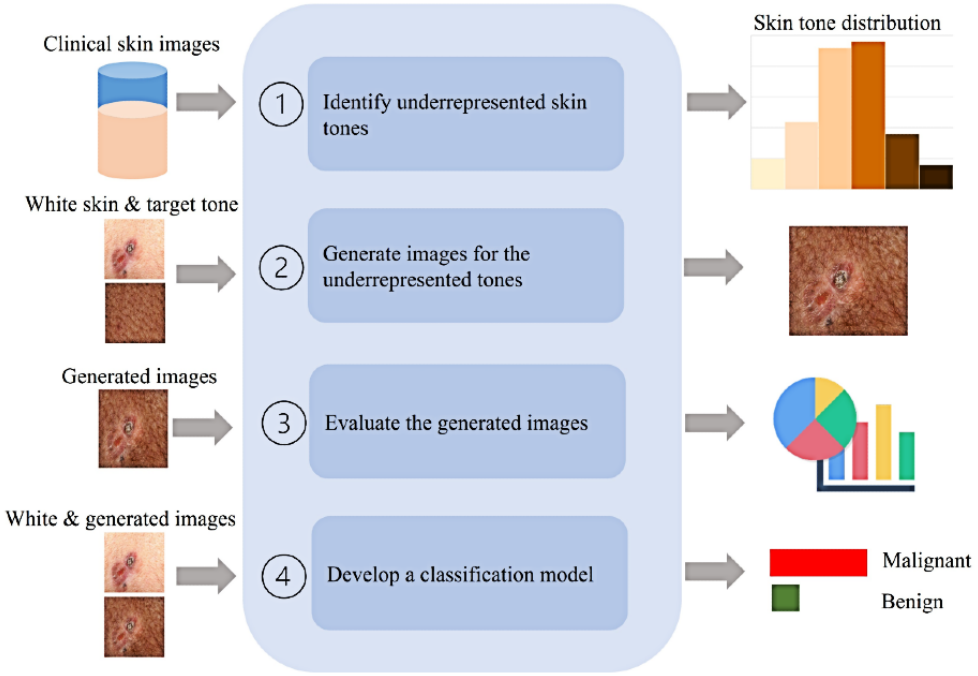


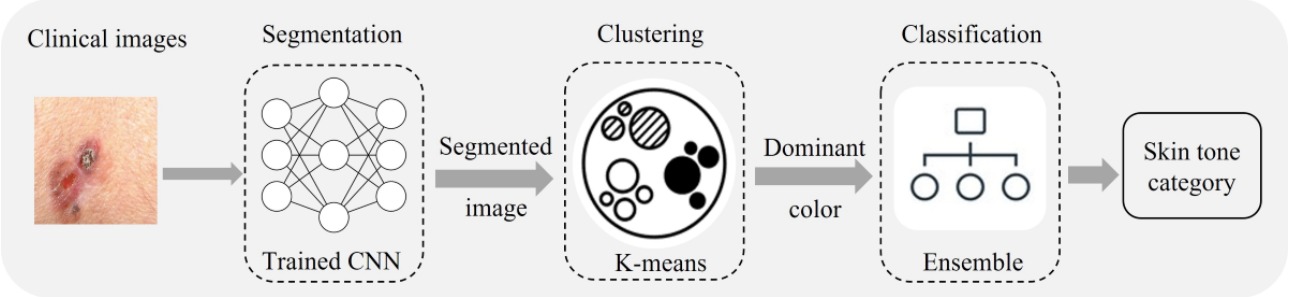
Table 1. Image distribution for training, finetuning, and internal validation (Set A) and for testing (Set B).

Class	Set A (n=1094), n (%)	Set B (n=607), n (%)
Malignant	634 (58)	508 (83.7)
Benign	460 (42)	99 (16.3)

Phase 1: Underrepresented Skin Tones Identification

The goal of this phase (Figure 2) is to analyze the images to determine the skin tone distribution and quantify the underrepresentation of the darker tones in the data sets.

Figure 2. Modules of phase 1. CNN: convolutional neural network.

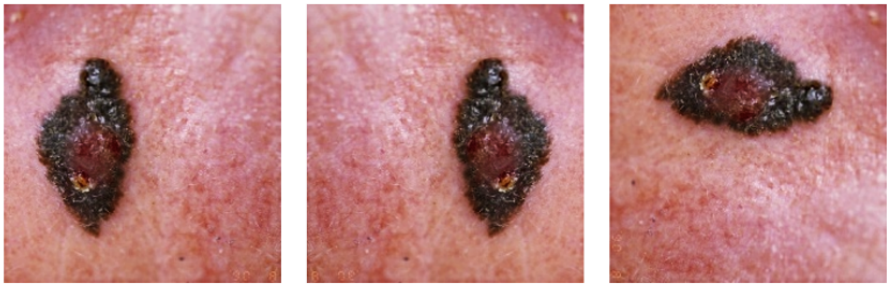


Skin Image Augmented Segmentation

Image data augmentation is the process of applying transformations on the images such as flipping, cropping, and rotating to increase the training set size, improve training data variance, and reduce overfitting [25]. A set of 500 clinical images was randomly selected from Set A; images were horizontally flipped and rotated by 90° to yield additional images for the data set (Figure 3).

The images (500 original and 1000 augmented) were utilized to train a segmentation network to separate the disease region from the underlying skin and allow the analysis of the skin color. Trained segmentation networks can be used to improve skin image classification [26]. A deep learning–based image segmentation network developed by Azad et al [27] was employed due to its high accuracy in segmenting dermoscopic skin images. The network was initially trained on 2594 dermoscopic skin images [27]. We adapted the network to segment clinical skin images through transfer learning, by performing an additional cycle of training on the original and augmented images (1500). The training data were randomly split into 1200 (80%) images for training, 150 (10%) images for validation, and 150 (10%) images for testing. Because segmentation classifies each pixel in the image as disease or skin, it can be considered a binary classification problem; therefore, the segmentation model was assessed using accuracy, sensitivity, specificity, AUC, and Jaccard similarity.

Figure 3. Data augmentation on skin image (A) original image (B) horizontal flipping (C) 90° rotation.



(A) (B) (C)

Skin Image Clustering

The segmented normal skin regions in the images were preprocessed to detect and remove any possible nonsegmented disease pixels resulting from a variation of the skin color, improve the quality of the images, and allow for accurate skin tone analyses. The contrast of the images was enhanced (Figure 4) using the contrast-limited adaptive histogram equalization, which has been widely applied in medical image enhancement, such as retinal fundus images [28], breast mammography [29], and bone fracture images [29]. The contrast enhancement algorithm divides the image into sections and creates a histogram for each section that is utilized to redistribute the brightness across the image. This method has outperformed other contrast enhancement methods because it limits the amplification of the contrast across the image and hence reduces noise [30].

Applying contrast adjustment helped to identify pigmented spots and made them easier to remove. A thresholding technique (Figure 5) that analyzed the updated image histogram to classify each pixel as foreground or background [31] was implemented to detect and remove any objects on the skin such as colored spots and hair. As a result, the image included only the pixels that truly reflected skin color.

Processed skin pixels were subsequently analyzed to determine the dominant skin tone using k-means clustering [32] to group the pixels based on color values. The number of clusters was selected to minimize the sum of squared errors and improve in-cluster cohesion [33]. The cluster that had the maximum number of pixels was considered the dominant cluster, and its center was considered the dominant color.

Figure 4. Skin image histogram (A) before and (B) after equalization.

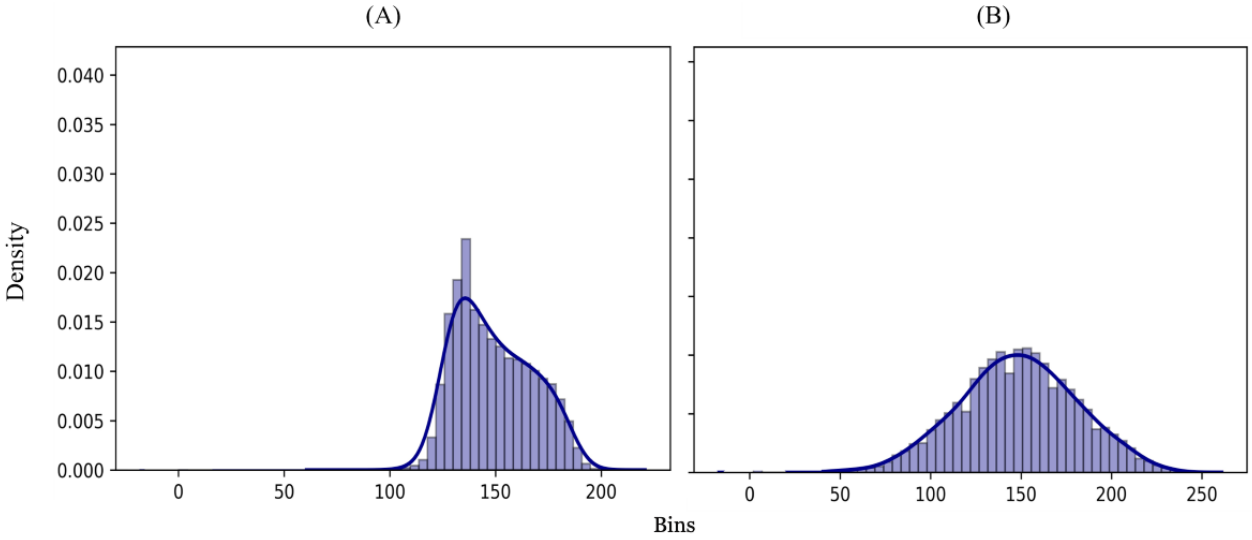
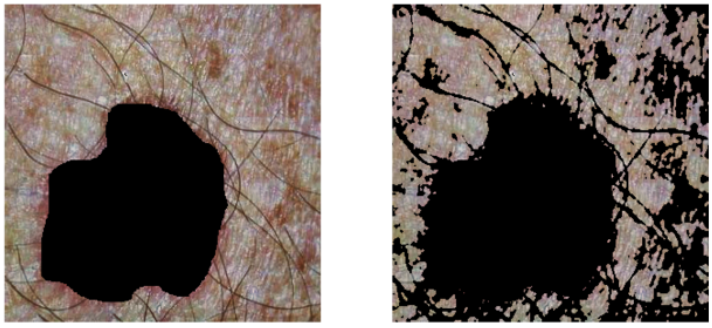


Figure 5. (A) Original and (B) contrast-adjusted (with thresholding) image.

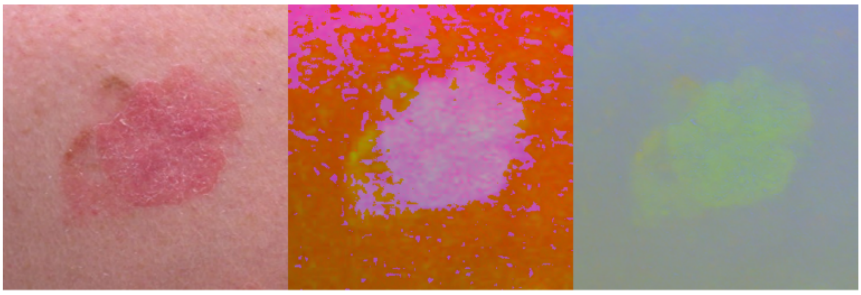


Skin Tone Classification

Although skin tone is easy to perceive, it is challenging to evaluate quantitively. Fitzpatrick [34] defined 6 skin color categories (very light, light, intermediate, tan, brown, and dark). The Melanin Index is one of the most reliable metrics to quantify human skin color based on skin reflectance (using a reflectance spectrophotometer) [35]. However, the use of the melanin index is limited to dermatologists as it requires specialized equipment. Individual typology angle is another metric to evaluate the skin tone category in which the skin color is utilized to calculate an angle that can be assigned to a skin category [36]. Unfortunately, the former skin categorization approach has exhibited inconsistencies and inaccuracies compared to the perceived skin tones [18].

We developed a skin categorization model that classified a dominant skin color into a skin category. An ensemble model, which included k-nearest neighbor [37], random forest [38], and naïve Bayes [39] methods, was implemented to classify the dominant skin color as very light, light, intermediate, tan, brown, or black. For training and validation, a set of 100 skin color variations, represented in the RGB (red, green, blue) color space, was collected from the human skin color database [40]. RGB features were processed to create supplementary color features from different color spaces such as HSV (hue, saturation, value) and Lab (L represents luminance, *a* represents the range from red to green, and b represents the range from blue to yellow) [41] to provide the model with sufficient color information (Figure 6). The model was tested on the dominant colors extracted from the clustering step and evaluated using accuracy and AUC.

Figure 6. (A) RGB (red, green, blue); (B) HSV (hue, saturation, value); and (C) Lab (luminance, red–green, blue–yellow) color spaces.

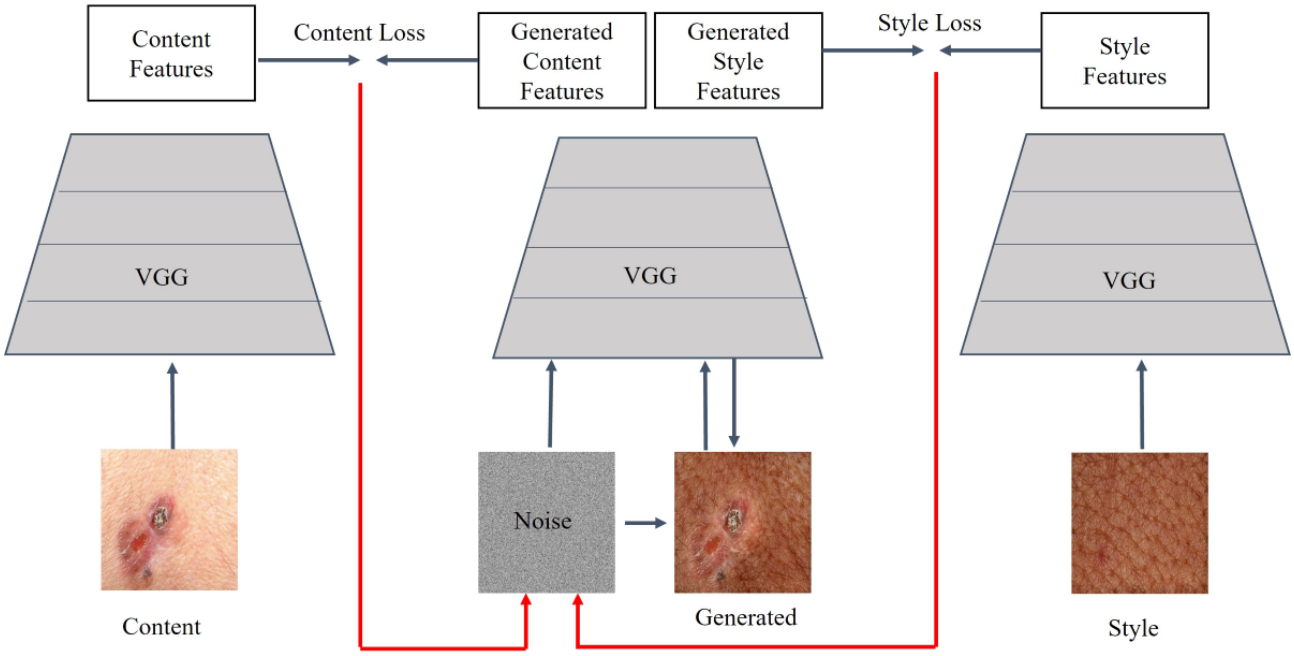


Phase 2: Image Generations for Underrepresented Tones

Style transfer and deep blending image generation methods will be investigated. Both methods are based on image feature extraction and blending using the Visual Geometry Group network trained on the ImageNet database with millions of images for object localization and recognition [42]. Given that convolutional neural networks trained using sufficient labeled data on object recognition are capable of extracting high-level feature representations regardless of the data set [43], style transfer and deep blending can be generalized to the skin image generation problem.

Style transfer has been mainly applied to create stylized artwork [44]. This method will be used to generate skin images with dark skin tones by extracting the features of (1) a content image containing the skin pathological condition and (2) a style image with the target skin color. A new image containing a weighted blend of both feature sets will then be generated starting from noise and iteratively improving by minimizing the content and style loss (Figure 7).

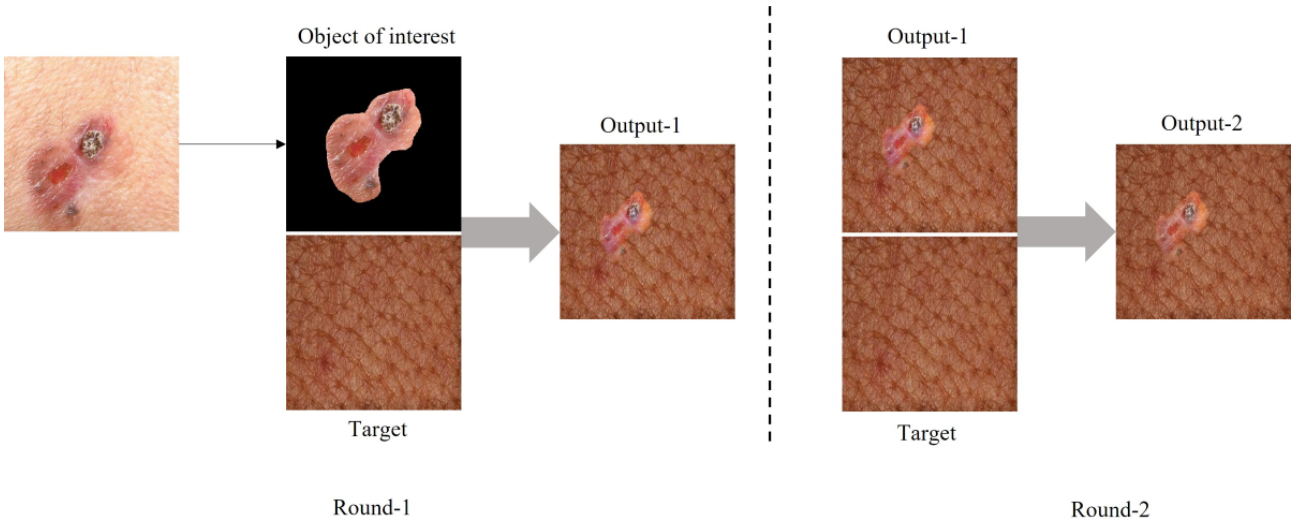
Figure 7. Style transfer procedure. VGG: Visual Geometry Group.



Deep blending has been used to blend an object with a target background image technically, as in style transfer; however deep blending has 3 main differences [45]: only the object of interest in the source (content) image is blended with the target (style) image, thus it requires a segmentation mask, a new loss metric is added to the content and style loss to minimize the sharp intensity change between the source object and target image, and 2 rounds of blending are performed—one with the source object and target image, and the second with the output image of the first round and the target image (Figure 8). We will use this method to (1) achieve a seamless blending of the content disease region and the target skin features and (2) use the image generated from the first step with the style image as input to the network to impose the target skin features (eg, color and texture) on the disease region and the blending boundary to provide a smooth realistic image.

The parameters of style transfer and deep blending methods, such as the number of network layers, the content to style weights, and the degree of lesion pigmentation, will be finetuned. The methods will be employed to generate images for underrepresented skin tones (based on the findings of phase 1).

Figure 8. Deep blending rounds.



Phase 3: Generated Image Evaluations

Overview

The images that are generated will be assessed quantitively and qualitatively. Primarily image realism and disease presentation will be evaluated through numerical image quality metrics as well as human expert and nonexpert rating. The best performing image generation technique will be utilized to generate diverse images to train the classification model.

Quantitative Evaluation

Quantitative evaluation of the generated images will be performed using 2 image quality assessment metrics—the blind reference-less image spatial quality evaluator, a reference-less metric that quantifies the loss of image realism in the presence of distortions by extracting 18 statistical features to assign a quality score with a support vector machine regressor [46], and the structural similarity index measure, which compares the structure, texture, and edges of a reference image (the original image) with a modified image and provides a similarity score [47].

Qualitative Evaluation

The human visual Turing test, wherein participants are asked to classify images as real or generated, will be conducted. Participants of this test will be medical personnel with varying experience and nonmedical personnel. The classification accuracy, false positive rate (the ratio of generated images classified as real), and true positive rate (the ratio of real images classified as real) will be calculated. Furthermore, a regression model will be implemented to study the significance of the participants' background in distinguishing the generated images.

Disease identification will also be conducted, which will include solely dermatologists with different years of experience as participants, to evaluate the accuracy of disease presentation in the generated images. The dermatologists will be asked to choose the disease that best describes a set of real and generated images with various malignant and benign conditions. The rate of correctly identified images will be calculated for each disease, image group (real or generated), and skin tone. In addition, disease misdiagnosis rates will be compared with that in published literature pertaining to misdiagnoses in skin of color.

Phase 4: Classification Model Development

The goal of this phase is to develop a malignancy detection model using real and generated diverse images. To develop the model, skin images from Set A and dark skin images generated in phase 2 will be utilized for training and validation. Set B will be used as an independent test set. Set A contains primarily images with light skin (collected from New Zealand). Set B was collected from Brazil where there are varying skin tones (based on Fitzpatrick classification) in the population compared with that of New Zealand [48]. The diversity in Set B, which was confirmed in phase 1, will allow the generalizability of the model to darker skin tones to be evaluated and the impact of the generated images on the classification accuracy to be determined.

Several classification networks will be trained and validated on Set A after handling data skewness. Data augmentation techniques will be utilized to balance class distribution and ensure that the model is not biased toward any class. Given that the developed model will be well balanced, no augmentation on Set B will be performed during the testing phase.

As training deep learning–based classification networks requires large data sets, adapting pretrained networks is important to make use of the network's calculated weights instead of starting from random weights which requires more training data. Transferring knowledge from networks pretrained on a large number of images, then enriching that knowledge to classify skin images helps to overcome the lack of data. Therefore, deep learning network architectures such as GoogLeNet [49] and ResNet [50], initially trained with millions of natural images from the ImageNet data set, will be adapted. Existing weights obtained from pretraining will be customized to fit skin image classification.

The classification process will follow 2 approaches to evaluate the effect of the generated images on skin tone diversity, classification accuracy, and generalizability. (1) The convolutional neural networks will be trained on 80% of Set A

(randomly selected) and their corresponding generated darker color images. The remaining 20% will be used for validation while building the model and to update the network weights. This approach will help increase the number of training instances and is expected to familiarize the network with diverse human skin tones. (2) The convolutional neural networks will be trained on the same 80% of Set A images and their corresponding augmented images. The remaining 20% will be utilized for validation, thus the training set and the validation set in both approaches have the same sizes and same original images. In both approaches, Set B will be utilized for testing.

Accuracy boosting will be attempted by integrating supplementary information as separate features, such as skin color category, lesion anatomic distribution, and lesion textual description (lesion color, shape, texture, clinical presentation, associated conditions such as scarring and inflammation). For example, in people of color, pigmented basal cell carcinoma is more prevalent, which will be captured by the generated images, however, some features, such as the solitary papule appearance, will be provided as textual description [51] to address missed appearance factors in the generated images, which will improve malignancy detection in the skin of color where lesions might look different or can be associated with unusual signs [51].

The models will be evaluated using accuracy and AUC. We will compare model classification performance to that of the dermatologists and report the dermatologists' diagnosis performance with the aid of the developed model as a second opinion. In addition, the correlation between performance measures and skin tone will be calculated.

Results

Phase 1 was initiated in September 2020 and completed in March 2021. Phase 2 was subsequently initiated and will be completed in March 2022. Phase 3 and phase 4 will be conducted in parallel; the study is expected to be completed by September 2022.

In Figure 9, a comparison between the segmentation model with and without training on clinical skin images shows that training improves all performance metrics. Accuracy increased from 0.88 (95% CI 0.8798-0.8802) to 0.94 (95% CI 0.9399-0.9401), sensitivity slightly increased from 0.72 (95% CI 0.7197-0.7203) to 0.73 (95% CI 0.7297-0.7303), specificity significantly increased from 0.91 (95% CI 0.9098-0.9102) to 0.98 (95% CI 0.9799-0.9801), AUC increased from 0.82 (95% CI 0.8155-0.8162) to 0.85 (95% CI 0.8538-0.8544), and Jaccard similarity increased from 0.88 (95% CI 0.8798-0.8802) to 0.94 (95% CI 0.9399-0.9401).

Segmentation masks that outlined the region of the disease improved with training (Figure 10). After image enhancement and thresholding, the bar of clusters' center color is plotted; the first cluster is the largest and dominant one (Figure 11). In classifying the dominant colors into skin tone categories (Figure 12), the model performed well (accuracy 0.95, 95% CI 0.86-1.0; AUC 0.98, 95% CI 0.92-1.0).

In Set A, more than 60% of the images were light skin, 20% were intermediate, and only 20% were tan, brown, and black; in Set B, 45.5% of the images were tan, brown, and black (Figure 13).

Figure 9. Performance measures of the segmentation before and after training.

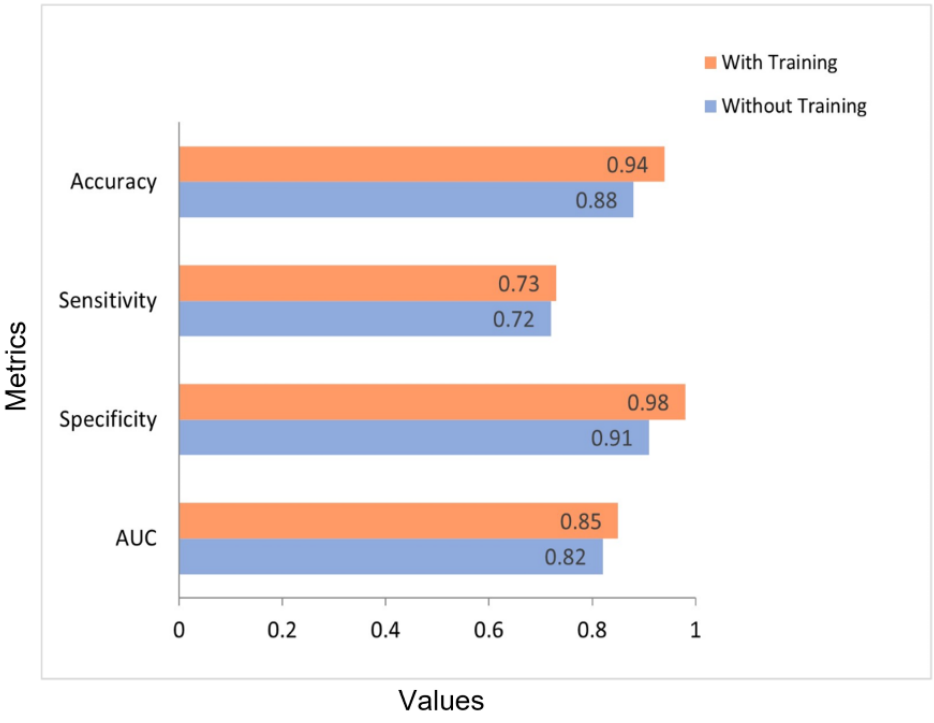


Figure 10. (A) input; (B) without training; and (C) with training.

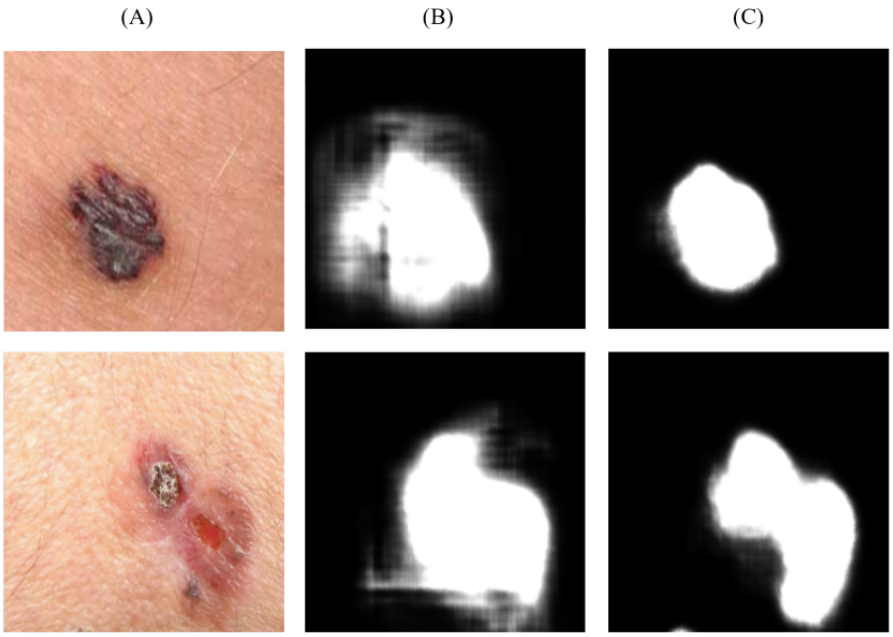


Figure 11. Clustering results.

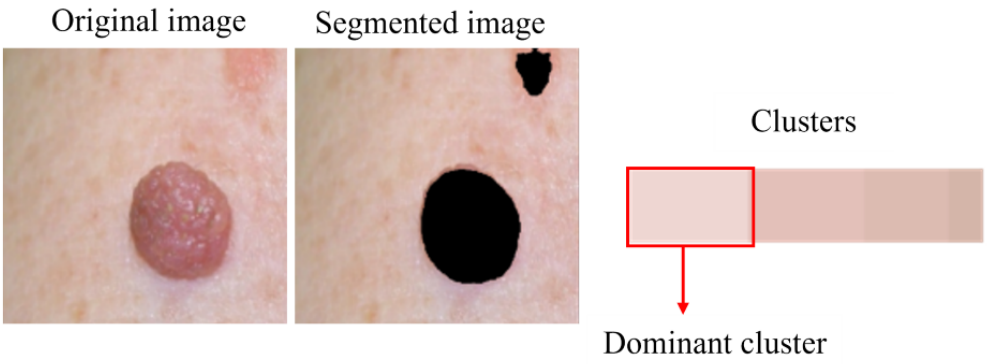
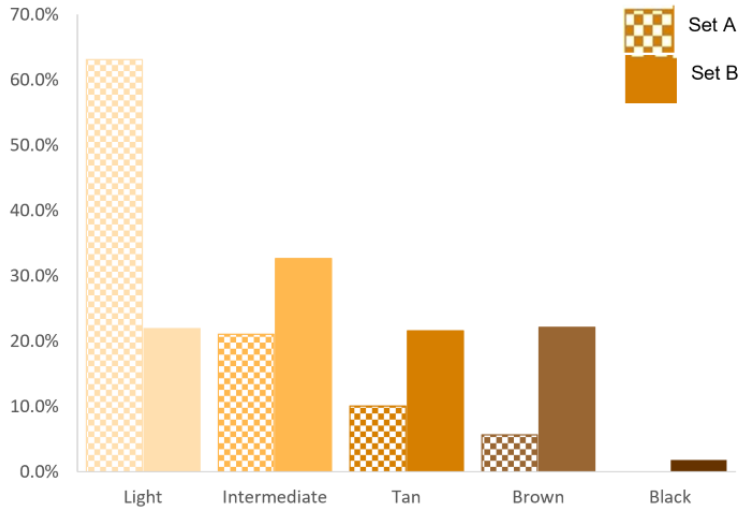


Figure 12. Skin tone categorization results.

Figure 13. Skin tone distribution in set A and set B.



Discussion

General

Currently, we have successfully completed the first phase and the second phase is in progress. Based on the results of phase 1, phase 2 will be configured to generate images for the underrepresented skin tones in Set A, thus a diverse image bank will be created. Phase 3 when completed will extensively evaluate the generated images to ensure their quality and high disease presentation. Since an image library that reflects the true breadth of human skin tones will become available, we will build generalizable deep learning models to detect malignancy. Moreover, the detection accuracy will be boosted by employing supplementary features extracted from disease clinical presentation and anatomic distribution. Furthermore, the correlation between classification accuracy and skin tones can be studied, and the performance of dermatologists with the aid of the skin cancer early detection system will be assessed in phase 4.

The development of the integrated artificial intelligence-based skin cancer early detection system for all skin tones incorporates 4 main milestones: identifying underrepresented skin tones, generating a diverse clinical image bank for various malignant and benign conditions, broadly evaluating the generated images, and developing a generalizable classifier to detect malignancy in any skin tone.

The system is designed to analyze clinical images to increase its usability as digital cameras are easily accessible. The system will advantage all skin tones, consequently increasing the dark skin tone inclusion. The system is also expected to raise the clinical index of suspicion and boost the detection rates of malignant lesions. Finally, the system will assist prioritize patients' referrals to expert dermatologists, for faster diagnosis and help dermatologists with malignancy detection.

The preliminary findings show that the segmentation components demonstrate high accuracy and the quality of the pilot-generated images is promising.

Comparison With Prior Work

Image generation using deep learning had been applied to improve pathology diversity and balance the data. Generative adversarial networks have been utilized in a prior study [52] to generate realistic dermoscopic skin images for various malignant and benign conditions to overcome data skewness. Unlike clinical images, dermoscopic images are captured while zooming in to focus on the disease, thus skin tone was not a factor.

Generative adversarial networks have been also utilized to generate clinical skin images for 8 skin conditions (eg, skin tag and melanoma). In [53], a semantic map encoding each input image was manually generated and given as part of the input while training, testing, and generating images, which significantly limited the system's applicability. Although it was possible to generate images with different skin tones using semantic maps, there was no focus on the darker skin tones. In addition, the ratio of generated images detected as real was relatively low (0.30) and the ratio of correctly identifying skin disease was 0.45. Moreover, the number of participants involved in image quality assessment was insufficient to draw significant conclusions.

To the best of our knowledge, to date, no published study has focused on generating clinical images that reflect the diversity of skin tones. Accompanied by the etiological factors and the anatomic distribution of skin cancer in people of color, the images that are generated will be a valuable resource that can be utilized in education and research, in addition, to its use in this study.

Limitations

The proposed system has limitations, the quality of the original images is vital in generating high-quality images. The system depends on publicly available data sets, thus there are data

availability limitations; however, we will overcome this limitation by implementing data-efficient and pretrained deep learning methods. Compared with the real images of darker skin tones, generated images might not include all factors that can affect the appearance of the skin lesions on people of color. We will mitigate this issue by reflecting the pigmentation on the lesion consistently with the skin color, and the classification model will be supplemented with textual features detailing key clinical presentations of different lesions on darker skin tones.

Implications

Until a real-image repository collected from patients with diverse skin tones becomes available, this study is the first step to filling the gap regarding skin tone diversity toward the goal of achieving racial equity in dermatology diagnosis. The generated image bank will increase the inclusion of dark skin images in dermatology research, the study will leverage the capabilities of deep learning–based cancer detection methods using these images. Our model can be trained in the future with real images of dark skin once dermatology atlases with a large number of dark skin pathologies become available.

Acknowledgments

This research is funded by the Natural Sciences and Engineering Research Council of Canada. The sponsor has no role in the design and conduct of the study.

Conflicts of Interest

None declared.

References

1. Buonsenso D, Roland D, Abela N, Shanmugavadivel D, Platt R, Davis T, DFTB Skin Deep core team. Time to improve the diversity of paediatric skin images. Lancet Child Adolesc Health 2020 Nov;4(11):e44. [doi: 10.1016/S2352-4642(20)30303-5] [Medline: 33068551]
2. Barve A, Shah M, Sheth P. Skin of color images in pediatric dermatology textbooks. Pediatr Dermatol 2021 Nov 16;38:144-145. [doi: 10.1111/pde.14760] [Medline: 34402099]
3. Buchanan Lunsford N, Berktold J, Holman DM, Stein K, Prempeh A, Yerkes A. Skin cancer knowledge, awareness, beliefs and preventive behaviors among black and hispanic men and women. Prev Med Rep 2018 Dec;12:203-209 [FREE Full text] [doi: 10.1016/j.pmedr.2018.09.017] [Medline: 30364862]
4. Bradford PT. Skin cancer in skin of color. Dermatol Nurs 2009;21(4):170-178 [FREE Full text] [Medline: 19691228]
5. Qiu CC, Brown AE, Lobitz GR, Shanker A, Hsu S. The color of skin: black diseases of the skin, nails, and mucosa. Clin Dermatol 2019;37(5):447-467. [doi: 10.1016/j.clindermatol.2019.08.003] [Medline: 31896402]
6. Jothishankar B, Stein SL. Impact of skin color and ethnicity. Clin Dermatol 2019;37(5):418-429. [doi: 10.1016/j.clindermatol.2019.07.009] [Medline: 31896399]
7. Hogarty DT, Su JC, Phan K, Attia M, Hossny M, Nahavandi S, et al. Artificial intelligence in dermatology-where we are and the way to the future: a review. Am J Clin Dermatol 2020 Feb;21(1):41-47. [doi: 10.1007/s40257-019-00462-6] [Medline: 31278649]
8. Young AT, Xiong M, Pfau J, Keiser MJ, Wei ML. Artificial intelligence in dermatology: a primer. J Invest Dermatol 2020 Aug;140(8):1504-1512 [FREE Full text] [doi: 10.1016/j.jid.2020.02.026] [Medline: 32229141]
9. Brinker TJ, Hekler A, Enk AH, Klode J, Hauschild A, Berking C, Collaborators. A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. Eur J Cancer 2019 Apr;111:148-154 [FREE Full text] [doi: 10.1016/j.ejca.2019.02.005] [Medline: 30852421]
10. Maron RC, Utikal JS, Hekler A, Hauschild A, Sattler E, Sondermann W, et al. Artificial intelligence and its effect on dermatologists' accuracy in dermoscopic melanoma image classification: web-based survey study. J Med Internet Res 2020 Sep 11;22(9):e18091 [FREE Full text] [doi: 10.2196/18091] [Medline: 32915161]
11. Haenssle HA, Fink C, Toberer F, Winkler J, Stolz W, Deinlein T, Reader Study Level ILevel II Groups. Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions. Ann Oncol 2020 Jan;31(1):137-143 [FREE Full text] [doi: 10.1016/j.annonc.2019.10.013] [Medline: 31912788]
12. Tschandl P, Rosendahl C, Akay BN, Argenziano G, Blum A, Braun RP, et al. Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks. JAMA Dermatol 2019 Jan 01;155(1):58-65 [FREE Full text] [doi: 10.1001/jamadermatol.2018.4378] [Medline: 30484822]
13. Mydlarski PR, Parsons LM, Pierscianowski TA, Kirchhof MG, Rosen CF, Purdy K, et al. Dermatologic training and practice in Canada: an in-depth review. J Cutan Med Surg 2020;24(3):297-303. [doi: 10.1177/1203475420914619] [Medline: 32293201]

14. Chan S, Reddy V, Myers B, Thibodeaux Q, Brownstone N, Liao W. Machine learning in dermatology: current applications, opportunities, and limitations. Dermatol Ther (Heidelb) 2020 Jun;10(3):365-386 [FREE Full text] [doi: 10.1007/s13555-020-00372-0] [Medline: 32253623]
15. Schlessinger DI, Chhor G, Gevaert O, Swetter SM, Ko J, Novoa RA. Artificial intelligence and dermatology: opportunities, challenges, and future directions. Semin Cutan Med Surg 2019 Mar 01;38(1):E31-E37. [doi: 10.12788/j.sder.2019] [Medline: 31051021]
16. Han SS, Kim MS, Lim W, Park GH, Park I, Chang SE. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. J Invest Dermatol 2018 Jul;138(7):1529-1538 [FREE Full text] [doi: 10.1016/j.jid.2018.01.028] [Medline: 29428356]
17. Navarrete-Dechent C, Dusza SW, Liopyris K, Marghoob AA, Halpern AC, Marchetti MA. Automated dermatological diagnosis: hype or reality? J Invest Dermatol 2018 Oct;138(10):2277-2279 [FREE Full text] [doi: 10.1016/j.jid.2018.04.040] [Medline: 29864435]
18. Kinyanjui N, Odonga T, Cintas C, Codella N, Panda R, Sattigeri P. Estimating skin tone and effects on classification performance in dermatology data sets. 2019 Presented at: NeurIPS Workshop on Fair ML for Health; December 14, 2019; Vancouver, Canada.
19. Pandya AG, Alexis AF, Berger TG, Wintroub BU. Increasing racial and ethnic diversity in dermatology: a call to action. J Am Acad Dermatol 2016 Mar;74(3):584-587. [doi: 10.1016/j.jaad.2015.10.044] [Medline: 26774427]
20. Pritchett EN, Pandya AG, Ferguson NN, Hu S, Ortega-Loayza AG, Lim HW. Diversity in dermatology: roadmap for improvement. J Am Acad Dermatol 2018 Aug;79(2):337-341. [doi: 10.1016/j.jaad.2018.04.003] [Medline: 29653209]
21. Mukwende M, Peter T, Margot T. Mind the gap: a handbook of clinical signs in black and brown skin. St George's University of London. 2020. URL: https://sgul.figshare.com/articles/online_resource/Mind_the_Gap_A_handbook_of_clinical_signs_in_Black_and_Brown_skin/12769988/1 [accessed 2022-02-15]
22. Oakley A. DermNet NZ. 2013. URL: https://dermnetnz.org/ [accessed 2022-02-15]
23. ISIC JID editorial images. International Skin Imaging Collaboration Archive. 2018. URL: https://www.isic-archive.com/#!/topWithHeader/onlyHeaderTop/gallery?filter=%5B%22meta.data [accessed 2022-02-15]
24. Silva S. Dermatology Atlas. 1999. URL: http://www.atlasdermatologico.com.br/index.jsf [accessed 2022-02-15]
25. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019 Jul 6;6(1):1-48. [doi: 10.1186/s40537-019-0197-0]
26. Maron RC, Hekler A, Krieghoff-Henning E, Schmitt M, Schlager JG, Utikal JS, et al. Reducing the impact of confounding factors on skin cancer classification via image segmentation: technical model study. J Med Internet Res 2021 Mar 25;23(3):e21695 [FREE Full text] [doi: 10.2196/21695] [Medline: 33764307]
27. Azad R, Asadi-Aghbolaghi M, Fathy M, Escalera S. Bi-directional ConvLSTM U-net with densley connected convolutions. 2019 Presented at: IEEE/CVF International Conference on Computer Vision Workshop; Oct 27, 2019; Seoul, Korea.
28. Sonali, Sahu S, Singh AK, Ghrera S, Elhoseny M. An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. Opt Laser Technol 2019 Feb;110:87-98. [doi: 10.1016/j.optlastec.2018.06.061]
29. Campos GFC, Mastelini SM, Aguiar GJ, Mantovani RG, Melo LFD, Barbon S. Machine learning hyperparameter selection for contrast limited adaptive histogram equalization. J Image Video Proc 2019 May 6;2019(1):59. [doi: 10.1186/s13640-019-0445-4]
30. Pizer SM, Amburn EP, Austin JD, Cromartie R, Geselowitz A, Greer T, et al. Adaptive histogram equalization and its variations. Comput Vis Graph Image Process 1987 Sep;39(3):355-368. [doi: 10.1016/s0734-189x(87)80186-x]
31. Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. J Electron Imaging 2004 Jan 01;13(1):146-165. [doi: 10.1117/1.1631315]
32. Mohanty A, Rajkumar S, Muzaffar Mir Z, Bardhan P. Analysis of color images using cluster based segmentation techniques. Int J Comput Appl 2013 Oct 18;79(2):42-47. [doi: 10.5120/13716-1487]
33. Shi C, Wei B, Wei S, Wang W, Liu H, Liu J. A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. J Wireless Com Network 2021 Feb 15;2021(1):11. [doi: 10.1186/s13638-021-01910-w]
34. Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. Arch Dermatol 1988 Jun;124(6):869-871. [Medline: 3377516]
35. Stamatas GN, Zmudzka BZ, Kollias N, Beer JZ. In vivo measurement of skin erythema and pigmentation: new means of implementation of diffuse reflectance spectroscopy with a commercial instrument. Br J Dermatol 2008 Sep;159(3):683-690. [doi: 10.1111/j.1365-2133.2008.08642.x] [Medline: 18510669]
36. Wilkes M, Wright CY, du Plessis JL, Reeder A. Fitzpatrick skin type, individual typology angle, and melanin index in an African population: steps toward universally applicable skin photosensitivity assessments. JAMA Dermatol 2015 Aug;151(8):902-903. [doi: 10.1001/jamadermatol.2015.0351] [Medline: 25923982]
37. Kubat M. Similarities: Nearest-Neighbor Classifiers. In: An Introduction to Machine Learning. Cham: Springer International Publishing; 2015:43-64.
38. Breiman L. Random forests. Machine Learning 2001;45:5-32. [doi: 10.1023/A:1010933404324]
39. Stephens CR, Huerta HF, Linares AR. When is the naive Bayes approximation not so naive? Mach Learn 2017 Jul 21;107(2):397-441. [doi: 10.1007/s10994-017-5658-0]

40. Scheme Color. URL: https://www.schemecolor.com/real-skin-tones-color-palette.php [accessed 2022-02-15]
41. Lionnie R, Alaydrus M. A comparison of human skin color detection for biometrie identification. 2017 Presented at: International Conference on Broadband Communication, Wireless Sensors and Powering; November 21-23, 2017; Jakarta, Indonesia.
42. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2015 Presented at: 3rd International Conference on Learning Representations; May 7-9, 2015; San Diego, USA.
43. Donahue J, Jia Y, Vinyals O, Hoffman J, Zhang N, Tzeng E, et al. DeCAF: a deep convolutional activation feature for generic visual recognition. 2014 Presented at: 31st International Conference on Machine Learning; June 21-26, 2014; Beijing, China.
44. Gatys L, Ecker A, Bethge M. Image style transfer using convolutional neural network. 2016 Presented at: IEEE Conference on Computer Vision and Pattern Recognition; June 27-30, 2016; Las Vegas, USA. [doi: 10.1109/cvpr.2016.265]
45. Zhang L, Wen T, Shi J. Deep image blending. 2020 Presented at: IEEE Winter Conference on Applications of Computer Vision; March 1-5, 2020; Snowmass Village, USA. [doi: 10.1109/wacv45572.2020.9093632]
46. Mittal A, Moorthy AK, Bovik AC. No-reference image quality assessment in the spatial domain. IEEE Trans Image Process 2012 Dec;21(12):4695-4708. [doi: 10.1109/TIP.2012.2214050] [Medline: 22910118]
47. Renieblas GP, Nogués AT, González AM, Gómez-Leon N, Del Castillo EG. Structural similarity index family for image quality assessment in radiological images. J Med Imaging (Bellingham) 2017 Jul;4(3):035501 [FREE Full text] [doi: 10.1117/1.JMI.4.3.035501] [Medline: 28924574]
48. Grimes PE, Sherrod Q. Structural and physiologic differences in the skin of darker racial ethnic groups structural. In: Aesthetics and Cosmetic Surgery for Darker Skin Types. Philadelphia: Wolters Kluwer, Lippincott Williams & Wilkins; 2008.
49. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. 2015 Presented at: IEEE Conference on Computer Vision and Pattern Recognition; June 7-12, 2015; Boston, USA. [doi: 10.1109/cvpr.2015.7298594]
50. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 Presented at: IEEE conference on computer vision and pattern recognition; June 27-30 2016; Las Vegas, USA. [doi: 10.1109/cvpr.2016.90]
51. Hogue L, Harvey VM. Basal cell carcinoma, squamous cell carcinoma, and cutaneous melanoma in skin of color patients. Dermatol Clin 2019 Oct;37(4):519-526. [doi: 10.1016/j.det.2019.05.009] [Medline: 31466591]
52. Baur C, Albarqouni S, Navab N. Generating highly realistic images of skin lesions with GANs. In: OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis. Cham: Springer; 2018:260-267.
53. Ghorbani A, Natarajan V, Coz D, Liu Y. DermGAN: Synthetic generation of clinical skin images with pathology. 2019 Presented at: NeurIPS workshop on Machine Learning for Health; December 13, 2019; Vancouver, Canada.

Abbreviations

AUC: area under the curve
ISIC: International Skin Imaging Collaboration
RGB: red, green, blue color space

Edited by A Mavragani; submitted 11.11.21; peer-reviewed by C Nelson, T Ndabu, E Krieghoff-Henning; comments to author 13.12.21; revised version received 23.12.21; accepted 04.01.22; published 08.03.22.

Please cite as:
Rezk E, Eltorki M, El-Dakhakhni W
Leveraging Artificial Intelligence to Improve the Diversity of Dermatological Skin Color Pathology: Protocol for an Algorithm Development and Validation Study
JMIR Res Protoc 2022;11(3):e34896
URL: https://www.researchprotocols.org/2022/3/e34896
doi:10.2196/34896
PMID:34983017

Protocol

The Essential Network (TEN): Protocol for an Implementation Study of a Digital-First Mental Health Solution for Australian Health Care Workers During COVID-19

Matthew James Coleshill[1,2], PhD; Peter Baldwin[1,2], PhD; Melissa Black[1,3], PhD; Jill Newby[1,3], PhD; Tanya Shrestha[1,2], BSc; Sam Haffar[1,2], BSc; Llewellyn Mills[4], PhD; Andrew Stensel[1,2], BSc; Nicole Cockayne[1,2], MPH; Jon Tennant[1,2], MEd; Samuel Harvey[1,2], MD; Helen Christensen[1,2], PhD

[1]Black Dog Institute, Sydney, Australia

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

[3]School of Psychology, University of New South Wales, Sydney, Australia

[4]Discipline of Addiction Medicine, Faculty Medicine and Health, University of Sydney, Sydney, Australia

Corresponding Author:
Matthew James Coleshill, PhD
Black Dog Institute
Hospital Road
Randwick, New South Wales
Sydney, 2031
Australia
Phone: 61 02 9065 9149
Email: m.coleshill@blackdog.org.au

Abstract

Background: The COVID-19 pandemic has placed health care workers (HCWs) under severe stress, compounded by barriers to seeking mental health support among HCWs. The Essential Network (TEN) is a blend of digital and person-to-person (blended care) mental health support services for HCWs, funded by the Australian Federal Department of Health as part of their national COVID-19 response strategy. TEN is designed as both a preventative measure and treatment for common mental health problems faced by HCWs. New blended services need to demonstrate improvements in mental health symptoms and test acceptability in their target audience, as well as review implementation strategies to improve engagement.

Objective: The primary objective of this implementation study is to design and test an implementation strategy to improve uptake of TEN. The secondary objectives are examining the acceptability of TEN among HCWs, changes in mental health outcomes associated with the use of TEN, and reductions in mental health stigma among HCWs following the use of TEN.

Methods: The implementation study contains 3 components: (1) a consultation study with up to 39 stakeholders or researchers with implementation experience to design an implementation strategy, (1) a longitudinal observational study of at least 105 HCWs to examine the acceptability of TEN and the effectiveness of TEN at 1 and 6 months in improving mental health (as assessed by the Distress Questionnaire [DQ-5], Patient Health Questionnaire [PHQ-9], Generalized Anxiety Disorder [GAD-7], Oldenburg Burnout Inventory [OBI-16], and Work and Social Adjustment Scale [WSAS]) and reducing mental health stigma (the Endorsed and Anticipated Stigma Inventory [EASI]), and (3) an implementation study where TEN service uptake analytics will be examined for 3 months before and after the introduction of the implementation strategy.

Results: The implementation strategy, designed with input from the consultation and observational studies, is expected to lead to an increased number of unique visits to the TEN website in the 3 months following the introduction of the implementation strategy. The observational study is expected to observe high service acceptability. Moderate improvements to general mental health (DQ-5, WSAS) and a reduction in workplace- and treatment-related mental health stigma (EASI) between the baseline and 1-month time points are expected.

Conclusions: TEN is a first-of-a-kind blended mental health service available to Australian HCWs. The results of this project have the potential to inform the implementation and development of blended care mental health services, as well as how such services can be effectively implemented during a crisis.

International Registered Report Identifier (IRRID): DERR1-10.2196/34601

(JMIR Res Protoc 2022;11(3):e34601) doi:10.2196/34601

KEYWORDS

blended care; mental health; burnout; health care workers; COVID-19; health care service; health service

Introduction

Background

The stress of the COVID-19 pandemic is placing health care workers (HCWs) at increased risk of poor mental health [1], with posttraumatic stress disorder (PTSD) being a major concern [2-5]. Mental health treatments can lower the risk of HCWs developing mental ill-health [1], yet few appropriate services are available at the necessary scale for all specializations of HCWs. The Australian health care workforce is currently estimated at 800,000 [6]. This figure will likely continue to grow alongside population growth, indicating that large-scale solutions are required. Early in the COVID-19 pandemic, researchers called for evidence-based, self-guided mental health services for HCWs, citing reluctance to seek help, coupled with the challenges of delivering quality support to thousands of time-poor consumers [7]. Blended care services that integrate digital (websites and apps) and person-to-person (including telehealth) services [8] can rapidly scale, while offering a personalized choice of evidence-based care options [9]. In this manner, blended care is a promising means of providing large-scale mental health services to HCWs during the ongoing COVID-19 pandemic.

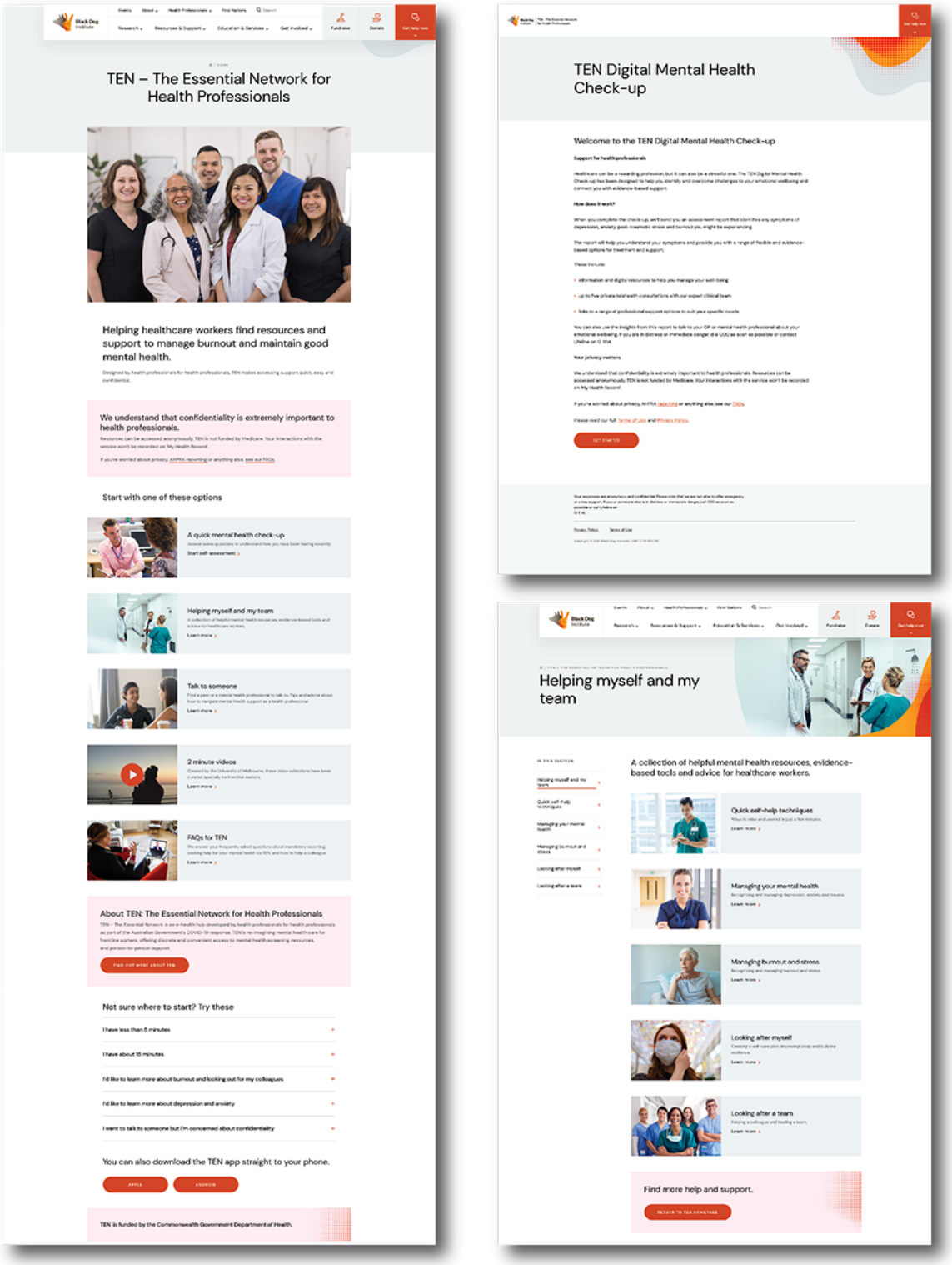
Over and above scalability, mental health services must be sensitive to the specialized needs of HCWs. As HCWs battle the pandemic, they face numerous sources of stress, such as fear of infection [10], unsupportive workplaces [11], and even watching colleagues die [12]. HCWs also vary in how they manage these stressors across time [13]. Therefore, HCWs need adaptable mental health services that can help with a range of concerns, from situational distress to moral injury [14] to PTSD [2,3]. Yet, even services tailored to HCWs have seen little uptake despite high reported demand [15]. One likely explanation is the lack of psychological safety many HCWs experience when accessing mental health care. Concerns about stigma [1,16], confidentiality, and discrimination from colleagues or employers [17,18] keep many HCWs from using the available services. Blended care can address some of these concerns through digital services (eg, anonymous chat-based therapies, accessing tailored resources without registration), while still providing the flexibility for HCWs who prefer person-to-person care to access these services.

To fully leverage the potential of blended care to support HCWs during the COVID-19 pandemic and beyond, services must be implemented in a way that ensures HCWs feel psychologically safe to use them. Thus far, little is known about implementing these services is a way that builds trust and increases uptake of available services. The few studies within the COVID era describe institution-specific implementation practices without data on the effectiveness of implementation strategies [19]. Studies implementing HCW-specific mental health services from previous pandemics are largely of poor quality and evaluate the effects of an intervention rather than implementation strategies to facilitate uptake and engagement [20]. These challenges exemplify implementation science in digital health, which is yet to produce a unified implementation framework to guide not only the reporting of strategies but also the evaluation of certain approaches [21,22]. Understanding how to effectively implement HCW-specific mental health services could boost the uptake of available care and inform the development of new services as the pandemic continues to unfold.

The Essential Network

The Essential Network (TEN) was funded by the Australian Government Department of Health as part of its national COVID-19 response strategy. TEN follows a sequential blended care approach [9] using an integrated service model that spans 4 care phases: (1) well-being promotion, (2) early detection and preventative interventions, (3) low-to-moderate-intensity services for HCWs experiencing preclinical distress, and (4) person-to-person clinical services for HCWs' mental health difficulties or clinical distress. In this sequential blended care approach, users engage with digital resources and treatment options, as required, and progress to person-to-person clinical services, where necessary. This selection of treatment options is made possible by a "digital ecosystem" that brought together a network of existing and new services from partner organizations with interest in the mental health of HCWs. Services are made available through a single web-based digital hub (Figure 1) accessible across desktop, laptop, and mobile web browsers. Digital services comprise an online mental health assessment, mental health resources and advice for assisting individuals and their colleagues, and links to other relevant resources and organizations (eg, psychiatrists who treat HCWs, peer support organizations). Person-to-person services comprise up to 5 individual clinical consultations with either a clinical psychologist or a psychiatrist delivered either via telehealth or in person at our clinic in Sydney, New South Wales (NSW), Australia. Users can select and explore the digital or person-to-person options that they believe best match their needs.

Figure 1. Example screenshots from the TEN website: home page (left), an online mental health assessment tool (top right), and evidence-based tools for practical self-help (bottom right). TEN: The Essential Network.



Implementation Framework

This study will examine the implementation of TEN, a blended mental health service for Australian HCWs established during COVID-19, funded by the Australian Federal Government. The Consolidated Framework for Implementation Research (CFIR) framework will be used to guide this examination as the constructs within the CFIR span both implementation and effectiveness within a health service context [23]. However, some adaptation of the CFIR will be required. The CFIR assumes that an intervention will be implemented in a single organization or industry, unlike TEN, which spans many organizational structures and professions. Similarly, the frameworks consolidated within the CFIR are largely aimed at implementing singular, intervention-based face-to-face services, rather than multichannel digital services like TEN. Previous

attempts to consolidate implementation strategies for digital mental health have focused on delivering specific digital programs as an alternative to traditional mental health services, rather than linking users to a range of care options in the way TEN does [21]. Work is underway to consolidate optimal digital health implementation strategies [22]; however, no such framework exists yet. For these reasons, in our study, the CFIR subconstruct definitions described by Damschroder et al [23] will be adapted to the digital and multiorganizational nature of TEN, guided by the digital-specific implementation strategies suggested by Graham et al [21] (see Table 1).

Study Objectives

The primary objective of this study is to work with HCW researchers, industry partners, and active TEN users to design and test an implementation strategy for TEN. The secondary objectives are to examine the acceptability of TEN, to measure any changes in psychosocial well-being associated with the use of TEN, and to determine whether TEN reduces mental health stigma among HCWs.

Table 1. CFIR[a] constructs and example interview questions.

CFIR subconstruct	Example interview question
Intervention characteristics	
Intervention source	N/A[b]. TEN[c] was externally developed.
Evidence strength and quality	Interviews: Questions regarding perceptions of the quality of the TEN website/offering and beliefs that the care options offered will benefit the mental health of HCWs[d].
Relative advantage	Interviews: Questions regarding awareness of TEN equivalents and perceived advantages of TEN.
Adaptability	Interviews: Questions regarding the capacity for TEN to be adapted to the needs of professional[e] groups.
Trialability	Interviews: Questions regarding the feasibility of implementation and the evaluation of TEN within a single organization or professional groups[e].
Complexity	Interviews: Questions regarding difficulty of implementing TEN.
Design quality and packaging	Interviews: Questions regarding perception of the quality of TEN presentation to organizations and users[e].
Cost	Interviews: Questions regarding resourcing (financial, personnel, etc) and opportunity costs of using TEN.
Outer setting	
Patient needs and resources	Interviews: Questions regarding knowledge of member/employee needs[e] and how these align with the organizational strategy.
Cosmopolitanism	Interviews: Questions regarding organizational networks.
Peer pressure	Interviews: Questions regarding knowledge of related organizational efforts and how these might affect the organizational endorsement of TEN.
External policies and incentives	Interviews: Questions regarding external policy considerations and how these might affect the organizational endorsement of TEN.
Inner setting	
Structural characteristics	Interviews: Questions regarding the social architecture, age, maturity, and size of an organization and profession[e].
Networks and communications	Interviews: Questions regarding how organizations communicate with members or employees (generally and in relation to mental health).
Culture	Interviews: Questions regarding norms, values, and basic assumptions of an organization and profession[e].
Implementation climate[f]	Interviews: Questions regarding current endorsement of TEN and whether TEN competes with other initiates for resources or strategic importance.
Readiness for implementation[f]	Interviews: Questions regarding readiness of an organization to assist with the implementation strategy[e].
Characteristics of individuals	
Knowledge and beliefs about the intervention	Interviews: Questions regarding the value and knowledge of both TEN and digital/blended mental health[e].
Self-efficacy	Interviews: Questions regarding the capabilities of both stakeholders and members/employees[e] to execute implementation strategies.
Individual stage of change	Interviews: Questions regarding the stage of change most members/employees[e] are likely in with respect to using TEN or blended mental health services[e].
Individual identification with organization	Interviews: Questions regarding how stakeholders perceive the relationship between their organization and members/employees[e].
Other personal attributes	Interviews: Questions regarding what personal traits, such as tolerance of ambiguity, intellectual ability, motivation, values, competence, capacity, and learning styles, of their members/employees[e] are likely to affect the implementation of TEN.

[a]CFIR: Consolidated Framework for Implementation Research.

[b]N/A: not applicable.

[c]TEN: The Essential Network.

[d]HCW: health care worker.

[e]Language adapted to the multiorganizational context of TEN.

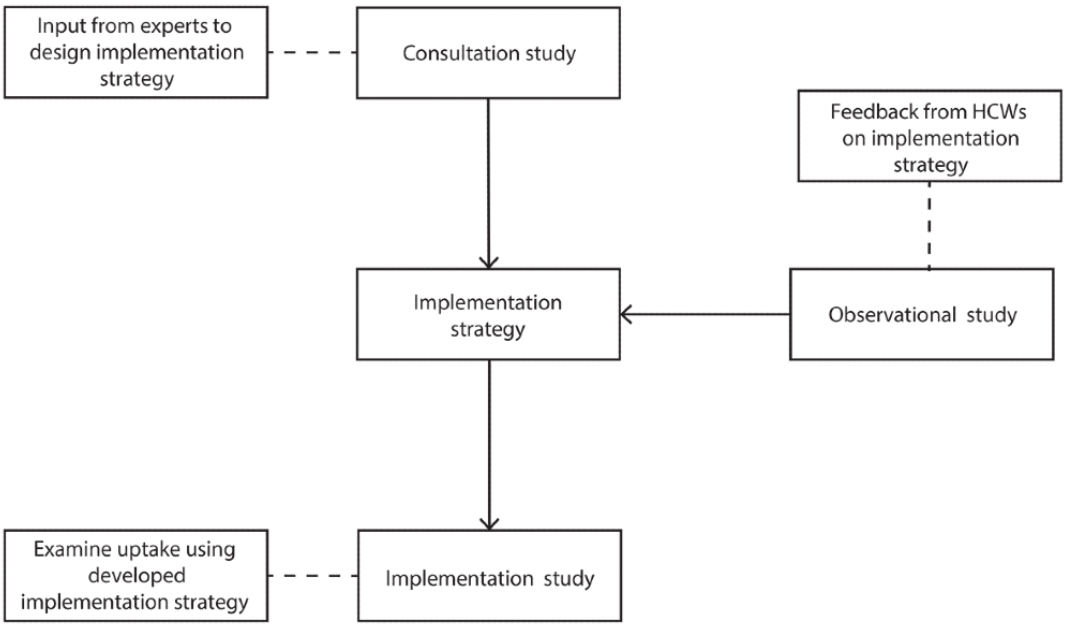
[f]Adaptation of original CFIR definitions to TEN.

Methods

Study Design

This study will follow an implementation-effectiveness hybrid design [24]. The protocol has 3 components (Figure 2). The *consultation study* component (Study 1) is a series of semistructured qualitative interviews conducted with stakeholders to explore the needs of Australian HCWs and develop an implementation strategy for TEN. The *observational study* component (Study 2) is a pre-/postobservational study with Australian HCWs evaluating the acceptability TEN, providing feedback on the implementation strategy, and examining any changes in mental health outcomes that are associated with use of the service. The *implementation study* (Study 3) component will be a pre-/postaudit of TEN users both before and after the implementation strategy is implemented [25].

Figure 2. Flowchart depicting the design and components of the implementation trial. HCW: health care worker.



Study Setting

There is no specific study site. Semistructured interviews carried out as part of the consultations will be conducted online via videocalls (eg, Zoom, Microsoft Teams). As a digital mental health hub, participants in the observational study will be able to participate in the study and access the TEN service from any location in Australia. If applicable, face-to-face consultations through the TEN Clinic may be conducted at the Black Dog Institute, Randwick, NSW 2031, or via telehealth.

Study 1: Consultation Study

Ethics

The consultation component received approval from the UNSW Human Research Ethics Panel (HC no. 3500) on June 7, 2021.

Participants

Participants in the consultation study will be representatives from Australian professional organizations (n=15-18) relevant to mental health or HCWs as well as researchers or staff from the Black Dog Institute (n=10-21) working on projects relevant to HCWs' mental health.

Recruitment

Participants for the consultation component will be approached directly to discuss participation or emailed an invitation to the study and provided with an online participant information statement and consent form.

Data Collection

Qualitative data will be captured in semistructured interviews with HCW researchers and industry partners who represent the interests and views of their respective disciplines. Interview questions will be structured to elicit information directly relevant to the key constructs constituting the CFIR and derived from the subconstruct descriptions provided by Damschroder et al [23]. Where appropriate, questions regarding some CFIR subconstructs will be adapted to a multiorganizational or digital context.

Study 2: Observational Study

Ethics

The observational component received approval from the UNSW Human Research Ethics Committee (HC no. 210394) on July 22, 2021.

Participants

Participants in the observational study will be self-identified Australian HCWs who currently reside in Australia and have sufficient English proficiency to participate. Participant status as HCWs will be verified using the Australian Health Practitioner Regulation Agency's register of practitioners. Participants who cannot be verified using the register of practitioners will be contacted on a case-by-case basis to confirm eligibility.

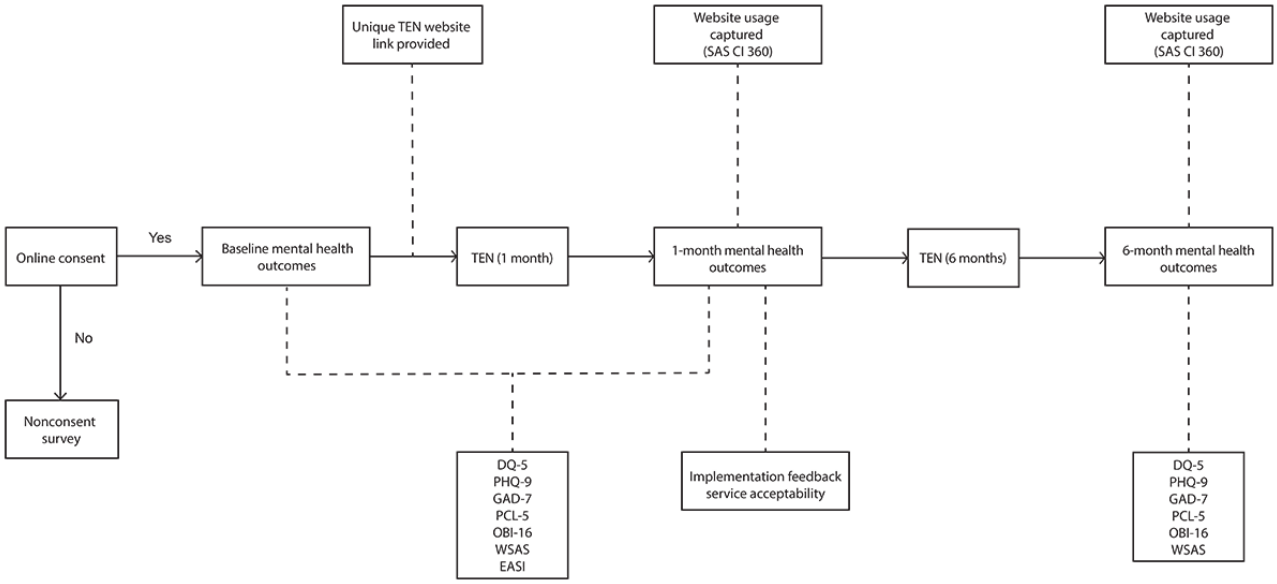
Recruitment

Participants for the observational study will be recruited through advertisements on social media, emails sent to health professionals on an internal Black Dog Institute mailing list, emails from TEN partner organizations, and information about the study on the Black Dog Institute and TEN websites. Prospective participants who engage with advertisements or study details on the TEN website will be redirected to the online participant information statement and consent form. Participants who elect not to participate in the study will be provided with a link to TEN and informed they may use the service freely and anonymously.

Data Collection

Quantitative data will be collected through an observational study with Australian HCWs (Figure 3). Participants will provide online consent using the e-Consent feature of REDCap (Vanderbilt University). People who decline to participate will be provided with a link to TEN and asked to complete a brief survey asking why they opted not to participate.

Figure 3. Flowchart depicting a progression through the observational study. DQ-5: Distress Questionnaire; EASI: Endorsed and Anticipated Stigma Inventory; GAD-7: Generalized Anxiety Disorder; OBI-16: Oldenburg Burnout Inventory; PCL-5: Posttraumatic Stress Disorder Checklist; PHQ-9: Patient Health Questionnaire; SAS CI 360: SAS Customer Intelligence 360; TEN: The Essential Network; WSAS: Work and Social Adjustment Scale.



After consenting, participants will provide information on demographics and mental health (Table 2), service acceptability (Table 3), and complete baseline mental health and stigma outcomes (Table 4). Mental health outcomes are identical to the TEN online assessment. To avoid the participants needing to repeat the online assessment on the TEN website after completing the baseline mental health outcomes, the survey baseline will provide the same assessment feedback. Participants will then be provided with a link to the TEN website and asked to engage with the service naturalistically for 6 months. The link provided to each participant is unique and associated with their consent record number. User engagement (eg, pages viewed) with the TEN website is recorded automatically through the SAS Customer Intelligence 360 (SAS CI 360) platform. SAS CI 360 is a commercial website analytics package that records website use. Using unique TEN website links sent to participants, their TEN website use will be extracted and linked to their other study data. Participants will be recommended to bookmark this link, and their unique link will be provided to them by email at baseline and after the 1-month follow-up.

Table 2. Questions examining demographic and baseline mental health.

Baseline question	Time point	Classification	Type of question
Gender	Baseline	Demographics	Single selection
Age	Baseline	Demographics	Single selection
Where are you located?	Baseline	Demographics	Single selection
What best describes your profession?	Baseline	Demographics	Single selection
Have you previously used the TEN[a] service?	Baseline	Mental health	Yes/no
COVID-19 has affected your mental health.	Baseline	Mental health	Likert scale
Have you ever seen a doctor, counsellor, or other health professional about your mental health?	Baseline	Mental health	Yes/no
If you were seeking help for a mental health problem, what type of service would you most likely choose?	Baseline	Mental health	Multiple selection
Which of the following were the most important reasons why you chose to participate in this study?	Baseline	Mental health	Multiple selection
Which parts of TEN are you most interested in?	Baseline	Mental health	Multiple selection
What is your current employment status as an HCW[b]?	6 months	Employment	Single selection

[a]TEN: The Essential Network.

[b]HCW: health care worker.

Table 3. Questions examining TEN[a] service acceptability.

Service acceptability question	Time point	Type of question
If you were seeking help for a mental health problem, what type of service would you most likely choose?	Baseline	Multiple selection
Which of the following were the most important reasons why you chose to participate in this study?	Baseline	Multiple selection
Which parts of TEN are you most interested in?	Baseline	Multiple selection
If you needed to, would you use the TEN Clinic if there was a service charge?	1 month	Yes/no
If you needed to, would you use the TEN Clinic if it was provided under Medicare?	1 month	Yes/no
TEN met my mental health and well-being needs.	1 month	Likert scale
TEN improved my awareness of my mental health.	1 month	Likert scale
TEN improved my psychological coping skills.	1 month	Likert scale
A mental health service I found through TEN improved my psychological coping skills.	1 month	Likert scale
I would recommend TEN to a colleague.	1 month	Likert scale
Which parts of TEN were most useful to you?	1 month	Multiple selection
Why did you find these parts of TEN useful?	1 month	Free text
Did TEN make it easier for you to find the mental health support you wanted?	1 month	Yes/no
Did TEN help you learn about new sources of support?	1 month	Yes/no
Did you have any other feedback about TEN?	1 month	Free text

[a]TEN: The Essential Network.

Table 4. Mental health and psychosocial instruments.

Questionnaire	Outcome
DQ-5[a]	Psychological distress
PHQ-9[b]	Depression
GAD-7[c]	Anxiety
PCL-5[d]	PTSD[e]
OBI-16[f]	Burnout
WSAS[g]	Social- and work-related impairment
EASI[h]	Mental health stigma

[a]DQ-5: Distress Questionnaire.
[b]PHQ-9: Patient Health Questionnaire.
[c]GAD-7: Generalized Anxiety Disorder.
[d]PCL-5: Posttraumatic Stress Disorder Checklist.
[e]PTSD: posttraumatic stress disorder.
[f]OBI-16: Oldenburg Burnout Inventory.
[g]WSAS: Work and Social Adjustment Scale.
[h]EASI: Endorsed and Anticipated Stigma Inventory.

After using TEN for 1 month, participants will be provided with a survey containing service acceptability (Table 3), mental health and stigma outcomes (Table 4), feedback on the implementation strategy (Table 5), and self-reported service engagement (Table 5). After completing the survey, participants will be informed that their TEN website usage will be monitored for a further 5 months to examine persistence with the service. After 6 months, participants will be provided with a follow-up survey about employment (Table 2) and mental health outcomes (Table 4).

Table 5. Questions examining the implementation strategy and service engagement.

Question	Type of question
Implementation feedback	
Would you agree that this would be an effective strategy for engaging most HCWs[a] with TEN[b]?	Likert scale
Would you agree that this would be an effective strategy for engaging your specific health care profession with TEN?	Likert scale
Do you have any thoughts about this strategy or suggestions for other strategies to engage more HCWs with TEN?	Free text
Service engagement	
How much did you use the TEN service over the month?	Likert scale
Roughly how many times would you say you accessed TEN over the month?	Numerical

[a]HCW: health care worker.
[b]TEN: The Essential Network.

TEN Service Acceptability

Service acceptability will be examined through survey questions provided to participants at both the 0- and 1-month time points. The questions will examine whether TEN met the participants' mental health needs and which TEN services they found most useful (Table 3). People who decline to participate in the study will be prompted with an optional survey with a single question asking them why they opted not to participate in the study.

Mental Health and Psychosocial Outcomes

Mental health and psychosocial outcomes will be examined through surveys provided to participants at the 0-, 1-, and 6-month time points. The questionnaires used will examine a range of outcomes, including the Distress Questionnaire (DQ-5) [25], the Patient Health Questionnaire (PHQ-9) [26], Generalized Anxiety Disorder (GAD-7) [27], the Posttraumatic Stress Disorder Checklist (PCL-5) [28], the Oldenburg Burnout Inventory (OBI-16) [29], the Work and Social Adjustment Scale (WSAS) [30], and the Endorsed and Anticipated Stigma Inventory (EASI) [31] (Table 4). To address workplace stigma unique to HCWs, a question addressing mandatory reporting concerns was added to the workplace stigma subscale of the EASI. Except for the EASI, these questionnaires are identical to those used in the TEN online assessments. Completed online assessments are automatically collected and stored on University of New South Wales (UNSW) servers. Any completed online

assessments will be linked back to the participants using their internet protocol (IP) address collected during online consent.

TEN Implementation Feedback

Feedback on the strategy will be examined through survey questions provided to observational study participants at the 1-month time point. Participants will first be presented with a brief overview of the implementation strategy. The questions after this overview will examine whether the implementation strategy is an effective strategy for engaging HCWs, as well as open-ended feedback on the implementation strategy (Table 5).

TEN Service Engagement

Engagement with TEN will be examined through survey questions provided to participants at the 1-month time point, as well as through website and service analytics. The questions will examine self-reported usage of TEN (Table 5). Website user analytics are automatically collected through the SAS CI 360 platform whenever a user interacts with the TEN website. These user analytics include the IP address of the user as well as which pages or resources were accessed and when. User analytics will be linked to the participant using the IP address provided during consent. TEN Clinic clinicians will also provide data from person-to-person clinical services, including whether a participant accessed the TEN Clinic and the number of sessions.

Sample Size

Sample size calculations were based on the primary analysis: a random intercepts mixed effects model estimating change in WSAS scores from baseline to 1 month in a single group of participants. Estimated scores at baseline were derived from 249 TEN mental health assessments performed between June 28 and August 28, 2021. These scores represent HCWs self-selecting to engage with the TEN service and, as such, are analogous to the observational study participants at baseline. The mean of scores in this group was 14.9 (SD 8.9). The cut-off score for a clinical case on the WSAS is 11; thus, to reduce from the mean untreated score of 14.9 to a subclinical score requires a 3.87-point reduction. With variance-covariance of random effects derived from the SD in WSAS scores from current TEN users of 8.9, and accounting for 20% attrition, a sample size of at least 105 is required to detect a true reduction of this size with 80% power and a significance level of α=.05.

Study 3: Implementation Study

Ethics

Ethics approval for the implementation study will be sought from the UNSW Human Research Ethics Committee following design of the implementation strategy.

Participants

Participants in the implementation study will be self-identified Australian HCWs.

Data Collection

Quantitative data will be collected through an audit of usage of the TEN website and TEN Clinic for a period of 3 months both before and after the implementation of the implementation strategy. Although the exact implementation strategy will be determined by the results of the consultation study, example strategies from the literature include identifying “champions” within an organization to drive uptake or develop incentives, such as continuing professional development activities [32]. TEN website data include completed online assessments automatically recorded on UNSW servers and website analytics (eg, number of unique users, pages viewed) automatically collected through the SAS CI 360 platform. TEN Clinic data include service analytics data, which are routinely captured by clinicians in the delivery of TEN Clinic services. Surface-level nonidentifiable TEN Clinic data will be examined as part of the implementation study (eg, number of referrals, number of consultations).

Results

Study 1: Consultation Study

Data Analysis

A framework analysis using the CFIR will be conducted to thematically analyze and interpret the semistructured interviews from the consultations. Each transcript will be analyzed separately by 2 investigators who will meet periodically throughout data collection to discuss emerging themes and resolve discrepancies. The sample size is expected to be sufficient for thematic saturation [33].

Study 2: Observational Study

Data Analysis

Descriptive statistics will be used to report the overall acceptability of TEN. The results from the observational study will be analyzed using linear mixed models to examine changes in reported psychosocial benefits over time (ie, baseline, 1 month). Patterns of service engagement and baseline outcome severity will be classified using cluster analyses or latent class analyses. These patterns will be used to examine how service engagement influences outcomes, as well as stratify participants baseline outcomes for subset analysis. Kaplan-Meier curves with TEN engagement data will be used to examine persistence with TEN over the course of the 6-month follow-up (ie, when participants ceased accessing the service). Missing data will be addressed using maximum likelihood estimation. This will answer the research questions by examining the acceptability of TEN among HCWs, patterns of service usage over time, any benefits observed after engaging with the TEN service, and persistence with the service over time.

Study 3: Implementation Study

Data Analysis

Descriptive statistics will be used to report usage of the TEN website and TEN Clinic over the 3-month period both pre- and postimplementation strategy. The *t* test will be used to compare overall usage of the TEN service for these 2 periods. Linear mixed models will examine latent growth in service usage following the implementation strategy.

Discussion

Summary

Blended digital and person-to-person care has significant potential to support mental health of HCWs through service flexibility and widespread availability. To date, however, the potential of blended care has been hampered by a lack of knowledge around the optimal implementation of such services. The CFIR framework provides a structured approach to implementation that will guide the development and application of implementation strategies for TEN. Using a implementation-effectiveness hybrid study, this implementation study will span 3 separate components. Through a series of consultations, an implementation strategy tailored to HCWs will be developed. Feedback on this strategy will be provided by HCWs participating in an observational study of TEN, from which data on the effectiveness of TEN will be acquired. Finally, the implementation strategy will be implemented and compared against the previous engagement strategies. The knowledge created by this project will inform the development and delivery of blended mental health services for HCWs, including how such services can be rapidly launched and implemented during times of crisis.

Potential Limitations

Although the implementation strategy has not yet been designed, the wide availability of TEN makes it difficult to examine site- or specialization-specific implementation strategies or service upgrades. Although the goal of TEN is to provide a blended care service for all Australian HCWs, this lack of specificity may hamper implementation within specific HCW specializations. The future of blended health is platforms that can be easily customized to different contexts, allowing for tailored implementation strategies and service upgrades. Similarly, the broad aims of the implementation of TEN may not be easily generalizable to other, more specific, contexts.

This implementation study and observational study will also be carried out during the ongoing COVID-19 pandemic. TEN itself was designed to support the mental health of HCWs during the pandemic. As such, changes in the COVID-19 context in Australia, such as outbreaks, restrictions, and the vaccination rollout, are likely to affect both uptake of the service and mental health outcomes. Such wider events will be considered when interpreting the results of both the implementation strategy and the observational study.

The observational study is also limited by a lack of an appropriate control arm and use of randomization. Common mental health outcomes, such as distress, anxiety, and depression, have the potential to improve over time without intervention. Although subgroup analyses examining usage of the service may shed some light on improvement without the use of the TEN service, the fact remains that a randomized controlled trial design with a suitable control would provide more robust data on efficacy.

Due to common concerns around mandatory reporting among HCWs [34], the TEN website was designed to be accessed anonymously (ie, without registration). Although this is a strength of the service itself, it poses problems for measuring participant engagement with the TEN website during a prospective study. The method used in the observational study—providing users with a unique website link—is error prone. Although participants are asked to bookmark the link and will be provided with the link via email, it remains that participants may simply access the TEN website without using their unique link. During data analysis, participant website analytics will be compared against self-reported usage to understand the extent of such errors.

Conclusion

By using a blended care approach, TEN can overcome common barriers to HCWs engaging with mental health services—primarily by allowing HCWs to anonymously access digital resources. Nonetheless, it remains that many HCWs will still choose not to engage with person-to-person services due to concerns around mandatory reporting. Further, although the lack of registration to access TEN ameliorates concerns around digital privacy, it requires computer skills that some HCWs may be lacking. Although these are issues beyond the scope of this evaluation, they need to be considered and addressed when designing and implementing mental health services for HCWs.

Acknowledgments

This study is supported by funding from the Australian Government Department of Health through The Prevention Hub. HC is supported by the National Health and Medical Research Council (NHMRC) Fellowship 1155614.

Authors' Contributions

MJC and PB were responsible for initial drafting of the manuscript. MB, JN, TS, S Haffar, LM, AS, NC, JT, S Harvey, and HC provided input on subsequent drafts of the manuscript. MJC, PB, and MB were responsible for the design of the implementation study. PB, MB, JN, S Harvey, and HC were responsible for the development and delivery of The Essential Network (TEN).

Conflicts of Interest

PB and MB are supported by funding for The Essential Network (TEN). MJC, PB, MB, JN, TS, S Harvey, AS, NC, JT, S Haffar, and HC are employed by the Black Dog Institute, which provides clinical services to various populations, including health professionals.

References

1. Kisely S, Warren N, McMahon L, Dalais C, Henry I, Siskind D. Occurrence, prevention, and management of the psychological effects of emerging virus outbreaks on healthcare workers: rapid review and meta-analysis. BMJ 2020 May 05;369:m1642 [FREE Full text] [doi: 10.1136/bmj.m1642] [Medline: 32371466]
2. Carmassi C, Foghi C, Dell'Oste V, Cordone A, Bertelloni CA, Bui E, et al. PTSD symptoms in healthcare workers facing the three coronavirus outbreaks: what can we expect after the COVID-19 pandemic. Psychiatry Res 2020 Oct;292:113312 [FREE Full text] [doi: 10.1016/j.psychres.2020.113312] [Medline: 32717711]
3. Dutheil F, Mondillon L, Navel V. PTSD as the second tsunami of the SARS-Cov-2 pandemic. Psychol Med 2021 Jul;51(10):1773-1774 [FREE Full text] [doi: 10.1017/S0033291720001336] [Medline: 32326997]
4. Sahebi A, Nejati-Zarnaqi B, Moayedi S, Yousefi K, Torres M, Golitaleb M. The prevalence of anxiety and depression among healthcare workers during the COVID-19 pandemic: an umbrella review of meta-analyses. Prog Neuropsychopharmacol Biol Psychiatry 2021 Apr 20;107:110247 [FREE Full text] [doi: 10.1016/j.pnpbp.2021.110247] [Medline: 33476692]
5. Young KP, Kolcz DL, O'Sullivan DM, Ferrand J, Fried J, Robinson K. Health care workers' mental health and quality of life during COVID-19: results from a mid-pandemic, national survey. Psychiatr Serv 2021 Feb 01;72(2):122-128. [doi: 10.1176/appi.ps.202000424] [Medline: 33267652]
6. Ahpra and National Boards. Statistics. URL: https://www.ahpra.gov.au/About-Ahpra/What-We-Do/Statistics.aspx [accessed 2022-02-28]
7. Yang L, Yin J, Wang D, Rahman A, Li X. Urgent need to develop evidence-based self-help interventions for mental health of healthcare workers in COVID-19 pandemic. Psychol Med 2021 Jul;51(10):1775-1776 [FREE Full text] [doi: 10.1017/S0033291720001385] [Medline: 32340642]
8. Wentzel J, van der Vaart R, Bohlmeijer ET, van Gemert-Pijnen JEWC. online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016 Feb 09;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]
9. Erbe D, Eichert H, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Sep 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
10. Choudhury T, Debski M, Wiper A, Abdelrahman A, Wild S, Chalil S, et al. COVID-19 pandemic: looking after the mental health of our healthcare workers. J Occup Environ Med 2020 Jul;62(7):e373-e376. [doi: 10.1097/JOM.0000000000001907] [Medline: 32730043]
11. Evanoff BA, Strickland JR, Dale AM, Hayibor L, Page E, Duncan JG, et al. Work-related and personal factors associated with mental well-being during the COVID-19 response: survey of health care and other workers. J Med Internet Res 2020 Aug 25;22(8):e21366 [FREE Full text] [doi: 10.2196/21366] [Medline: 32763891]
12. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020 Apr 07;323(13):1239-1242. [doi: 10.1001/jama.2020.2648] [Medline: 32091533]
13. Walton M, Murray E, Christian MD. Mental health care for medical staff and affiliated healthcare workers during the COVID-19 pandemic. Eur Heart J Acute Cardiovasc Care 2020 Apr;9(3):241-247 [FREE Full text] [doi: 10.1177/2048872620922795] [Medline: 32342698]
14. Griffin BJ, Purcell N, Burkman K, Litz BT, Bryan CJ, Schmitz M, et al. Moral injury: an integrative review. J Trauma Stress 2019 Jun;32(3):350-362. [doi: 10.1002/jts.22362] [Medline: 30688367]
15. Chen Q, Liang M, Li Y, Guo J, Fei D, Wang L, et al. Mental health care for medical staff in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):e15-e16 [FREE Full text] [doi: 10.1016/S2215-0366(20)30078-X] [Medline: 32085839]
16. Knaak S, Mantler E, Szeto A. Mental illness-related stigma in healthcare: barriers to access and care and evidence-based solutions. Healthc Manage Forum 2017 Mar;30(2):111-116 [FREE Full text] [doi: 10.1177/0840470416679413] [Medline: 28929889]
17. Moll S, Eakin JM, Franche R, Strike C. When health care workers experience mental ill health: institutional practices of silence. Qual Health Res 2013 Feb;23(2):167-179. [doi: 10.1177/1049732312466296] [Medline: 23132130]
18. Beran RG. Mandatory notification of impaired doctors. Intern Med J 2014 Dec;44(12a):1161-1165. [doi: 10.1111/imj.12604] [Medline: 25442756]
19. Miotto K, Sanford J, Brymer MJ, Bursch B, Pynoos RS. Implementing an emotional support and mental health response plan for healthcare workers during the COVID-19 pandemic. Psychol Trauma 2020 Aug;12(S1):S165-S167. [doi: 10.1037/tra0000918] [Medline: 32525378]
20. Zaçe D, Hoxhaj I, Orfino A, Viteritti AM, Janiri L, Di Pietro ML. Interventions to address mental health issues in healthcare workers during infectious disease outbreaks: a systematic review. J Psychiatr Res 2021 Apr;136:319-333 [FREE Full text] [doi: 10.1016/j.jpsychires.2021.02.019] [Medline: 33636688]

21. Graham AK, Lattie EG, Powell BJ, Lyon AR, Smith JD, Schueller SM, et al. Implementation strategies for digital mental health interventions in health care settings. Am Psychol 2020 Nov;75(8):1080-1092 [FREE Full text] [doi: 10.1037/amp0000686] [Medline: 33252946]
22. Soobiah C, Cooper M, Kishimoto V, Bhatia RS, Scott T, Maloney S, et al. Identifying optimal frameworks to implement or evaluate digital health interventions: a scoping review protocol. BMJ Open 2020 Aug 13;10(8):e037643 [FREE Full text] [doi: 10.1136/bmjopen-2020-037643] [Medline: 32792444]
23. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
24. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
25. Batterham PJ, Sunderland M, Carragher N, Calear AL, Mackinnon AJ, Slade T. The Distress Questionnaire-5: population screener for psychological distress was more accurate than the K6/K10. J Clin Epidemiol 2016 Mar;71:35-42. [doi: 10.1016/j.jclinepi.2015.10.005] [Medline: 26464194]
26. 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]
27. Spitzer RL, Kroenke K, Williams JBW, 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]
28. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. J Trauma Stress 2015 Dec;28(6):489-498. [doi: 10.1002/jts.22059] [Medline: 26606250]
29. Halbesleben JRB, Demerouti E. The construct validity of an alternative measure of burnout: investigating the English translation of the Oldenburg Burnout Inventory. Work Stress 2005 Jul;19(3):208-220. [doi: 10.1080/02678370500340728]
30. Mundt JC, Marks IM, Shear MK, Greist JH. 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]
31. Vogt D, Di Leone BAL, Wang JM, Sayer NA, Pineles SL, Litz BT. Endorsed and Anticipated Stigma Inventory (EASI): a tool for assessing beliefs about mental illness and mental health treatment among military personnel and veterans. Psychol Serv 2014 Feb;11(1):105-113. [doi: 10.1037/a0032780] [Medline: 24274110]
32. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015 Feb 12;10(1):21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
33. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020;15(5):e0232076 [FREE Full text] [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]
34. Edwards JL, Crisp DA. Seeking help for psychological distress: barriers for mental health professionals. Aust J Psychol 2020 Nov 20;69(3):218-225. [doi: 10.1111/ajpy.12146]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
DQ-5: Distress Questionnaire
EASI: Endorsed and Anticipated Stigma Inventory
GAD-7: Generalized Anxiety Disorder
HCW: health care worker
IP: internet protocol
OBI-16: Oldenburg Burnout Inventory
PCL-5: Posttraumatic Stress Disorder Checklist
PHQ-9: Patient Health Questionnaire
PTSD: posttraumatic stress disorder
SAS CI 360: SAS Customer Intelligence 360
TEN: The Essential Network
UNSW: University of New South Wales
WSAS: Work and Social Adjustment Scale

Edited by T Leung; submitted 01.11.21; peer-reviewed by D Erbe, B Rollman; comments to author 20.12.21; revised version received 26.01.22; accepted 27.01.22; published 09.03.22.

Please cite as:
Coleshill MJ, Baldwin P, Black M, Newby J, Shrestha T, Haffar S, Mills L, Stensel A, Cockayne N, Tennant J, Harvey S, Christensen H
The Essential Network (TEN): Protocol for an Implementation Study of a Digital-First Mental Health Solution for Australian Health Care Workers During COVID-19
JMIR Res Protoc 2022;11(3):e34601
URL: https://www.researchprotocols.org/2022/3/e34601
doi:10.2196/34601
PMID:35148269

Protocol

Impact of Biological and Lifestyle Factors on Cognitive Aging and Work Ability in the Dortmund Vital Study: Protocol of an Interdisciplinary, Cross-sectional, and Longitudinal Study

Patrick D Gajewski[1*], PhD; Stephan Getzmann[1*], PhD; Peter Bröde[1], PhD; Michael Burke[1], PhD; Cristina Cadenas[1], PhD; Silvia Capellino[1], PhD; Maren Claus[1], PhD; Erhan Genç[1], PhD; Klaus Golka[1], PhD; Jan G Hengstler[1], PhD; Thomas Kleinsorge[1], PhD; Rosemarie Marchan[1], PhD; Michael A Nitsche[1], PhD; Jörg Reinders[1], PhD; Christoph van Thriel[1], PhD; Carsten Watzl[1], PhD; Edmund Wascher[1], PhD

Leibniz Research Centre for Working Environment and Human Factors (IfADo) at the Technical University of Dortmund, Dortmund, Germany
[*]these authors contributed equally

Corresponding Author:
Patrick D Gajewski, PhD
Leibniz Research Centre for Working Environment and Human Factors (IfADo) at the Technical University of Dortmund
Ardeystr. 67
Dortmund, 44139
Germany
Phone: 49 231 1084 383
Fax: 49 231 1084 401
Email: gajewski@ifado.de

Abstract

Background: Previous research revealed several biological and environmental factors modulating cognitive functioning over a human's lifespan. However, the relationships and interactions between biological factors (eg, genetic polymorphisms, immunological parameters, metabolic products, or infectious diseases) and environmental factors (eg, lifestyle, physical activity, nutrition, and work type or stress at work) as well as their impact on cognitive functions across the lifespan are still poorly understood with respect to their complexity.

Objective: The goal of the Dortmund Vital Study is to validate previous hypotheses as well as generate and validate new hypotheses about the relationships among aging, working conditions, genetic makeup, stress, metabolic functions, the cardiovascular system, the immune system, and mental performance over the human lifespan with a focus on healthy working adults. The Dortmund Vital Study is a multidisciplinary study involving the Departments of Ergonomics, Immunology, Psychology and Neurosciences, and Toxicology at the Leibniz Research Centre for Working Environment and Human Factors at the Technical University of Dortmund (IfADo) in Germany, as well as several national and international partners.

Methods: The Dortmund Vital Study is designed as a combined cross-sectional and longitudinal study. Approximately 600 healthy subjects aged between 20 and 70 years will participate. A wide range of demographic, psychological, behavioral, sensory, cardiovascular, immunological, and biochemical data, a comprehensive electroencephalography (EEG)-based cognitive test battery as well as structural and functional magnetic resonance imaging (MRI) have been included in the study.

Results: The study was approved by the Ethics Committee of IfADo in October 2015. The baseline testing was conducted between 2016 and 2021 and will be repeated every 5 years (3 follow-up measures until 2035). As of March 2020 (until the outbreak of the COVID-19 pandemic), 593 participants have been enrolled. Some results from the cross-sectional part of the study were already published, further results will be published soon. Longitudinal data will be analyzed and published by 2025.

Conclusions: We anticipate that the study will shed light on sources of interindividual differences in the alterations of cognitive functioning with increasing age and reveal biological and lifestyle markers contributing to work ability, longevity, and healthy aging on the one hand, and to risk factors for cognitive decline, mild cognitive impairment, or even dementia on the other hand.

Trial Registration: ClinicalTrials.gov NCT05155397; https://clinicaltrials.gov/ct2/show/NCT05155397

International Registered Report Identifier (IRRID): DERR1-10.2196/32352

(JMIR Res Protoc 2022;11(3):e32352) doi:10.2196/32352

KEYWORDS

cognitive aging; genetic polymorphisms; immunology; metabolism; latent infections; stress; occupational health; cardiovascular system; neuropsychology; magnetic resonance imaging; electroencephalography; lifespan; lifestyle; longitudinal study; biomarkers; longevity; aging.

Introduction

Background

Cognitive performance in humans is a complex phenomenon that is influenced by numerous variables. Among them are age [1,2], various infections [3,4], inflammatory processes often due to the release of cytokines from immune cells [5,6], metabolic alterations such as hyperammonemia [7,8], lifestyle factors including physical and mental activity, type of work, nutrition and stress [9-16], as well as numerous genetic variants [17,18] that can be analyzed by polygenic scores [19]. Most studies in this field of research have focused on specific individual factors of influence and did not consider the development of individuals over a significant part of their lifetime. To bridge this gap, we initiated the Dortmund Vital Study that assesses cognitive performance in a longitudinal study design using a complex multiscale approach to understand the influence of the most important lifestyle, occupational, and biological factors on cognitive performance.

Research Questions and Aims

The aim of the Dortmund Vital Study is to evaluate the effect of several endogenous and exogenous parameters and their complex interaction on behavior and the underlying neuronal activity. This will be achieved by implementing cross-sectional and longitudinal designs. By repeated measurements of the endogenous and exogenous parameters over an extended period, risk factors for possible cognitive impairment associated with depression, occupational burnout, or age-related diseases, such as mild cognitive impairment or dementia, and nonpathological cognitive decline shall be identified. More importantly, the Dortmund Vital Study focuses on the following research aspects, which are currently poorly understood and largely unexplored:

1. Liver Immune Behavior-Brain Axis: A core question is how specific aspects of cognitive performance and brain activities as well as the brain structure are affected by metabolic products, inflammatory mediators, or viral infections, including concentrations of metabolites such as creatinine, ammonia, or immunological parameters in the blood such as inflammatory cytokines, and novel metrics of immunological age. Combinations of standardized cognitive tests and electroencephalographic techniques with computer-based tests, as well as structural and functional magnetic resonance imaging (MRI) will be used to explore cognitive performance with a much greater level of detail compared to previous studies.
2. Multidimensional and Multichannel (Omics) Analyses: The roles of metabolic parameters or inflammation mediators have traditionally been determined by analyzing individual effects of the selected parameters per study. However, with currently available techniques and technologies, it is possible to measure a wide range of parameters (ie, metabolites, proteins, RNA species, DNA variants, and immune cells) simultaneously, using the so-called “omics” techniques or multichannel analysis. Consequently, a more comprehensive picture is obtained, which can then be evaluated using systems biology techniques to understand the interactions among the individual factors.
3. Age-Related Perspective: Temporal dynamics of the interplay between endogenous versus exogenous factors and cognitive functions, as well as work ability over the lifespan will be assessed. In particular, the effect of metabolic products, infections or inflammatory processes, and the brain structure and function on neurobehavioral parameters in young, middle, and old age will be analyzed in a cross-sectional design.
4. Disease-Related Perspective: With the present design, the derivation and evaluation of risk factors for widespread mental and neurodegenerative diseases such as depression, burnout, mild cognitive impairment, or dementia will be addressed. Moreover, epigenetic approaches that investigate the influence of environmental changes on the genome, as well as on the brain structure and function will be feasible. Epigenetic dysregulation is also important for the development of immunological and neuronal diseases and complex developmental disorders.
5. Genetic Perspective: Although it is known that genes influence cognitive performance, which in turn is associated with the structural and functional network efficiency of the brain, functional understanding of the specific brain parameters that mediate the relationship between individual genetic variations or polygenic scores and cognitive performance is incomplete. Given that the current study will capture all relevant neurocognitive measures, it will be possible to investigate the interactions among genes, brain parameters, and cognitive performance.
6. Longitudinal Perspective: Most existing studies have only used cross-sectional study designs that limit the interpretation of results, given that the results obtained in cross-sectional studies do not imply causality. Therefore, this study has been additionally designed as a longitudinal study with postmeasures every 5 years. The advantage of such a longitudinal study is that parameters, namely particular metabolic or immunological states, genetic variations, or brain metrics, can be recorded at an early stage and then be tracked to test whether they result in any effect later; moreover, one can investigate how the state of these parameters changes over the course of a person's life. This design allows causal inferences.

The benefit of the Dortmund Vital Study will be that profound knowledge regarding the interactions among metabolism, the immune system, genetics, the brain structure and function, and cognition will be gained. In the long run, this knowledge will lead to empirically based intervention strategies to preserve cognitive functioning, work ability, and promote healthy cognitive aging. The strength of the study design lies in the fact

that variables with an impact on cognitive performance and their interactions can be determined comprehensively and that factors whose significance is still unknown today will be identified and evaluated. Due to the complex design and several measurement parameters, a wide range of research questions is possible. Given below are 4 of several hypotheses possible in the study.

Hypothesis 1: Background and Mechanisms of Stress-Related Exhaustion Disorder

Burnout, a synonym for stress-related exhaustion disorder is a widespread phenomenon, but the underlying mechanisms have not yet been sufficiently explored. In addition, there is no clear distinction between burnout and depression [20,21], and there is currently limited evidence connecting burnout and cognitive functions [22,23]. However, it is known that the reward system in the orbitofrontal brain areas is altered owing to depression and subclinical depressive symptoms, which is evident in electrophysiological correlates [24]. As a result, the perception of mistakes or negative feedback is increased whereas the sensitivity to rewards is decreased [25]. The reward system is based on complex cognitive abilities, the so-called executive functions, which enable goal-directed behavior. Therefore, it is important to compare the effects of burnout and depression using the same parameters and measures and identify similarities and differences with respect to executive functions between the 2 disorders. We expect that individuals with high scores on the burnout scales will perform poorly in cognitive tasks and in tests measuring executive functions, such as working memory, task switching, decision-making, or distractibility. Additionally, we aim to explore possible differences in the resting-state electroencephalography (EEG) and functional MRI brain activity between burnout and nonburnout employees; we expect reduced alpha power and cortical hyperactivity in the former group [26] that may be the neuronal correlate of hypersensitivity to negative events. However, no effects are expected in tasks that measure basal cognitive functions such as vigilance or processing speed. Moreover, a relationship between burnout and immunological parameters [27,28] or metrics of immunological age can be assumed. For example, the concentration of hair cortisol as an index of long-term stress is expected to be positively related to the severity of burnout symptoms and reduced cognitive performance [29,30].

Hypothesis 2: Influences of Biological Factors on Cognitive Functioning Over the Lifespan

Considerable progress has been made in recent years to identify biological factors that influence cognition in older adults. These include genetic polymorphisms, such as the brain-derived neurotrophic factor (BDNF) $Val^{66}Met$ [31], catechol-O-methyltransferase (COMT) $Val^{158}Met$ [32], or latent infectious diseases (eg, *Toxoplasma gondii*) [33,34] that affect brain metabolism. However, studies with young individuals often show no effects [3]. Our hypothesis is that the effects of genes or neuropathologically relevant infections increase with age, which is probably due to the decreasing integrity of neuronal networks as one ages. To test this hypothesis, we aim to compare the performance of young, middle-aged, and old participants with and without infections, such as *Toxoplasma gondii* or COVID-19, or those with a specific gene expression pattern, with the expectation that there are larger inter- and intraindividual differences in the cognitive parameters in older adults compared to younger adults. Moreover, as this is a long-term study, we have the opportunity to test these hypotheses of altered performance in the same individuals over many years. We also aim to identify the immunological age of the participants based on various immunological parameters [35], which will allow us to determine if immunological age rather than chronological age is associated with changes in cognitive functioning.

Hypothesis 3: Interaction of Damaging and Protective Factors

Cognitive age is influenced by several internal and external factors. Apart from genetic makeup, environmental and lifestyle variables such as education, type of work, habitual physical and mental activity, nutrition, and stress at work play a modulating role and explain the high degree of variability observed with cognitive performance at advanced ages [36,37]. Thus, we aim to determine which of these factors critically influence the development of cognitive competence and the so-called cognitive reserve in older adults [38]. Consequently, we can discern whether, for example, long-term stress neutralizes the benefits of advanced education on cognitive aging, or whether regular physical activity can compensate for the negative effects of less advanced education. On the other hand, we expect that psychosocial stress (whether during childhood or adulthood) influences immunological or metabolic parameters [39] and that individuals suffering from permanent stress, which has been demonstrated for particular occupational groups, will show corresponding changes in their metabolism and cognition [40].

Hypothesis 4: Analysis of Factors Affecting Work Ability Across the Lifespan

As employees age, their physical and mental abilities tend to decline and the risks of accidents at work and work-related diseases increase [41,42]. Work ability is defined as the relationship between individual resources and specific work requirements, and it is the result of interactions between job requirements in terms of physical and mental strain, capacities and skills of the employees, as well as their health status and subjective evaluation of functioning in a given working environment. The instrument used to evaluate the ability to work is the Work Ability Index (WAI) [43]. It considers specific psychosocial and physical factors related to performing a given type of work, as well as the employee's mental and physical resources. Therefore, it is important to understand the relationships among age, lifestyle, quality of life, work-related factors, stress-related impairments, and cognitive abilities, as well as examine work ability measured by the WAI. Additionally, we will consider individual health risks, such as cardiovascular, metabolic, anthropometric, or immunological contributors to work ability. The longitudinal study design helps assess changes in several risk factors associated with compromised work ability to promote healthy aging in working environments and reduce the likelihood of early retirement.

Methods

Study Population

The participants of the Dortmund Vital Study (trial registration number: NCT05155397) represent a sample of a generally healthy population of a western society aged between 20 and 70 years. We define “healthy” in a rather broad sense and allow the inclusion of individuals who are smokers, drink alcohol, are overweight, or have a history of diseases without having severe symptoms. Moreover, we do not impose restrictions regarding education or occupation to enhance the representativity of the sample. Participants are recruited via an internet site, newspaper advertisements, reports and announcements in local print and radio media, public information events, social media, and flyers throughout the city of Dortmund, Germany. In addition, some larger companies in the region were contacted and asked to inform their employees about participating in the study. The number of subjects resulted from a biostatistical estimation (power analysis), in which the size of the expected effects and their variance were considered. The study parameters will be collected repeatedly for 4 time points, namely at the start of the study (baseline), and then 5, 10, and 15 years later (time points for measuring within-subject factors) and between 3 age groups (between-subject factor ie, Age Group: young, middle-aged, and old). Of particular interest here is the interaction between the age group and the measurement time point, which will provide insight into the different trajectories within the specific age groups. Thus, power estimation was conducted exemplarily for repeated-measures analyses of variance (ANOVA) of within-between interactions. The sample size was determined based on a small effect size of f=0.1 for the mixed ANOVA model with repeated measures (4 time points) and interactions with the Age Group factor (3 categories), given an error probability of α=.05 and a power of 0.95, resulting in a sample size of 264 individuals (determined with G*Power, a free-source statistical software package) [44]. However, having a larger number of subjects is desirable owing to expected dropouts over the long duration of the study. Therefore, a group size of approximately 600 was targeted, as we assumed a dropout rate of up to 20% per time point of the follow-up measures. Assuming this dropout rate, the number of remaining participants 20 years later would be approximately 300, which roughly corresponds to the estimated sample size, as summarized in Table 1.

Table 1. Longitudinal study design showing birth cohorts of the study from 1946 to 1996 with measurements at 5-year intervals[a].

Birth cohort	Age at T1 2016-2021	Age at T2 2021-2026	Age at T3 2026-2031	Age at T4 2031-2036
1996	20 years	25 years	30 years	35 years
1946	70 years	75 years	80 years	85 years
n	600	480	384	307

[a]The ages of the youngest (20 years old at baseline) and oldest (70 years old at baseline) cohorts and the expected number of subjects at 4 test points (T1 to T4) are indicated.

Representativeness of the Sample

To verify the representativeness of the sample and thus the generalizability of the future findings, four different aspects are considered:

1. Age Distribution and Gender: We enroll almost the same number of participants in each age group and compare the proportion in each decade of life (as well as the proportion of women and men) between the participants in this study with the corresponding proportion of the general population in Germany.
2. Genetic Representativeness and Homogeneity: We test whether the distribution of genetic polymorphisms in the sample does not differ from the Hardy-Weinberg equilibrium (HWE).
3. Cognitive Abilities: We test whether neuropsychological tests measuring crucial cognitive functions are comparable with reference values in the literature.
4. Education and Occupation: We compare levels of education and the proportion of employed individuals in the present sample and in the general population in Germany.

First, until August 2021, totally 593 participants have been enrolled. Compared with the average of 20- to 70-year-old people living in Germany [45], the participants of the Dortmund Vital Study are slightly younger (44.2 vs 45.8 years on average), and the proportion of women is higher (61.5% vs 49.6%). The proportions of participants in the Dortmund Vital Study in 5 age groups (20 to 29 years: n = 122, 20.6%; 30 to 39 years: n = 113, 19.1%; 40 to 49 years: n =110, 18.5%; 50 to 59 years: n =146, 24.6%; 60 to 70 years: n =102, 17.2) correspond largely to the proportions in the German population (20 to 29 years: 17.2 %; 30 to 39 years: 19.6 %; 40 to 49 years: 18.2 %; 50 to 59 years: 24%; 60 to 70 years: 21%).

Second, to ensure that the sample of participants of the Dortmund Vital Study is representative for the population in a genetic manner, the HWE is used to compare the expected and observed proportion of common homozygote, heterozygote, and rare homozygote variants of the single nucleotide polymorphisms (SNPs) calculated by the internet-based HWE calculator [46]. A deviation from the HWE was found in only 1 out of the 19 measured SNPs (IL-12), as shown in Table 2.

Third, we compared the scores in the neuropsychological tests between the present study and the corresponding reference or normative values available in the test manuals or literature. The Mini Mental State Examination (MMSE) used for detecting early signs of cognitive impairment shows a mean of 28.1 points (SD 1.9) out of a possible 30 points in our participants aged 60 years and older (n=102) and 28.3 points (SD 1.6) in a healthy sample of 204 participants in the same age range in the study by Votruba et al [47]. For the Beck Depression Inventory, the

mean score in our sample is 5.7 points (SD 6.4), which is well below the cutoff for minimal depression (9 points), and similar to the mean of a healthy control group (n=583) described in the test manual, which scored 7.4 points (SD 7.3) [48]. The Multiple-Choice Vocabulary Test (MWT-B) that measures premorbid (crystalized) intelligence shows a mean score of 31 points (SD 3.2) that corresponds to a mean IQ of 115. A healthy control group of adults (n=102) in a study by Satzger et al [49] reached a mean IQ of 121 (the test usually overestimates the IQ obtained with typical intelligence tests). For the digit span test from the Hamburg-Wechsler-Intelligenztest für Erwachsene, Revision (HAWIE-R), the German version of the Wechsler Adult Intelligence Scale-III [50], the mean number of correct responses is 14.1 (SD 3.6) in the present sample and 15.0 (SD 3.2) in a healthy adult sample (n=100) in the HAWIE-R manual [51]. Similarly, the Digit Symbol Test from the same test battery shows a mean score of 59.1 points (SD 12.3) for the Dortmund Vital Study participants and 50.6 (SD 10.3) points in the HAWIE-R manual. For the Stroop Test from the Nürnberger Alters Inventar (NAI) [52], the time to complete the word list for participants aged between 55 to 69 years is 14.7 (SD 2.3) seconds in the present study (n=144) and 16.0 (SD 2.0) seconds in the reference values for healthy older persons in the manual (n=78). For the color naming list of the Stroop Test, the corresponding times are 21.1 (SD 3.7) versus 24.0 (SD 7.0) seconds, and for the interference list, they are 36.7 (SD 7.4) versus 44.0 (SD 13.0) seconds. The capacity for learning new words assessed by the Verbal Learning and Memory Test (VLMT) [53], the German version of the Rey Auditory Verbal Learning Test (RAVLT) [54], shows a mean sum of 54.8 points (SD 4.3), indicating the number of successfully learned words in our sample (n=525), and exactly 54.8 points in a norm sample (n=515) provided in the test manual. Furthermore, the D2 test, measuring selective attention and attentional endurance (D2-R) [55] reveals for our participants aged between 20 and 60 years (n=489), the mean number of correctly crossed symbols is 158 (SD 36.7), whereas the normative value for the same age group (n=976) in the manual is 153 (SD 28.4). Finally, for the 2 subtests from the Leistungs-Prüf-System (LPS, meaning performance test system) validated for people older than 50 years [56], our participants aged 50 years and older (n=217) had a mean of 24.7 (SD 5.1) correct responses in the logical reasoning and 19.5 (SD 6.4) in the spatial rotation subtests, whereas the controls in the LPS manual reached 22.0 (SD 5.2) and 17.6 (SD 6.0) correct responses, respectively.

Fourth, with respect to education, 28% (n=161) of the participants of the Dortmund Vital Study have secondary degree, 30.3% (n=174) have a high school diploma and 41.6% (n= 239) have a university degree, compared to the corresponding proportions in the German population (23.5%, 33.5%, and 18.5%, respectively). Furthermore, 68.6% of the participants in the Dortmund Vital Study are employed (full-, half-, or part-time) compared to 67.7% of the general population in Germany [45].

In summary, we consider the sample of the present study as representative in terms of age distribution, genetics, depressive symptoms, cognitive parameters, and occupation. However, in contrast to the general population, the ratio of female participants and that of university degree holders appear to be higher than in the general population.

Inclusion Criteria

The study includes adults without history of severe diseases, namely neurological diseases such as dementia, Parkinson disease, or stroke; cardiovascular diseases; bleeding tendency; oncological diseases; psychiatric disorders such as schizophrenia, obsessive-compulsive disorder, anxiety disorders, or severe depression; head injuries; head surgery; head implants; eye diseases (cataract, glaucoma, or blindness); accidents that limit physical fitness and mobility; and those who do not use psychotropic drugs and neuroleptics. Medications that did not lead to exclusion from the study are blood thinners, hormones, antihypertensives, and cholesterol reducers. Participants have normal or corrected-to-normal vision and hearing and fulfill standard inclusion criteria for MRI measurements.

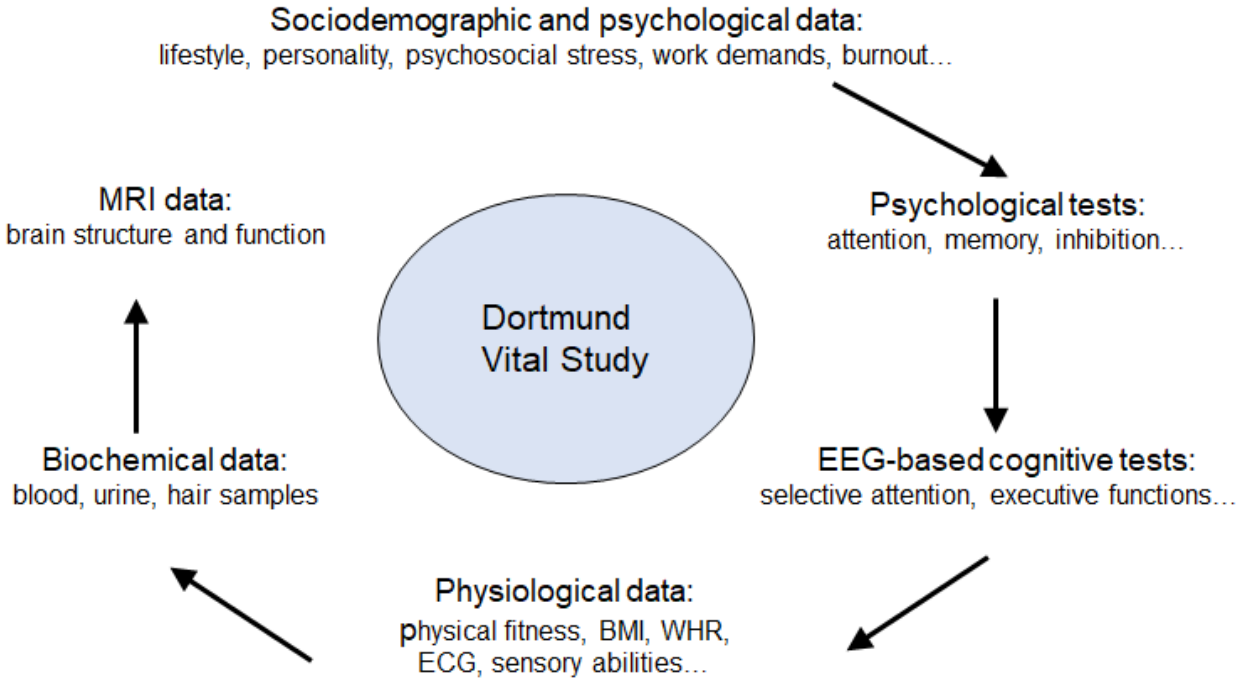
Telephone Interview

Once the participants register via a contact form, they receive a telephone call where an interview is conducted to check their inclusion criteria and obtain personal data, such as age, contact information, education, current occupation (type, full time/part-time), planned changes in the next years (relocation, retirement), pre-existing and current illnesses, physical restrictions, medication, willingness to visit the Leibniz Research Centre for Working Environment and Human Factors at the Technical University of Dortmund (IfADo) for tests on 2 independent days, and the willingness to participate in follow-up tests in 5, 10, and 15 years.

Measures

The Dortmund Vital Study uses a wide range of instruments to measure biological and environmental parameters that potentially affect cognitive performance across the lifespan, as shown in Figure 1.

Figure 1. Schematic illustration of the measures used in the Dortmund Vital Study. The arrows indicate the order in which the data are collected. Beginning with the first post-test, an MRI will be included into the test battery. EEG: electroencephalography, ECG: electrocardiography, MRI: magnetic resonance imaging.



Questionnaires

The participants receive several questionnaires via mail that should be filled in at home and brought to the first test session. The battery includes nonstandardized and standardized questionnaires. Nonstandardized questionnaires obtain sociodemographic and lifestyle variables (marital status, children, education, proficiency in languages other than the native language, type and history of employment, history of physical activity, nutrition, smoking, social activities, hobbies, using digital media, caregiving of family members), as well as request information on vision and use of glasses, necessary medical information for blood collection, and a rating of subjective time perception. Moreover, for the follow-up measurements starting in the middle of 2021, a questionnaire addressing COVID-19–specific aspects has been included that gathers information on, for example, the participants' experience with the pandemic, how they coped with the pandemic, and the consequences of a COVID-19 infection (if applicable). Standardized questionnaires are used to measure the following: depressive symptoms (Becks Depression Inventory) [57], personality traits (Big Five Personality traits, NEO Five-Factor Inventory) [58], traumatic experiences during childhood with the adapted version of the Childhood Trauma Questionnaire [59], chronotype (D-MEQ) [60], cognitive failures in daily life (Cognitive Failure Questionnaire) [61], Grit personality trait (Grit Scale) [62], handedness (Edinburgh Handedness Inventory) [63], emotional dissonance [64], job control [65], physical activity (Lüdenscheid Physical Activity Questionnaire) [66], burnout (Maslach Burnout Inventory-General Survey [67], and Oldenburg Burnout Inventory [68], stress reactivity (Perceived Stress Reactivity Scale) [69], affectivity (Positive and Negative Affect Schedule) [70], psychosocial stress (Psychosocial Stress Questionnaire) [71], general self-control and self-control at work [72], chronic stress (Trier Inventory of Chronic Stress) [73], quality of life using the short version of the World Health Organization Quality of Life questionnaire (WHOQoL-BREF) [74], and work ability (WAI) [43,75]. For the follow-up tests, the Copenhagen Psychosocial Questionnaire-III [76] will be included to assess psychosocial work demands. Some of the questionnaires were adapted to suit the purpose of the present study.

Sensory Testing

To evaluate basic sensory abilities, audiometry and visual acuity tests are conducted. Audiometric thresholds are tested for 10 pure-tone frequencies (125, 250, 500, 750, 1000, 2000, 3000, 4000, 6000, and 8000 Hz) for the left and right ears (Oscilla USB100, Inmedico).

Visual acuity is measured using Vistec's Optovist according to the DIN 58220-3 standard "Visual acuity testing - Part 3: Test for use in expertise" for far vision with the right and left eye separately (monocular) and with both eyes together (binocular), and if available with correction for distance (glasses for far vision). The "inclined optometer" [77] is used to determine the zones of sufficient binocular vision at horizontal gaze inclination. For this, the near and far points are obtained, if available with distance correction (glasses for far vision). Moreover, the participants complete a questionnaire regarding the glasses they wear, and the visual and musculoskeletal complaints during the use of computer monitors at the workplace.

Neuropsychological Assessment

A wide range of cognitive functions are evaluated using the following standardized neuropsychological tests: evaluate the memory span and working memory (WM) using digit span

forward and backward from HAWIE-R [51], the German version of WAIS-IIIR [50]; semantic memory in written and spoken versions (Word-Fluency Test from LPS) [56]; selective attention and attentional endurance (D2-R) [55]; lateralization and motor functions (Perdue Pegboard Test) [78]; crystallized intelligence (MWT-B) [79]; general cognitive status (MMSE) [80], different aspects of verbal memory like learning performance and retrieval (VLMT) [53], a German version of the RAVLT [54], psychomotor performance and speed of processing (Digit Symbol Test from HAWIE) [51], interference control and inhibition (Stroop Test) [81] from NAI [52], task switching (Trail Making Test, TMT-A and TMT-B) [82], and 2 subtests from the performance testing system [56] measuring logical reasoning and spatial rotation (refer to [83] for details of the tests). For the follow-up measurements, Raven's 2 [84] will be included to estimate fluid intelligence and the number connection test [85] for measuring processing speed. Finally, all cognitive measures will be used to generate the *g*-factor also known as general intelligence. The *g*-factor is a construct developed in psychometric studies of cognitive abilities and human intelligence. It is based on the observation that performance of different cognitive tasks is positively correlated, reflecting the fact that an individual's performance of 1 type of cognitive task tends to be comparable to that person's performance of other cognitive tasks [86].

Computer-Based Cognitive Tests With EEG Recording

Behavior and electrical brain activity are simultaneously measured using 2 separate test batteries consisting of a total of 11 computer-based cognitive tasks measuring crucial cognitive functions. The test batteries are applied on 2 examination days. Before the start of the test batteries on days 1 and 2, EEG activity is measured for 2 minutes with eyes open, and 2 minutes with eyes closed to evaluate the resting-state oscillatory brain activity. To assess potential time-on-task effects on brain activity, these EEG measurements are repeated posttesting after the cognitive tests are completed.

The computerized test battery on day 1 includes the following tasks:

Bar Task [87]

The aim of this task is to evaluate attentional performance in a perceptual control task. The task is to respond to luminance changes in 1 of 2 symmetrically presented bars and to press the corresponding response key.

Psychomotor Vigilance Test [88]

The standard 10-minute psychomotor vigilance test measures sustained or vigilant attention by recording response times to visual stimuli that occur at random interstimulus intervals. The 4 interstimulus intervals are 2, 3, 5, and 8 seconds.

Simon Task [89,90]

The Simon task measures stimulus-response compatibility and conflict processing. The Simon effect elucidated in the task refers to the observation that spatially arranged responses to nonspatial stimulus features (eg, shape, color) are faster when the task-irrelevant stimulus location and the response are on the same side than when they are on opposite sides.

AX-Continuous Performance Task (AX-CPT) [91]

The AX-CPT is used to measure updating and strategy learning. It has been commonly used to examine shifts in the use of proactive and reactive cognitive control. The AX-CPT requires participants to respond to a certain cue-probe pair (ie, target cue-target probe; AX trials) and to withhold their response or make an alternate response or use other cue-probe pairs. The proactive control mode has been associated with cue-driven processing. In contrast, the reactive control mode has been associated with probe-driven processing. The shift between these alternative control modes can be assessed by comparing different cue-probe combinations.

Speech-in-Noise Perception Task [92]

This task measures speech understanding and auditory distractibility. This is examined in a simulated auditory stock market scenario in which the subjects must respond to a target company and its market value. This target company is included in 50% of the trials, the price of which is 50% above or below the critical value of 5. The target company and value are presented in the presence of 2 competing companies to provide distractive stimuli that must be inhibited.

The test battery on the second testing day primarily assesses executive functions.

n-Back Task [83,93]

The *n*-back task is assumed to be a measure of WM capacity because it requires maintaining, continuous updating, and processing of information. The task consists of a 2-choice condition with low WM-load (0-back) and a 2-back condition with high WM-load. Participants are successively presented with a series of visual stimuli; for each stimulus, they are asked whether it matches a stimulus presented *n* trials before. For example, in a *2*-back task, in which the stimuli consist of letters, participants must decide whether the current letter is the same as the letter in trial *n*–2.

Task Switching (Cue- and Memory-Based Task Switching) [94,95]

This task switching paradigm is frequently used to evaluate several control processes by applying only 1 experimental task consisting of different experimental blocks. It enables assessment of task preparation, WM, interference processing, and switching processes, as well as their interaction. First, participants are asked to perform 3 different single task conditions using the same stimulus material (numerical, parity, and font size tasks) in separate blocks of trials. In the next step, they are asked to switch between the 3 tasks in a randomized order using a cue stimulus that signals the relevant task in advance. In the last block, they switch between the tasks every third trial without any cue stimuli, requiring memorizing and recalling the task sequence.

Auditory Distraction [96,97]

This auditory task evaluates the ability to focus on a given task and ignore concurrent distracting stimuli. Participants perform a duration discrimination task on a random sequence of long- and short-tone stimuli, which either have a standard pitch (80%) or a deviant pitch (20%). EEG correlates of involuntary shifts

in attention to the task-irrelevant deviant stimuli and subsequent reorientation to the task-relevant stimulus feature are evaluated.

Interference Processing [81,98]

The Stroop task measures the susceptibility to interference and the capacity to inhibit irrelevant stimuli and prepotent responses. Participants must indicate the meaning or the color of color words whereas the word color is either congruent or incongruent with the word's meaning. One-half of the trials are congruent (name and color are the same), the other half are incongruent, and both types of trials are presented in a randomized order. As word reading is an automated response and produces no or little interference in congruent trials, naming of the font color and inhibiting the word's meaning in an incongruent trial is a complex executive function that produces strong interference.

Cognitive Inhibition (Go/NoGo) [99,100]

A standard task to evaluate inhibitory control is the Go/NoGo task, in which participants are asked to respond under time pressure to frequent stimuli (letter K) while refraining from responding to the rare stimuli (letter T). First, participants conduct a baseline block to estimate their mean response time, which later serves as a time limit in the test block. In case participants respond slower than the individual time limit, visual feedback prompts them to respond faster (which usually leads to a higher rate of false alarm, ie, failure of inhibition).

Spatial Selective Attention (Visual Search) [101,102]

The visual search task measures visual selective attention. Participants search for a target item presented together with 8 distractor items (a matrix consisting of colored arrows with different orientations). In half of the trials, 1 of 2 predefined targets are present, whereas in the other half, only distractors are presented. Subjects respond to indicate that they have detected a target. The 2 dependent measures that are most commonly studied are reaction time and the ratio of detected targets.

For the measurement on day 1, the EEG is recorded from 64 electrodes positioned according to the extended 10-20 system using a Brain Amp DC amplifier. The data are filtered online at 250 Hz DC, and the sampling rate is set at 1000 Hz. On day 2, the EEG is recorded from 32 active electrodes positioned according to the extended 10-20 system, using a BioSemi system (BioSemi Instrumentation). The sampling rate is 2048 Hz. A common mode sense (CMS) active electrode and a driven right leg (DRL) passive electrode are used. These 2 electrodes form a feedback loop, which drives the average potential of the subject. The reference and ground electrodes are integrated into the CMS und DRL loop. For the follow-up measurements, an extended electrode montage with 64 channels will be used.

Structural and Functional MRI

With the start of the second measurement series beginning in late 2021, structural and functional MRI will be included to complement previous neurocognitive measurements. The resting-state MRI data will be recorded to quantify functional network properties and multishell diffusion-weighted imaging as well as multiparametric mapping (qMRI) to estimate the local and network architecture of gray and white matter. In addition, arterial spin labeling will be performed to quantify cerebral blood flow and magnetic resonance spectroscopy to record various parameters of brain metabolism. In the abdomen, the distribution and concentration of visceral and subcutaneous fat as well as the fat and iron concentration of the liver will be measured.

Measurement of Biological Parameters

Physical Fitness Test, and Cardiovascular and Anthropometric Parameters

Participants' current physical performance is measured using a bicycle ergometer with a physical work capacity (PWC-130) cycle test. The aim of this test is to predict the absolute power output at a projected heart rate of 130 beats per minute. The relative power output is calculated by the power-to-weight ratio. In addition, the heart rate, electrocardiography (ECG) during rest, as well as the systolic and diastolic blood pressures before and during ergometry are measured. Additionally, the height, weight, waist-to-hip ratio, and BMI are obtained from each participant.

Immunological Parameters

We collected 80 ml peripheral venous blood from all participants. Immunological parameters are determined by analyzing the absolute numbers of lymphocytes, quantitative and qualitative changes in the composition of lymphocyte subsets, concentration of cytokines in serum, namely interleukin (IL)-1b, interferon (IFN)-alpha, IFN-gamma, tumor necrosis factor (TNF)-alpha, monocyte chemoattractant protein (MCP-1, IL-6, IL-8, IL-10, IL-12p70, IL-17A, IL-18, IL-23, and IL-33) and functional activities of T cells and natural killer cells [103]. A metric of immune age [35], which was recently developed from a longitudinal cohort using cell subset phenotyping, functional responses of cells to cytokine stimulations and whole blood gene expression data, was adapted to our data by approximation using principal component regression based on the composition of lymphocyte subsets (natural killer and T cells, CD4+, CD8+, and CD8+CD28 T cells). Additionally, peripheral blood mononuclear cells (CD14 positive monocytes) as well as proteins (C-reactive protein), RNA, and DNA are extracted and analyzed using Omics methods.

Metabolic Parameters

A urine sample is collected early in the morning before starting the tests, from which creatinine and calcium oxalate are extracted later. Furthermore, ammonia concentration, leucocytes, erythrocytes, hematocrit, monocytes, lymphocytes, cell volume, thrombocytes, triglycerides, cholesterol, high- and low-density lipoprotein cholesterol, glycosylated hemoglobin, glucose, C-reactive protein, and creatinine are measured in venous blood. Later, we plan to measure choline, betaine, and glycerophosphocholine in the blood.

Endocrine Parameters

If possible, a hair sample is taken from which hair cortisol concentrations are measured as an index of long-term stress [104].

Infections

The presence of a latent *Toxoplasma gondii* infection is explored using specific IgG antibodies. In the follow-up testing, a measure of COVID-19 antibodies will also be obtained.

Genetic Parameters (SNPs)

Isolation of genomic DNA of leukocytes is performed according to standard procedures [105]. Several genetic polymorphisms potentially related to the structure and function of the central nervous system are selected. The analyzed SNPs are presented in Table 2.

Table 2. Single nucleotide polymorphisms measured in the Dortmund Vital Study (N=528) with the number of observed and expected common homozygotes, heterozygotes, and rare homozygotes in parentheses computed using the Hardy-Weinberg equilibrium calculated on the internet [46][a].

Genotype	rs (reference SNP[b] cluster ID)	Allele observed (n)	Allele expected (n)	χ^2 (*df*)
Apo[c]-E2, E3, E4				
	7412	0=CC (443), 1=CT (83), 2=TT (3)	0=CC (444), 1=CT (81), 2=TT (4)	0.14 (2)
	429358	0=CC (10), 1=CT (127), 2=TT (389)	0=CC (10), 1=CT (126), 2=TT (389)	0.01 (2)
BDNF[d] Val[66]Met	6265	0=GG (365), 1=GA (148), 2=AA (15)	0=GG (365), 1=GA (148), 2=AA (15)	0.00 (2)
COMT-1[e]	4633	0=CC (128), 1=CT (256), 2=TT (144)	0=CC (124), 1=CT (263), 2=TT (140)	0.45 (2)
COMT-2 Val[158]Met	4680	0=AA (144), 1=AG (256), 2=GG (128)	0=AA (140), 1=AG (263), 2=GG (124)	0.45 (2)
DRD2[f]	6277	0=CC (112), 1=CT (275), 2=TT (141)	0=CC (118), 1=CT (263), 2=TT (146)	1.06 (2)
DRD1[g]-48A/G	4532	0=CC (39), 1=CT (129), 2=TT (93)	0=CC (41), 1=CT (124), 2=TT (95)	0.27 (2)
CHRNA6[h]-1	1072003	0=CC (346), 1=CG (161), 2=GG (21)	0=CC (344), 1=CG (163), 2=GG (19)	0.17 (2)
CHRNA6-3	2304297	0=AA (32), 1=AG (200), 2=GG (296)	0=AA (33), 1=AG (198), 2=GG (297)	0.05 (2)
CHRNB3[i]-1	13280604	0=AA (311), 1=AG (192), 2=GG (25)	0=AA (314), 1=AG (186), 2=GG (28)	0.40 (2)
CHRNB3-2	4950	0=AA (311), 1=AG (191), 2=GG (26)	0=AA (313), 1=AG (187), 2=GG (28)	0.23 (2)
GPCPD1[j] (EDI3)	6116869	0=GG (198), 1=GT (249), 2=TT (79)	0=GG (198), 1=GT (249), 2=TT (79)	0.00 (2)
GRIN2A[k]	1969060	0=CC (19), 1=CT (165), 2=TT (342)	0=CC (19), 1=CT (164), 2=TT (343)	0.02 (2)
GRIN2A	8057394	0=GG (285), 1=GC (207), 2=CC (36)	0=GG (286), 1=GC (205), 2=CC (37)	0.03 (2)
GRIN2B[l]	890	0=GG (126), 1=TG (254), 2=TT (148)	0=GG (121), 1=TG (263), 2=TT (143)	0.69 (2)
IL[m]-1beta	16944	0=GG (235), 1=GA (221), 2=AA (70)	0=GG (227), 1=GA (237), 2=AA (62)	2.43 (2)
IL-6	1800795	0=CC (85), 1=CG (265), 2=GG (176)	0=CC (90), 1=CG (255), 2=GG (181)	0.78 (2)
IL-12A	568408	0=AA (9), 1=AG (141), 2=GG (178)	0=AA (19), 1=AG (120), 2=GG (188)	*9.53* (2)
TNF-alpha[n]	1800629	0=AA (15), 1=AG (141), 2=GG (372)	0=AA (14), 1=AG (143), 2=GG (370)	0.13 (2)

[a]The chi-square test indicates the conformity between the expected and observed distribution. Significant deviances from the HWE are italicized.

[b]SNP: single nucleotide polymorphism.

[c]APO: apolipoprotein.

[d]BDNF: brain-derived neurotrophic factor.

[e]COMT: catechol-O-methyltransferase.

[f]DRD2: dopamine receptor D2.

[g]DRD1: dopamine receptor D1.

[h]CHRNA6: cholinergic receptor nicotinic alpha 6.

[i]CHRNB3: cholinergic receptor nicotinic beta 3.

[j]GPCPD1: glycerophosphocholine phosphodiesterase.

[k]GRIN2A: glutamate ionotropic receptor NMDA type subunit 2A.

[l]GRIN2B: glutamate ionotropic receptor NMDA type subunit 2B.

[m]IL: interleukin.

[n]TNF-alpha: tumor necrosis factor-alpha.

In future analyses, we intend to evaluate polygenic scores (PGS) because based on the findings of several recent large-scale genome-wide association studies, it became increasingly clear that thousands of alleles across the genome (polygenic architecture) contribute to interindividual variations in cognitive performance with small effect sizes [19]. DNA genotyping will be carried out using the Illumina Infinium Global Screening Array 3.0 with major depressive disorder and Psych content, and genome-wide PGS (weighted sums of each participant's trait-associated alleles across all SNPs) will be created using publicly available summary statistics for cognitive performance and other available behavioral phenotypes of interest.

Research Data Management

Types of Data

The types of data generated are qualitative and quantitative questionnaire data, sensory and psychometric scores, behavioral data from psychological experiments, EEG data, MRI data (starting with the second measurement wave), concentrations of immunological parameters, endocrinological data, genetic polymorphisms and PGS, ECG and cardiovascular data, anthropometric data, and the concentrations of metabolites.

Measures taken for quality assurance and quality management comprise continuous control of raw data, sample data analysis, data validation through split-half, test-retest reliability between measures, and data validation by statistical analysis.

The used formats of the generated EEG data are MATLAB (*.m), BrainVision Core Data Format (each recording consisting of a *.vhdr, *.vmrk, *.eeg file triplet) and BioSemi EEG data (*.bdf). SPSS (*.sav, *.spv) and R-data are used for statistical analyses. Integration of all data types takes place in a Structured Query Language (SQL) database. If appropriate, raw data (eg, immunological data, SNPs, questionnaires, and neuropsychological tests) are included in the database. Some types of data, namely preprocessed EEG data, are included in the database after averaging for each person, task, and condition separately. Useful variables and total scores are extracted from the large number of questionnaires and neuropsychological tests. The EEG and MRI raw data will be later converted into the Brain Imaging Data Structure (BIDS) format [106,107]. The aim of BIDS is to create standards allowing researchers to readily organize and share study data within and between laboratories.

Storage, Selection, and Retention Period

All data are stored on IfADo servers as working copies, as proof of good scientific practice, for reuse, and for legal and contractual requirements. Research data are stored and analyzed without reference to the personal data of the subjects. Personal subject information (such as name and contact data required for re-invitation) are physically separated from the research data (pseudonymization). Only a few selected persons have access to the subject's identities and the subject IDs that are stored in written form in a secure location. The anonymized research data are stored in server rooms at different locations throughout the institute. The amount of data is approximately 1 TB per year excluding MRI data, which is estimated to be approximately 2 TB per year. Technologies used for data storage include an NAS (Network Attached Storage) server with the RAID5 (redundant array of independent disks) configuration. Additional backup copies are regularly made and checked. Paper and pencil tests and questionnaires are stored in a central archive of the institute. Isolated peripheral blood mononuclear cells, serum, and urine probes are frozen for future use.

Data Access and Use

To access the data, scientists from IfADo, as well as external cooperation partners who plan to analyze data from the Dortmund Vital Study fill in a proposal form that includes a short description of the project and the respective hypotheses, responsible persons, cooperation partners, data usage, and analysis strategy. The requested research data will be made available in an anonymized form after consultation with the scientists responsible for the data. Responsible persons include project managers and coordinators of the Dortmund Vital Study in consultation with the IfADo Research Data Management Unit.

After primary analyses and publication of the main results, the data and the scripts used for data analyses will be made available in repositories for secondary analyses. The transfer agreement will be prepared by the coordinators of the Dortmund Vital Study in consultation with the IfADo Research Data Management Unit. Data access will be restricted and will require a structured project proposal. Access to the SQL database will be password protected. Interoperability will be guaranteed by metadata included in the SQL database and common data formats (eg, BIDS) to enhance exchange, management, and documentation. Data identifiers (digital object identifiers) will be assigned at least until the end of the projects.

Organization, Management, and Policies

Central organizational support for data management is provided by the Research Data Management Unit at IfADo.

The Dortmund Vital Study is funded by the institute's budget. Thus, the study design, collection, management, analysis, interpretation of data, writing of the report, and the decision to submit the report for publication is not influenced by or biased toward any sponsor.

Permissions for the tests, questionnaires, and software used that are subject to copyright were obtained from the corresponding publishers. If no license was required, permission was granted by the authors of the corresponding questionnaire or test.

Data Analysis

Detailed descriptions of the specific data analysis methods depending on the type of measure will be provided in subsequent publications.

Generally, for all quantitative variables, descriptive statistics, frequencies, and distributions will be calculated for all participants. Summary statistics including the mean, SD, minimum values, and maximum values will be provided for quality assurance. Qualitative variables will be categorized before performing any statistical analysis. Immunological, biochemical, medical, optometric, and audiometric data as well as coding for genetic polymorphisms are integrated in the SQL

database. Data from psychometric tests in the SQL database will be verified by using an automated MATLAB script for plausibility and outliers. EEG and behavioral data from experimental tasks will be analyzed with EEGLAB [108] using scripts. The main parameters, such as individual amplitudes or latencies of transient components, time-frequency parameters, or behavioral data, such as reaction times, error rates, and SDs, will be automatically written in the SQL database.

Depending on the research topics and questions, different methods will be used for data analysis. For the cross-sectional analyses, mixed ANOVA, Analyses of Covariance (ANCOVA), 1-way ANOVA, or *t* tests will be used. Furthermore, the Pearson correlation coefficient or appropriate regression models will be used to identify predictors of the variables of interest. Moreover, the structural equation modeling approach and novel decoding techniques will be employed when appropriate.

For the longitudinal data, mixed model ANOVA including within-subject factors, such as Time Point of Measure, and between-subject factors, such as Age Group, will be employed. Interactions will be analyzed using simple ANOVA or *t* tests. Additionally, covariance analysis (ANCOVA) will be conducted to control for potential confounding factors like gender and education. It is not intended to impute missing data for ANOVA or ANCOVA. In case of a few or moderate number of missing data, the analyses will be conducted with the complete cases. In case of a larger number of missing data over the long period, advanced statistical approaches like the Generalized Estimating Equation (GEE) [109] or Mixed Effect Regression (MER) [110] will be used.

Ethics Approval

The first run of the Dortmund Vital Study (2016-2021) was conducted with approval from the local Ethics Committee of the IfADo in 2015. Follow-up testing was approved in July 2021 by the Ethics Committee of the IfADo (approval number: A93-3).

Results

The baseline testing was conducted between 2016 and 2021. As of August 2021, 593 participants were tested. In October 2021, the first follow-up measure was started. Some initial results from the cross-sectional part of the study were already published [83, 111-116]. Longitudinal data will be analyzed, and the first publications are estimated for 2025.

Discussion

Study Overview

The Dortmund Vital Study is a multidisciplinary long-term project, combining a cross-sectional and longitudinal study design. The goal is to investigate the influence of a wide range of biological and environmental factors on cognitive functions and their neurophysiological correlates as well as on the immune system and other physiological functions and structures. The combination of well-elaborated experimental paradigms reflecting basic cognitive functions, modern EEG methodology, MRI scans, lifestyle data, and the analysis of biochemical parameters will facilitate the development and evaluation of specific hypotheses on the mechanisms of healthy and pathological aging from different perspectives and disciplines. Work-related human factors (eg, type of work), work ability, and the role of work conditions (eg, stress or job satisfaction) will be of particular interest. Moreover, the Dortmund Vital Study has the potential to shed tremendous light on the complex interactions between immunological, genetic, metabolic, and brain-related parameters in healthy adults. Specifically, parameters obtained by MRI and EEG-related measures can be evaluated as a function of the PGS, metabolic products, concentration of immune cells, immune age, and infections such as Toxoplasmosis or COVID-19 that are largely unexplored. The same is true for environmental and lifestyle factors that influence brain activity and behavior. As human behavior and cognition reflects a combination of these factors, the Dortmund Vital Study has the potential to elucidate several important findings. Crucial points are the developmental aspect and the progress or decline of several functions over the lifespan. This design enables analysis of changes of inter- and intraindividual variability with increasing age and lead to conclusions, such as which factors in the past have the largest impact on the present measures or pathological outcomes like burnout, depression, mild cognitive impairment, or dementia. In conclusion, this multidomain and interdisciplinary study will help elucidate underlying mechanisms of healthy and pathological aging.

Limitations and Future Challenges

A challenging problem in longitudinal study designs is the dropout rate, especially during long test periods like in this study. The dropout rate can be affected by unpredictable events like the present COVID-19 pandemic. A reduced number of participants during follow-up produces missing values, shrinks the entire sample size, reduces the statistical power, and may lead to a small sample size at the end of the study relative to the number of analyzed variables.

Thus, it is important to analyze the reasons for dropouts and whether the participants are missing at random or not at random (ie, due to unsystematic or systematic reasons), which may differently affect the results. For example, dropouts due to lack of interest in further participation in the study or change of residence can be treated as random. In contrast, dropouts due to illness or death could reflect a systematic reason for missing data, given that elderly participants entering the study have a much higher risk of illness or death in the long period spanning 20 years. This may lead to a disproportionate dropout in this age group compared to the younger groups. Moreover, a pandemic like COVID-19 reflects a systematic source of dropouts.

Missing data have important consequences for data analysis. Repeated measures ANOVA or MANOVA requires fully complete data sets for each repeated measure. That is, without data imputation, only a reduced (but complete) data set can be used for analysis. There are several methods to counteract this using statistical approaches. For example, the missing data can be replaced by the last measured data or by a simple or conditional mean of the variable resulting from a predicted value. However, data imputation is generally not recommended,

and instead, other methods should be used to deal with missing data sets [117].

For the present design, a GEE [109] model will be used, which is robust to misspecifications of the repeated measures' correlation structure, time irregularities, and does not require excluding participants with incomplete data sets. An alternative method would be MER [110], in which random effects can serve to describe each participant's trend over time. MER assesses longitudinal changes of several outcomes and is even more robust to missing data than the GEE. Both models allow time-invariant predictors (eg, biological gender, genotype) and time-varying predictors (eg, age, metabolic, immunological, or cognitive parameters), and handle irregularly timed and missing data without the need for imputation [118]. An advantage of these procedures is that both can model time-varying predictions useful for understanding changes in cognitive aging.

Nevertheless, the relatively large number of participants in this study that primarily aims to analyze cognitive changes using neuropsychology, EEG, and MRI is unique and even taking dropouts into account allows answering important scientific questions. For example, when interested in age-related changes in an electrophysiological parameter, a group of approximately 40 participants would still be sufficient to track the changes across the span of 20 years. Generally, we expect the dropout rate to remain constant and to be able to analyze the data with a sufficiently large sample at least for the first follow-up measures 5 and 10 years after the beginning of the study.

Acknowledgments

We wish to thank the following technical staff at the Leibniz Research Centre for Working Environment and Human Factors at the Technical University of Dortmund: Claudia Brockhaus, Pia Deltenre, Barbara Foschi, Judith Geiger, Silke Joiko, Nicole Klaschik, Miriam König, Kirsten Liesenhoff-Henze, Karin Lukaszewski, Trung-Tin Pham, Carola Reiffen, and Christiane Westedt as well as numerous student assistants for conducting telephone interviews and organizing and conducting testing and data preprocessing. Furthermore, we would like to thank Ludger, Tobias Blanke, and Ewald Alshuth for developing the software and the database and their technical support, and finally special thanks to all participants of the study and the 2 anonymous reviewers for the earlier version of the article. The publication of this article was funded by the Open Access Fund of the Leibniz Association.

Authors' Contributions

All authors contributed to the design of the study. Being the principal investigator of the study, EW supervised the organization of the measures, and selected, designed, and analyzed the electroencephalography test battery. PDG and SG selected the tests and questionnaires, conducted data acquisition and integration, supervised and registered the study, and wrote the article. SG, PDG, and EW organized research data management, with support from PB, MC, and JR from the Leibniz Research Centre for Working Environment and Human Factors at the Technical University of Dortmund (IfADo) Research Data Management Unit. PB, SC, MC, and CW are responsible for recording the immunological parameters and the corresponding data analysis. MB, EG, and MN designed the magnetic resonance imaging (MRI) sequences and are responsible for MRI data acquisition and analysis. TK is responsible for selection of work-related questionnaires. CC and RM are responsible for the inclusion and analyses of metabolic parameters. KG, JGH, JR, and CvT are responsible for recording the anthropometric, cardiovascular, endocrine, genetic, infectious, and metabolic parameters, and the corresponding data analysis. All authors critically revised and approved the manuscript.

Conflicts of Interest

None declared.

References

1. Verhaeghen P, Cerella J. Aging, executive control, and attention: a review of meta-analyses. Neurosci Biobehav Rev 2002 Nov;26(7):849-857. [doi: 10.1016/s0149-7634(02)00071-4] [Medline: 12470697]
2. Salthouse TA. Trajectories of normal cognitive aging. Psychol Aging 2019 Feb;34(1):17-24 [FREE Full text] [doi: 10.1037/pag0000288] [Medline: 30211596]
3. Colzato L, Zhang W, Beste C, Stock AK. Dissociating direct and indirect effects: a theoretical framework of how latent toxoplasmosis affects cognitive profile across the lifespan. Neurobiol Aging 2021 Jun;102:119-128 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2021.02.007] [Medline: 33765425]
4. Pistarini C, Fiabane E, Houdayer E, Vassallo C, Manera MR, Alemanno F. Cognitive and emotional disturbances due to COVID-19: an exploratory study in the rehabilitation setting. Front Neurol 2021 May;12:643646 [FREE Full text] [doi: 10.3389/fneur.2021.643646] [Medline: 34079511]
5. Marsland AL, Gianaros PJ, Kuan DC, Sheu LK, Krajina K, Manuck SB. Brain morphology links systemic inflammation to cognitive function in midlife adults. Brain Behav Immun 2015 Aug;48:195-204 [FREE Full text] [doi: 10.1016/j.bbi.2015.03.015] [Medline: 25882911]

6. Sartori A, Vance D, Slater L, Crowe M. The impact of inflammation on cognitive function in older adults: implications for healthcare practice and research. J Neurosci Nurs 2012 Aug;44(4):206-217 [FREE Full text] [doi: 10.1097/JNN.0b013e3182527690] [Medline: 22743812]
7. Wilkinson DJ, Smeeton NJ, Watt PW. Ammonia metabolism, the brain and fatigue; revisiting the link. Prog Neurobiol 2010 Jul;91(3):200-219. [doi: 10.1016/j.pneurobio.2010.01.012] [Medline: 20138956]
8. Kosenko E, Kaminsky Y, Solomadin I, Marov N, Venediktova N, Felipo V, et al. Acute ammonia neurotoxicity in vivo involves increase in cytoplasmic protein P53 without alterations in other markers of apoptosis. J Neurosci Res 2007 Aug;85(11):2491-2499. [doi: 10.1002/jnr.21385] [Medline: 17551980]
9. Ballesteros S, Kraft E, Santana S, Tziraki C. Maintaining older brain functionality: a targeted review. Neurosci Biobehav Rev 2015 Aug;55:453-477. [doi: 10.1016/j.neubiorev.2015.06.008] [Medline: 26054789]
10. Greenwood PM, Parasuraman R. Neuronal and cognitive plasticity: a neurocognitive framework for ameliorating cognitive aging. Front Aging Neurosci 2010 Nov;2:150 [FREE Full text] [doi: 10.3389/fnagi.2010.00150] [Medline: 21151819]
11. Gajewski PD, Falkenstein M. Neurocognition of aging in working environments. ZAF 2011 Aug;44(4):307-320. [doi: 10.1007/s12651-011-0090-6]
12. Gajewski P, Falkenstein M. Lifestyle and interventions for improving cognitive performance in older adults. In: Raab M, Lobinger B, Hoffmann S, Pizzera A, Laborde S, editors. Performance Psychology. Perception, Action, Cognition, and Emotion. Oxford: Elsevier Academic Press; 2016:189-203.
13. Gajewski P, Falkenstein M. Physical activity and inactivity impacts on cognitive and emotional functioning in later life. In: Oxford Research Encyclopedia of Psychology. Oxford: Oxford University Press; 2018:1-44.
14. Carrera-Bastos P, Fontes M, O'Keefe J, Lindeberg S, Cordain L. The western diet and lifestyle and diseases of civilization. Res Rep Clin Cardiol 2011 Mar:15-35. [doi: 10.2147/rrcc.s16919]
15. Alexander J, Hillier A, Smith R, Tivarus M, Beversdorf DQ. Beta-adrenergic modulation of cognitive flexibility during stress. J Cogn Neurosci 2007 Mar;19(3):468-478. [doi: 10.1162/jocn.2007.19.3.468] [Medline: 17335395]
16. Lupien SJ, McEwen BS, Gunnar MR, Heim C. Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat Rev Neurosci 2009 Jun;10(6):434-445. [doi: 10.1038/nrn2639] [Medline: 19401723]
17. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun 2018 May;9(1):2098 [FREE Full text] [doi: 10.1038/s41467-018-04362-x] [Medline: 29844566]
18. Erickson KI, Kim JS, Suever BL, Voss MW, Francis BM, Kramer AF. Genetic contributions to age-related decline in executive function: a 10-year longitudinal study of COMT and BDNF polymorphisms. Front Hum Neurosci 2008 Sep;2:11 [FREE Full text] [doi: 10.3389/neuro.09.011.2008] [Medline: 18958211]
19. Plomin R, von Stumm S. The new genetics of intelligence. Nat Rev Genet 2018 Mar;19(3):148-159 [FREE Full text] [doi: 10.1038/nrg.2017.104] [Medline: 29335645]
20. Bakker AB, Schaufeli WB, Demerouti E, Janssen PP, Van Der Hulst R, Brouwer J. Using equity theory to examine the difference between burnout and depression. Anxiety Stress Coping 2000 Jul;13(3):247-268. [doi: 10.1080/10615800008549265]
21. Bianchi R, Schonfeld IS, Laurent E. Burnout–depression overlap: a review. Clin Psychol Rev 2015 Mar;36:28-41. [doi: 10.1016/j.cpr.2015.01.004] [Medline: 25638755]
22. Gajewski PD, Boden S, Freude G, Potter GG, Claus M, Bröde P, et al. Executive control, ERP and pro-inflammatory activity in emotionally exhausted middle-aged employees. Comparison between subclinical burnout and mild to moderate depression. Psychoneuroendocrinology 2017 Dec;86:176-186. [doi: 10.1016/j.psyneuen.2017.09.017] [Medline: 28972891]
23. Potter G, Hatch D, Hagy H, Radüntz T, Gajewski P, Falkenstein M, et al. Slower information processing speed is associated with persistent burnout symptoms but not depression symptoms in nursing workers. J Clin Exp Neuropsychol 2021 Feb;43(1):33-45. [doi: 10.1080/13803395.2020.1863340] [Medline: 33402015]
24. Wacker J, Dillon D, Pizzagalli D. The role of the nucleus accumbens and rostral anterior cingulate cortex in anhedonia: integration of resting EEG, fMRI, and volumetric techniques. Neuroimage 2009 May;46(1):327-337 [FREE Full text] [doi: 10.1016/j.neuroimage.2009.01.058] [Medline: 19457367]
25. Cardenas L, Tremblay LK, Naranjo CA, Herrmann N, Zack M, Busto UE. Brain reward system activity in major depression and comorbid nicotine dependence. J Pharmacol Exp Ther 2002 Sep;302(3):1265-1271. [doi: 10.1124/jpet.302.3.1265] [Medline: 12183688]
26. Golonka K, Gawlowska M, Mojsa-Kaja J, Marek T. Psychophysiological characteristics of burnout syndrome: resting-state EEG analysis. Biomed Res Int 2019 Jul;2019:3764354 [FREE Full text] [doi: 10.1155/2019/3764354] [Medline: 31467886]
27. De Vente W, Olff M, Van Amsterdam JGC, Kamphuis JH, Emmelkamp PMG. Physiological differences between burnout patients and healthy controls: blood pressure, heart rate, and cortisol responses. Occup Environ Med 2003 Jun;60(Suppl 1):i54-i61 [FREE Full text] [doi: 10.1136/oem.60.suppl_1.i54] [Medline: 12782748]
28. Danhof-Pont MB, van Veen T, Zitman FG. Biomarkers in burnout: a systematic review. J Psychosom Res 2011 Jun;70(6):505-524. [doi: 10.1016/j.jpsychores.2010.10.012] [Medline: 21624574]

29. Russell E, Koren G, Rieder M, Van Uum S. Hair cortisol as a biological marker of chronic stress: current status, future directions and unanswered questions. Psychoneuroendocrinology 2012 May;37(5):589-601. [doi: 10.1016/j.psyneuen.2011.09.009] [Medline: 21974976]
30. Staufenbiel SM, Penninx BW, Spijker AT, Elzinga BM, van Rossum EF. Hair cortisol, stress exposure, and mental health in humans: a systematic review. Psychoneuroendocrinology 2013 Aug;38(8):1220-1235. [doi: 10.1016/j.psyneuen.2012.11.015] [Medline: 23253896]
31. Toh YL, Ng T, Tan M, Tan A, Chan A. Impact of brain-derived neurotrophic factor genetic polymorphism on cognition: a systematic review. Brain Behav 2018 Jul;8(7):e01009 [FREE Full text] [doi: 10.1002/brb3.1009] [Medline: 29858545]
32. Tunbridge EM, Narajos M, Harrison CH, Beresford C, Cipriani A, Harrison PJ. Which dopamine polymorphisms are functional? Systematic review and meta-analysis of COMT, DAT, DBH, DDC, DRD1-5, MAOA, MAOB, TH, VMAT1, and VMAT2. Biol Psychiatry 2019 Oct;86(8):608-620 [FREE Full text] [doi: 10.1016/j.biopsych.2019.05.014] [Medline: 31303260]
33. Gajewski PD, Falkenstein M, Hengstler JG, Golka K. Toxoplasma gondii impairs memory in infected seniors. Brain Behav Immun 2014 Feb;36:193-199. [doi: 10.1016/j.bbi.2013.11.019] [Medline: 24321215]
34. Ortiz-Guerrero G, Gonzalez-Reyes RE, de-la-Torre A, Medina-Rincón G, Nava-Mesa MO. Pathophysiological mechanisms of cognitive impairment and neurodegeneration by infection. Brain Sci 2020 Jun;10(6):369 [FREE Full text] [doi: 10.3390/brainsci10060369] [Medline: 32545619]
35. Alpert A, Pickman Y, Leipold M, Rosenberg-Hasson Y, Ji X, Gaujoux R, et al. A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. Nat Med 2019 Mar;25(3):487-495 [FREE Full text] [doi: 10.1038/s41591-019-0381-y] [Medline: 30842675]
36. Alvares Pereira G, Silva Nunes MV, Alzola P, Contador I. Cognitive reserve and brain maintenance in aging and dementia: an integrative review. Appl Neuropsychol Adult 2021 Jan:1-11. [doi: 10.1080/23279095.2021.1872079] [Medline: 33492168]
37. Hultsch D, Strauss E, Hunter M, MacDonald S. Intraindividual variability, cognition, and aging. In: Craik FI, Salthouse TA, editors. The Handbook of Aging and Cognition. New York: Psychology Press; 2008:491-556.
38. Stern Y. Cognitive reserve. Neuropsychologia 2009 Aug;47(10):2015-2028 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2009.03.004] [Medline: 19467352]
39. Parade SH, Huffhines L, Daniels TE, Stroud LR, Nugent NR, Tyrka AR. A systematic review of childhood maltreatment and DNA methylation: candidate gene and epigenome-wide approaches. Transl Psychiatry 2021 Feb;11(1):134 [FREE Full text] [doi: 10.1038/s41398-021-01207-y] [Medline: 33608499]
40. Grippo AJ, Johnson AK. Stress, depression and cardiovascular dysregulation: a review of neurobiological mechanisms and the integration of research from preclinical disease models. Stress 2009 Jan;12(1):1-21 [FREE Full text] [doi: 10.1080/10253890802046281] [Medline: 19116888]
41. Ilmarinen J. Towards a longer worklife! : ageing and the quality of worklife in the European Union. In: Finnish Institute of Occupational Health. Helsinki: Finnish Institute of Occupational Health, Ministry of Social Affairs and Health; 2005.
42. Tuomi K, Huuhtanen P, Nykyri E, Ilmarinen J. Promotion of work ability, the quality of work and retirement. Occup Med (Lond) 2001 Aug;51(5):318-324. [doi: 10.1093/occmed/51.5.318] [Medline: 11473138]
43. Tuomi K, Ilmarinen J, Jahkola A, Katajarinne L, Tulkki A. Work Ability Index, 2nd Revised Edition. Helsinki: Finnish Institute of Occupational Health; 1998.
44. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/BF03193146] [Medline: 17695343]
45. Statistisches Bundesamt (German Federal Statistical Office). DESTATIS. 2020 Dec. URL: https://www-genesis.destatis.de/genesis/online [accessed 2021-08-16]
46. Equilibrium Hardy-Weinberg. WpCalc. URL: https://wpcalc.com/en/equilibrium-hardy-weinberg/ [accessed 2021-10-15]
47. Votruba KL, Persad C, Giordani B. Cognitive deficits in healthy elderly population with "normal" scores on the mini-mental state examination. J Geriatr Psychiatry Neurol 2016 May;29(3):126-132. [doi: 10.1177/0891988716629858] [Medline: 26850856]
48. Hautzinger M, Keller F, Kühner C. Das Beck Depressionsinventar II. In: Deutsche Bearbeitung und Handbuch zum BDI II. Frankfurt am Main: Harcourt Test Services; 2006.
49. Satzger W, Fessmann H, Engel RR. Liefern HAWIE-R, WST und MWT-B vergleichbare IQ-Werte? Zeitschrift für Differentielle und Diagnostische Psychologie 2002 Jun;23(2):159-170. [doi: 10.1024//0170-1789.23.2.159]
50. Wechsler D. WAIS-IIIR Manual. New York: The Psychological Corporation; 1997.
51. Tewes U. Hamburg-Wechsler-Intelligenztest für Erwachsene (HAWIE-R) (2. Ed.). Bern: Huber; 1994.
52. Oswald W, Fleischmann U. Nürnberger Alters Inventar (NAI) (4. Ed.). Göttingen: Hogrefe Verlag; 1997.
53. Helmstaedter C, Lendt M, Lux S. Verbaler Lern- und Merkfähigkeitstest. Göttingen: Hogrefe; 2001.
54. Somerville J, Tremont G, Stern RA. The Boston Qualitative Scoring System as a measure of executive functioning in Rey-Osterrieth Complex Figure Performance. J Clin Exp Neuropsychol 2010 Aug;22(5):613-621. [doi: 10.1076/1380-3395(200010)22:5;1-9;FT613]
55. Brickenkamp R, Schmidt-Atzert L, Liepmann D. d2-R: Test d2 - Revision. Göttingen: Hogrefe; 2010.
56. Sturm W, Wilmes K, Horn W. LPS 50+ Leistungsprüfsystem für 50 bis 90-Jährige (2nd Ed.). Göttingen: Hogrefe; 2015.

57. Beck A, Steer R, Brown G. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation; 1996.
58. Costa JP, McCrae R. Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Investory (NEO-FFI) Professional Manual. Odessa: Psychological Assesment Ressources, Inc; 1992.
59. Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. Child Abuse Negl 2003 Feb;27(2):169-190. [doi: 10.1016/S0145-2134(02)00541-0]
60. Griefahn B, Kunemund C, Brode P, Mehnert P. Zur Validitat der deutschen Ubersetzung des Morningness-Eveningness-Questionnaires von Horne und Ostberg. The validity of a German version of the Morningness-Eveningness-Questionnaire developed by Horne and Ostberg. Somnologie 2001 May;5(2):71-80. [doi: 10.1046/j.1439-054x.2001.01149.x]
61. Broadbent D, Cooper P, FitzGerald P, Parkes K. The Cognitive Failures Questionnaire (CFQ) and its correlates. Br J Clin Psychol 1982 Feb;21(1):1-16. [doi: 10.1111/j.2044-8260.1982.tb01421.x] [Medline: 7126941]
62. Duckworth AL, Quinn PD. Development and validation of the short grit scale (grit-s). J Pers Assess 2009 Mar;91(2):166-174. [doi: 10.1080/00223890802634290] [Medline: 19205937]
63. Oldfield R. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 1971 Mar;9(1):97-113. [doi: 10.1016/0028-3932(71)90067-4]
64. Neubach B, Schmidt K. Selbstkontrolle als Arbeitsanforderung - Rekonzeptualisierung und Validierung eines Messinstruments. Zeitschrift für Arbeits- und Organisationspsychologie A&O 2006 Apr;50(2):103-109. [doi: 10.1026/0932-4089.50.2.103]
65. Jackson PR, Wall TD, Martin R, Davids K. New measures of job control, cognitive demand, and production responsibility. J Appl Psychol 1993 Oct;78(5):753-762. [doi: 10.1037/0021-9010.78.5.753]
66. Höltke V, Jakob E. Lüdenscheider Aktivitätsfragebogen zum Risikofaktor Bewegungsmangel. Lüdenscheid: Sportmedizin Hellersen; 2002.
67. Schutte N, Toppinen S, Kalimo R, Schaufeli W. The factorial validity of the Maslach Burnout Inventory-General Survey (MBI-GS) across occupational groups and nations. J Occup Org Psychol 2000;73:53-66. [doi: 10.1348/096317900166877]
68. Demerouti E, Bakker AB, Nachreiner F, Schaufeli WB. The job demands-resources model of burnout. J Appl Psychol 2001 Jun;86(3):499-512. [doi: 10.1037/0021-9010.86.3.499]
69. Schlotz W, Yim IS, Zoccola PM, Jansen L, Schulz P. The Perceived Stress Reactivity Scale: measurement invariance, stability, and validity in three countries. Psychol Assess 2011 Mar;23(1):80-94. [doi: 10.1037/a0021148] [Medline: 21280954]
70. Crawford JR, Henry JD. The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. Br J Clin Psychol 2004 Sep;43(3):245-265. [doi: 10.1348/0144665031752934] [Medline: 15333231]
71. Fliege H, Rose M, Arck P, Levenstein S, Klapp BF. Validierung des “Perceived Stress Questionnaire“ (PSQ) an einer deutschen Stichprobe. Diagnostica 2001 Jul;47(3):142-152. [doi: 10.1026//0012-1924.47.3.142]
72. Neubach B, Schmidt KH. Entwicklung und Validierung von Skalen zur Erfassung verschiedener Selbstkontrollanforderungen bei der Arbeit Development and validation of scales assessing different job-related self-control demands. Zeitschrift für Arbeitswissenschaften 2007;61:35-45. [doi: 10.1037/t76218-000]
73. Schulz P, Schlotz W, Becker P. Trierer Inventar zum chronischen Stress (TICS). Göttingen: Hogrefe; 2004.
74. World Health Organization. WHOQOL–BREF: introduction, administration, scoring and generic version of the assessment. Field Trial Version. In: Rapley M, editor. Quality of Life Research. A Critical Introduction. London: Sage Publications; 1996:229-248.
75. Ilmarinen J. The Work Ability Index (WAI). Occup Med 2006 Oct;57(2):160. [doi: 10.1093/occmed/kqm008]
76. Burr H, Berthelsen H, Moncada S, Nübling M, Dupret E, Demiral Y, international COPSOQ Network. The third version of the Copenhagen Psychosocial Questionnaire. Saf Health Work 2019 Dec;10(4):482-503 [FREE Full text] [doi: 10.1016/j.shaw.2019.10.002] [Medline: 31890332]
77. König M, Haensel C, Jaschinski W. How to place the computer monitor: measurements of vertical zones of clear vision with presbyopic corrections. Clin Exp Optom 2015 May;98(3):244-253 [FREE Full text] [doi: 10.1111/cxo.12274] [Medline: 25963115]
78. Desrosiers J, Hébert R, Bravo G, Dutil E. The Purdue Pegboard Test: normative data for people aged 60 and over. Disabil Rehabil 1995 Jul;17(5):217-224. [doi: 10.3109/09638289509166638] [Medline: 7626768]
79. Lehrl S. Mehrfach-Wortwahl-Test (MWT). Erlangen: Medizinische Verlagsgesellschaft; 1995.
80. Folstein M, Folstein S, 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]
81. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935;18(6):643-662. [doi: 10.1037/h0054651]
82. Reitan R. Trail Making Test: Manual for Administration and Scoring. Tucson, AZ: Reitan Neuropsychology Laboratory; 1992.

83. Gajewski PD, Hanisch E, Falkenstein M, Thönes S, Wascher E. What does the n-back task measure as we get older? Relations between working-memory measures and other cognitive functions across the lifespan. Front Psychol 2018 Nov;9:2208 [FREE Full text] [doi: 10.3389/fpsyg.2018.02208] [Medline: 30534095]
84. Raven J, Raven J. Raven progressive matrices. In: McCallum RS, editor. Handbook of Nonverbal Assessment. Boston: Springer; 2003:223-237.
85. Oswald W. Zahlen-Verbindungs-Test ZVT (3rd Ed.). Göttingen: Hogrefe; 2016.
86. Deary I. Intelligence: A Very Short Introduction. Oxford: Oxford University Press; 2001.
87. Schneider D, Wascher E. Mechanisms of target localization in visual change detection: an interplay of gating and filtering. Behav Brain Res 2013 Nov;256:311-319. [doi: 10.1016/j.bbr.2013.08.046] [Medline: 24001756]
88. Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behav Res Meth Instrum Comput 1985 Nov;17(6):652-655. [doi: 10.3758/BF03200977]
89. Simon JR. Reactions toward the source of stimulation. J Exp Psychol 1969 Jul;81(1):174-176. [doi: 10.1037/h0027448] [Medline: 5812172]
90. Hoppe K, Küper K, Wascher E. Sequential modulations in a combined horizontal and vertical Simon task: Is there ERP evidence for feature integration effects? Front Psychol 2017 Jun;8:1094 [FREE Full text] [doi: 10.3389/fpsyg.2017.01094] [Medline: 28713305]
91. Cohen JD, Barch DM, Carter C, Servan-Schreiber D. Context-processing deficits in schizophrenia: converging evidence from three theoretically motivated cognitive tasks. J Abnorm Psychol 1999 Feb;108(1):120-133. [doi: 10.1037/0021-843x.108.1.120]
92. Getzmann S, Wascher E. Visually guided auditory attention in a dynamic "cocktail-party" speech perception task: ERP evidence for age-related differences. Hear Res 2017 Feb;344:98-108. [doi: 10.1016/j.heares.2016.11.001] [Medline: 27825858]
93. Kirchner WK. Age differences in short-term retention of rapidly changing information. J Exp Psychol 1958 Apr;55(4):352-358. [doi: 10.1037/h0043688] [Medline: 13539317]
94. Rogers RD, Monsell S. Costs of a predictible switch between simple cognitive tasks. J Exp Psychol Gen 1995 Jun;124(2):207-231. [doi: 10.1037/0096-3445.124.2.207]
95. Gajewski PD, Ferdinand NK, Kray J, Falkenstein M. Understanding sources of adult age differences in task switching: evidence from behavioral and ERP studies. Neurosci Biobehav Rev 2018 Sep;92:255-275 [FREE Full text] [doi: 10.1016/j.neubiorev.2018.05.029] [Medline: 29885425]
96. Schröger E, Wolff C. Behavioral and electrophysiological effects of task-irrelevant sound change: a new distraction paradigm. Cogn Brain Res 1998 Jul;7(1):71-87. [doi: 10.1016/s0926-6410(98)00013-5]
97. Getzmann S, Falkenstein M, Gajewski PD. Neuro-behavioral correlates of post-deviance distraction in middle-aged and old adults. J Psychophysiol 2014 Sep;28(3):178-186. [doi: 10.1027/0269-8803/a000124]
98. Gajewski PD, Falkenstein M, Thönes S, Wascher E. Stroop task performance across the lifespan: high cognitive reserve in older age is associated with enhanced proactive and reactive interference control. Neuroimage 2020 Feb;207:116430 [FREE Full text] [doi: 10.1016/j.neuroimage.2019.116430] [Medline: 31805383]
99. Karlin L, Martz MJ, Mordkoff AM. Motor performance and sensory-evoked potentials. Electroencephalogr Clin Neurophysiol 1970 Mar;28(3):307-313. [doi: 10.1016/0013-4694(70)90167-7]
100. Gajewski PD, Falkenstein M. Effects of task complexity on ERP components in Go/Nogo tasks. Int J Psychophysiol 2013 Mar;87(3):273-278. [doi: 10.1016/j.ijpsycho.2012.08.007] [Medline: 22906814]
101. Treisman AM, Gelade G. A feature-integration theory of attention. Cogn Psychol 1980 Jan;12(1):97-136. [doi: 10.1016/0010-0285(80)90005-5]
102. Gajewski PD, Hengstler JG, Golka K, Falkenstein M, Beste C. The functional tumor necrosis factor-α (308A/G) polymorphism modulates attentional selection in elderly individuals. Neurobiol Aging 2013 Nov;34(11):2694.e1-2694.e12. [doi: 10.1016/j.neurobiolaging.2013.04.017] [Medline: 23673311]
103. Claus M, Dychus N, Ebel M, Damaschke J, Maydych V, Wolf OT, et al. Measuring the immune system: a comprehensive approach for the analysis of immune functions in humans. Arch Toxicol 2016 Oct;90(10):2481-2495. [doi: 10.1007/s00204-016-1809-5] [Medline: 27522652]
104. Stalder T, Kirschbaum C. Analysis of cortisol in hair–state of the art and future directions. Brain Behav Immun 2012 Oct;26(7):1019-1029. [doi: 10.1016/j.bbi.2012.02.002] [Medline: 22366690]
105. Lehmann M, Selinski S, Blaszkewicz M, Orlich M, Ovsiannikov D, Moormann O, et al. Rs710521[A] on chromosome 3q28 close to TP63 is associated with increased urinary bladder cancer risk. Arch Toxicol 2010 Dec;84(12):967-978. [doi: 10.1007/s00204-010-0617-6] [Medline: 21063684]
106. Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data 2016 Jun;3(1):160044 [FREE Full text] [doi: 10.1038/sdata.2016.44] [Medline: 27326542]
107. Pernet CR, Appelhoff S, Gorgolewski KJ, Flandin G, Phillips C, Delorme A, et al. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Sci Data 2019 Jun;6(1):103 [FREE Full text] [doi: 10.1038/s41597-019-0104-8] [Medline: 31239435]

108. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004 Mar;134(1):9-21. [doi: 10.1016/j.jneumeth.2003.10.009] [Medline: 15102499]
109. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986 Apr;73(1):13-22. [doi: 10.1093/biomet/73.1.13]
110. Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982 Dec;38(4):963-974. [doi: 10.2307/2529876]
111. Bierbrauer A, Kunz L, Gomes CA, Luhmann M, Deuker L, Getzmann S, et al. Unmasking selective path integration deficits in Alzheimer's disease risk carriers. Sci Adv 2020 Aug;6(35):eaba1394 [FREE Full text] [doi: 10.1126/sciadv.aba1394] [Medline: 32923622]
112. Thönes S, Wascher E, Gajewski P, Getzmann S. Time Hurries on but Does not Fly in Older Age — No Effect of Depressive Symptoms. Timing Time Percept 2021 Jan 13;9(3):241-256. [doi: 10.1163/22134468-bja10007]
113. Sharifian F, Schneider D, Arnau S, Wascher E. Decoding of cognitive processes involved in the continuous performance task. Int J Psychophysiol 2021 Sep;167:57-68 [FREE Full text] [doi: 10.1016/j.ijpsycho.2021.06.012] [Medline: 34216693]
114. Getzmann S, Digutsch J, Kleinsorge T. COVID-19 Pandemic and Personality: Agreeable People Are More Stressed by the Feeling of Missing. Int J Environ Res Public Health 2021 Oct 13;18(20) [FREE Full text] [doi: 10.3390/ijerph182010759] [Medline: 34682500]
115. Metzen D, Genç E, Getzmann S, Larra MF, Wascher E, Ocklenburg S. Frontal and parietal EEG alpha asymmetry: a large-scale investigation of short-term reliability on distinct EEG systems. Brain Struct Funct 2022 Mar;227(2):725-740 [FREE Full text] [doi: 10.1007/s00429-021-02399-1] [Medline: 34676455]
116. Getzmann S, Arnau S, Gajewski PD, Wascher E. When long appears short: Effects of auditory distraction on event-related potential correlates of time perception. Eur J Neurosci 2022 Jan 14;55(1):121-137. [doi: 10.1111/ejn.15553] [Medline: 34859527]
117. National Research Council (US) Panel on Handling Missing Data in Clinical Trials. The Prevention and Treatment of Missing Data in Clinical Trials. Washington (DC): National Academies Press (US); 2010.
118. Garcia TP, Marder K. Statistical approaches to longitudinal data analysis in neurodegenerative diseases: Huntington's disease as a model. Curr Neurol Neurosci Rep 2017 Feb;17(2):14 [FREE Full text] [doi: 10.1007/s11910-017-0723-4] [Medline: 28229396]

Abbreviations

ANCOVA: Analyses of Covariance
ANOVA: analyses of variance
AX-CPT: AX-continuous performance task
BDNF: brain-derived neurotrophic factor
CMS: Common Mode Sense
COMT: catechol-O-methyltransferase
DRL: driven right leg
ECG: electrocardiography
EEG: electroencephalography
HAWIE-R: Hamburg-Wechsler-Intelligenztest für Erwachsene, Revision
HWE: Hardy-Weinberg equilibrium
GEE: Generalized Estimating Equation
IFN: interferon
IL: interleukin
LPS: Leistungs-Prüf-System
MER: Mixed Effect Regression
MMSE: Mini Mental State Examination
MRI: magnetic resonance imaging
MWT-B: Multiple-Choice Vocabulary Test
NAS: Network Attached Storage
PGS: polygenic scores
RAVLT: Rey Auditory Verbal Learning Test
SNP: single nucleotide polymorphism
TNF: tumor necrosis factor
VLMT: Verbal Learning and Memory Test
WAI: Work Ability Index
WAIS-III: Wechsler Adult Intelligence Scale
WM: working memory

Edited by G Eysenbach; submitted 05.10.21; peer-reviewed by S Ballesteros, P Heyn; comments to author 10.11.21; revised version received 14.12.21; accepted 15.12.21; published 14.03.22.

Please cite as:
Gajewski PD, Getzmann S, Bröde P, Burke M, Cadenas C, Capellino S, Claus M, Genç E, Golka K, Hengstler JG, Kleinsorge T, Marchan R, Nitsche MA, Reinders J, van Thriel C, Watzl C, Wascher E
Impact of Biological and Lifestyle Factors on Cognitive Aging and Work Ability in the Dortmund Vital Study: Protocol of an Interdisciplinary, Cross-sectional, and Longitudinal Study
JMIR Res Protoc 2022;11(3):e32352
URL: https://www.researchprotocols.org/2022/3/e32352
doi:10.2196/32352
PMID:35285810

Protocol

The Physical Activity and Fitness in Childhood Cancer Survivors (PACCS) Study: Protocol for an International Mixed Methods Study

Hanne C Lie[1], PhD; Sigmund Anderssen[2], PhD; Corina Silvia Rueegg[3], PhD; Truls Raastad[4], PhD; May Grydeland[4], PhD; Lene Thorsen[5,6], PhD; Trine Stensrud[2], PhD; Elisabeth Edvardsen[4], PhD; Marie Hamilton Larsen[1], PhD; Ingrid Kristin Torsvik[7], MD, PhD; Lars Peder Bovim[8], PhD; Miriam Götte[9], PhD; Päivi Maria Lähteenmäki[10], MD, PhD; Susi Kriemler[11], PhD; Hanne Bækgaard Larsen[12,13], PhD; Martin Kaj Fridh[12], PhD; Kristin Ørstavik[14], MD, PhD; Henrik Brun[15], MD, PhD; Iren Matthews[16], MD, PhD; Else Hornset[17]; Ellen Ruud[18,19], MD, PhD

[1]Department of Behavioural Medicine, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway

[2]Department of Sports Medicine, Norwegian School of Sport Sciences, Oslo, Norway

[3]Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway

[4]Department of Physical Performance, Norwegian School of Sport Sciences, Oslo, Norway

[5]National Advisory Unit on Late Effects after Cancer Treatment, Department of Oncology, Division of Cancer Medicine, Oslo University Hospital, Oslo, Norway

[6]Department of Clinical Service, Division of Cancer Medicine, Oslo University Hospital, Oslo, Norway

[7]Department of Pediatrics, Haukeland University Hospital, Bergen, Norway

[8]Department of Health and Functioning, Faculty of Health and Social Sciences, Western Norway University of Applied Sciences, Bergen, Norway

[9]Department of Pediatric Hematology/Oncology, Clinic for Pediatrics III, West German Cancer Center, University Hospital Essen, Essen, Germany

[10]Department of Pediatric and Adolescent Hematology/Oncology, Turku University Hospital, University of Turku, Turku, Finland

[11]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland

[12]Department of Pediatrics and Adolescent Medicine, Copenhagen University Hospital (Rigshospitalet), Copenhagen, Denmark

[13]Institute for Clinical Medicine, Faculty of Health Science, The University of Copenhagen, Copenhagen, Denmark

[14]Department of Neurology, Section for Clinical Neurophysiology, Oslo University Hospital, Oslo, Norway

[15]Department of Pediatric Cardiology, Oslo University Hospital, Oslo, Norway

[16]Department of Paediatric Allergy and Pulmonology, Oslo University Hospital, Oslo, Norway

[17]Norwegian Childhood Cancer Society, Oslo, Norway

[18]Department of Pediatric Hematology and Oncology, Oslo University Hospital, Oslo, Norway

[19]Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding Author:
Hanne C Lie, PhD
Department of Behavioural Medicine, Institute of Basic Medical Sciences, Faculty of Medicine
University of Oslo
P.B. 1111 Blindern
Oslo, 0317
Norway
Phone: 47 41360470
Email: h.c.lie@medisin.uio.no

Abstract

Background: Survivors of childhood cancer represent a growing population with a long life expectancy but high risks of treatment-induced morbidity and premature mortality. Regular physical activity (PA) may improve their long-term health; however, high-quality empirical knowledge is sparse.

Objective: The Physical Activity and Fitness in Childhood Cancer Survivors (PACCS) study comprises 4 work packages (WPs) aiming for the objective determination of PA and self-reported health behavior, fatigue, and quality of life (WP 1); physical fitness determination (WP 2); the evaluation of barriers to and facilitators of PA (WP 1 and 3); and the feasibility testing of an intervention to increase PA and physical fitness (WP 4).

Methods: The PACCS study will use a mixed methods design, combining patient-reported outcome measures and objective clinical and physiological assessments with qualitative data gathering methods. A total of 500 survivors of childhood cancer aged 9 to 18 years with ≥1 year after treatment completion will be recruited in follow-up care clinics in Norway, Denmark, Finland, Germany, and Switzerland. All participants will participate in WP 1, of which approximately 150, 40, and 30 will be recruited to WP 2, WP3, and WP 4, respectively. The reference material for WP 1 is available from existing studies, whereas WP 2 will recruit healthy controls. PA levels will be measured using ActiGraph accelerometers and self-reports. Validated questionnaires will be used to assess health behaviors, fatigue, and quality of life. Physical fitness will be measured by a cardiopulmonary exercise test, isometric muscle strength tests, and muscle power and endurance tests. Limiting factors will be identified via neurological, pulmonary, and cardiac evaluations and the assessment of body composition and muscle size. Semistructured, qualitative interviews, analyzed using systematic text condensation, will identify the perceived barriers to and facilitators of PA for survivors of childhood cancer. In WP 4, we will evaluate the feasibility of a 6-month personalized PA intervention with the involvement of local structures.

Results: Ethical approvals have been secured at all participating sites (Norwegian Regional Committee for Medical Research Ethics [2016/953 and 2018/739]; the Oslo University Hospital Data Protection Officer; equivalent institutions in Finland, Denmark [file H-19032270], Germany, and Switzerland [Ethics Committee of Northwestern and Central Switzerland, project ID: 2019-00410]). Data collection for WP 1 to 3 is complete. This will be completed by July 2022 for WP 4. Several publications are already in preparation, and 2 have been published.

Conclusions: The PACCS study will generate high-quality knowledge that will contribute to the development of an evidence-based PA intervention for young survivors of childhood cancer to improve their long-term care and health. We will identify physiological, psychological, and social barriers to PA that can be targeted in interventions with immediate benefits for young survivors of childhood cancer in need of rehabilitation.

International Registered Report Identifier (IRRID): DERR1-10.2196/35838

(JMIR Res Protoc 2022;11(3):e35838) doi:10.2196/35838

KEYWORDS

childhood cancer survivor; physical activity; physical fitness; barriers; intervention; quality of life; fatigue

Introduction

Background

Improved diagnostics and treatment of pediatric cancer over the past decades have drastically increased 5-year survival rates to >80% [1]. Therefore, the population of survivors is rapidly growing, and there are currently an estimated ≥500,000 survivors of childhood cancer in Europe [2]. Survivors of childhood cancer represent a vulnerable population, which is expected to face high, long-term risks of morbidity [3,4] and early mortality because of the late effects of cancer and its treatment [5-7]. Late effects can arise years or decades beyond treatment completion; by the age of 50 years, approximately all survivors of childhood cancer experience multiple chronic health conditions [8]. Late effects include, among others, cardiorespiratory disorders, endocrine dysfunction, secondary cancers, musculoskeletal deficits, metabolic syndrome, early frailty, neurocognitive impairments, and fatigue, resulting in reduced quality of life (QoL) [3,4,8]. Increased knowledge of the negative consequences of treatment has prompted efforts to improve the long-term QoL and health outcomes of survivors of childhood cancer.

There is increasing evidence for a range of positive effects of regular physical activity (PA) on reducing risks of late effects, early mortality, and improving QoL in both adult [9-16] and young survivors of childhood cancer [10,16,17]. However, reduced levels of PA and physical fitness (PF) in survivors of childhood cancer compared with controls have been consistently reported, albeit with substantial heterogeneity across studies [9]. Objectively measured PA levels in childhood and early adolescent survivors of childhood cancer support these findings, suggesting considerably reduced PA levels compared with controls in the small-scale studies that exist [18]. Moreover, studies among adult survivors of childhood cancer suggest that reduced PA levels persist for a substantial proportion of survivors well into adulthood [19,20]. Thus, there is a need for effective PA interventions among young survivors of childhood cancer to establish healthy PA routines at an early age. A modest body of research indicates the effectiveness of PA interventions on PA behavior and physical, psychosocial, cognitive, and social outcomes for groups of young survivors of childhood cancer [10,16,17]; however, PA recommendations are currently limited by important knowledge gaps and methodological shortcomings, including low sample sizes, single-center samples, and large heterogeneity across studies [9,10,16,17]. Data from large-scale studies of survivors of childhood cancer representing multiple diagnostic groups and countries, using device-based measurements, are still lacking.

Knowledge gaps regarding PF in young survivors of childhood cancer are even more prevalent than for PA [21]; however, findings indicate that PF is also persistently poorer among young survivors of childhood cancer with reduced cardiorespiratory fitness (9%-23%) [22,23] and muscle strength (10%-20%) [9,24] than the age-matched controls. Interestingly, the difference in PF between survivors of childhood cancer and healthy controls seems to progress years after treatment [22]. Research within our group suggests that Norwegian adult survivors of childhood leukemia and lymphoma experience reduced exercise capacity because of the cardiotoxic effect of treatment [25] and lower

QoL compared with cancer-free controls even >20 years after diagnosis [26]. Thus, low levels of PA and PF may exacerbate already increased risk of late effects, functional limitations, and reduced QoL for survivors of childhood cancer well into adulthood [27,28].

Research on young survivors' perspectives on what facilitates and hinders participation in PA is also scarce [29,30]. The observed lower levels of PA among young survivors of childhood cancer than in the general population are often attributed to a range of physical and psychosocial factors, including (1) adverse consequences of cancer therapies on PF and function [17]; (2) anxiety and overprotective attitudes toward PA [31]; and (3) the cancer and treatment may coincide with the period in life when children start organized sports, thus contributing to an ability gap to healthy peers [22]. Late effects and PF deficits interfere with physical, psychological, and social functioning and the ability to participate in daily and social activities in approximately 50% of adult survivors of childhood cancer [32]. Therefore, limitations to engaging in PA can create a *vicious circle* of reduced activity levels, which further negatively influence long-term development, health, and functioning. Therefore, knowledge of perceived barriers and facilitators for survivors of childhood cancer regarding PA and exercise programs is needed to develop effective and sustainable PA interventions.

Objectives

The overall aim of the Physical Activity and Fitness in Childhood Cancer Survivors (PACCS) study is to address the abovementioned research gaps and generate knowledge that will contribute to the development of an evidence-based PA intervention for young survivors of childhood cancer. The PACCS study is run by an international multidisciplinary consortium, including Norway, Finland, Denmark, Germany, and Switzerland, and includes four work packages (WPs) with the following main objectives:

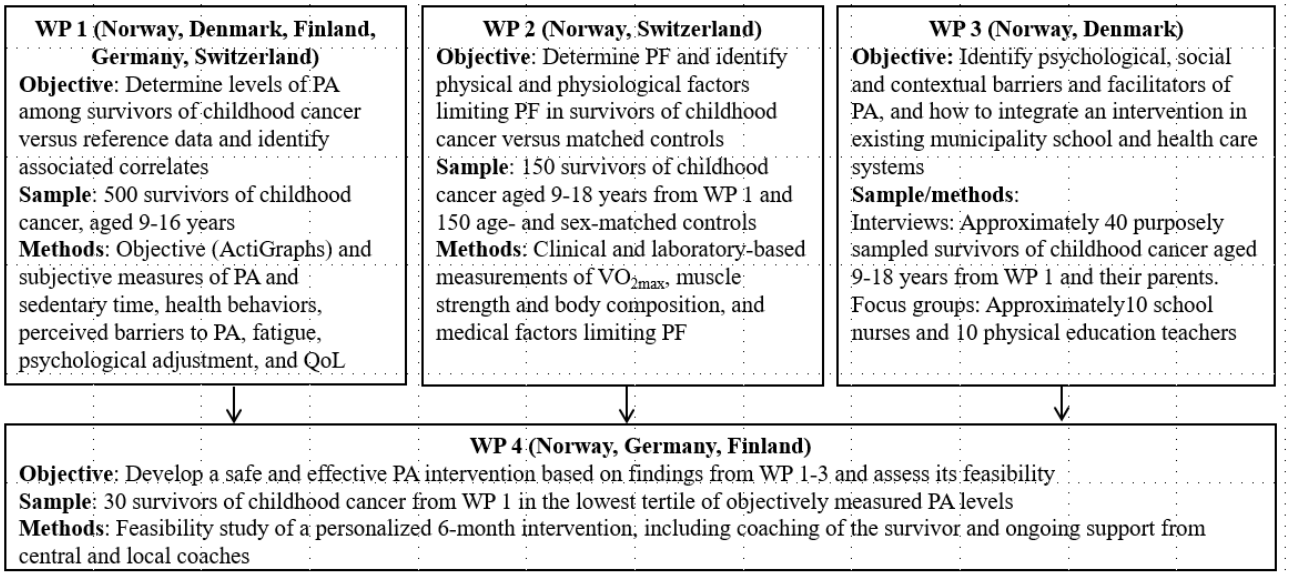
1. Objectively determine levels of PA and sedentary time among survivors of childhood cancer compared with reference data and identify physiological, psychological, and social correlates, including QoL and fatigue (WP 1)
2. Objectively determine PF and identify physical and physiological factors limiting PF in survivors of childhood cancer compared with controls (WP 2)
3. Identify physical, psychological, social, and contextual factors that act as barriers to and facilitators for engagement in PA in survivors of childhood cancer; furthermore, to identify opportunities for collaboration between specialist and municipality health care services to establish systematic PA interventions locally for survivors of childhood cancer (WP 3)
4. Develop and feasibility test a PA intervention as a preparation for a future well-designed large randomized controlled trial based on results from WP 1 to 3 (WP 4)

Methods

Study Design, Study Population, and Inclusion Criteria

The PACCS study [33] is an international multicenter study based on a mixed methods design that combines clinical and physiological assessment methods, patient-reported outcome measures, qualitative interviews, and a feasibility intervention. WPs 1, 2, and 3 have a cross-sectional design, whereas WP 4 will be a single-arm (intervention only), multicenter, open PA feasibility trial. The study population comprises young survivors of childhood cancer (aged 9-18 years) recruited from seven medical centers in Europe: 2 in Norway (Oslo and Bergen) and Finland (Turku and Tampere) and 1 in Denmark (Copenhagen), Germany (Essen), and Switzerland (Basel). Survivors of childhood cancer with *any* malignant disease, at least 1 year after completion of cancer treatment, and those attending outpatient follow-up care consultations are eligible for inclusion. Survivors of childhood cancer will be excluded if there are language difficulties or limited cognitive functioning. Figure 1 provides an overview of the different WPs.

Figure 1. Overview of the objectives and methods for the 4 work packages (WPs) of the Physical Activity and Fitness in Childhood Cancer Survivors study.



We aimed to include a minimum of 500 survivors of childhood cancer aged 9 to 16 years in WP 1. We aimed to consecutively enroll approximately 150 participants from WP 1 to WP 2. The age group in WP 2 will be 9 to 18 years, as 1 to 2 years may, in some sites, elapse from inclusion into WP 1 to participation in WP 2. In addition, in WP 2, 150 age- and sex-matched controls will be enrolled. A subsample of 40 selected WP 1 participants and one of their parents will be invited to participate in qualitative interviews in WP 3. Participants will be purposely sampled to maximize the variation of key sociodemographic and clinical factors, including activity levels and rural or urban places of residence. Approximately 10 school nurses and 10 physical education teachers will also be recruited. In WP 4, around 30 survivors of childhood cancer aged 12 to 18 years within the lowest tertial of total PA, as assessed by the accelerometer in WP 1, without contraindications for vigorous PA, will be recruited among the participants of WP 1. As WP 4 is a pilot feasibility study, we aim to recruit survivors of different diagnoses, age ranges, and genders. All countries and study sites will participate in WP 1; Norway and Switzerland will participate in WP 2; Norway and Denmark will participate in WP 3; and Norway, Finland, and Germany will participate in WP 4.

Assessments, Power Calculations, and Statistical Analyses

WP 1

Overview

Outcomes of WP 1 include objective and subjective assessments of PA and sedentary behavior, barriers to PA, fatigue, psychological adjustment, and QoL. The self-reported measures, as well as demographic variables, will be collected using electronic questionnaires. An overview of the key variables assessed in WP 1 is presented in Table 1.

Table 1. Overview of factors assessed at each study site for WPs[a] 1, 2, and 4.

Variables	WP 1; all sites	WP 2			WP 4; Oslo, Bergen, Turku, and Essen
		Oslo	Bergen	Basel	
Objectively assessed					
Physical activity (counts per minute; sedentary time; light, moderate, and vigorous activity)[b]	✓	✓	✓	✓	✓
Echocardiography		✓	✓		
Lung function		✓	✓	✓	
Neurological tests		✓	✓		
Cardiopulmonary Exercise Testing		✓[c]	✓[c]	✓[d]	✓[c]
Knee extension strength		✓	✓	✓	✓[e]
Hand grip		✓	✓	✓	
Chest press strength		✓	✓	✓	✓[f]
Countermovement jump		✓	✓	✓	
1-Minute sit-to-stand test		✓	✓	✓	✓
Body composition (DXA[g])		✓	✓	✓	
Muscle thickness (ultrasound)		✓	✓		
Blood volume and hemoglobin mass		✓			
Anthropometric variables (height and weight)	✓	✓	✓	✓	✓
Puberty stage (Tanner)		✓	✓	✓	
Medical variables (diagnosis, time since diagnosis, treatment, time since treatment, medication use, and comorbidities)	✓	✓	✓	✓	
Questionnaires					
Demographic variables (age and gender)[h,i]	✓	✓	✓	✓	✓
Self-reported PA[j,k], including recreational sports participation and active commuting habits[h,i]	✓				
Physical education[l]	✓				
Perceived PA barriers: 15 items covering reasons for not being physical active[m]	✓				✓
Attitude toward PA (perceived physical competence, perceived enjoyment, and motivation)					✓
Television or screen time[n]	✓				
Diet, including sugar rich beverages[h,o]	✓				
Sleep[i]	✓				
Quality of life (PedsQL[p])	✓				✓
Fatigue (PedsQL MDF[q] scale)	✓				✓
Psychological adjustment (SDQ[r])	✓				

[a]WP: work package.

[b]For WP 1, 2, and 4: ActiGraph model GT3X+, ActiGraph LLC; for WP 4: real-time tracking of PA over 6 months (time spent in activity, type of activity, and intensity and steps) using a Polar watch (Polar Global).

[c]Cardiopulmonary exercise testing on a treadmill.

[d]Cardiopulmonary exercise testing on a bike.

[e]Isometric knee extension strength is tested on the custom-built strength ergometer similar to WP 2 in Bergen and Oslo and on sites' own commercial devices in Turku and Essen.

[f]Upper arm strength is tested as isometric chest press on the custom-built strength ergometer similar to WP 2 in Bergen and Oslo and on sites' own commercial device in Essen. Turku performed a dynamic upper extremity lifting test in a standing position.

[g]DXA: dual-energy x-ray absorptiometry.

[h]Items from the Health in Adolescents study [34].

[i]Items from the Physical Activity in Norwegian Children Study [35].

[j]PA: physical activity.

[k]Number of counts per minute (cpm; primary outcome); fulfillment of physical activity (PA) recommendations (>60 minutes of moderate to vigorous PA per day); PA intensity categories (counts translated into metabolic energy equivalents of intensities): sedentary time (<100 cpm), light (100-1999 cpm), moderate (2000-5999 cpm), and vigorous (≥6000 cpm).

[l]Items from the Relevance of Physical Activity Contexts in the Everyday Life of Adolescents study [36].

[m]Constructed for the purpose of the Physical Activity and Fitness in Childhood Cancer Survivors study based on literature review and expert and user input.

[n]Items from the Health Behavior in School-aged Children study [37].

[o]Items from the UngKost [38].

[p]PedsQL: Pediatric Quality of Life Inventory [39].

[q]PedsQL MDF: Pediatric Quality of Life Inventory Multidimensional Fatigue scale [40].

[r]SDQ: Strength and Difficulties Questionnaire [41].

Assessment of PA and Sedentary Time by Accelerometer

PA and sedentary time will be assessed objectively by accelerometers worn for 7 days during the awake time to measure accelerations of +6 G to –6 G (ActiGraph model GT3X+, ActiGraph LLC). The sample rate of the accelerometer is set to measure the raw signals at 30 Hz. Total PA will be described as the number of counts per minute (cpm), which is a measure of the volume of PA assessed by the accelerometer. The accelerometer data will be uploaded by all collaborators to a central server located at the Norwegian School of Sport Sciences and analyzed at one site. Accelerometer measurements with at least four valid measurement days (≥480 min/day) will be included in the analysis.

The activity counts from the accelerometers will be translated into metabolic energy equivalents of intensities as sedentary time (<100 cpm), light (100-1999 cpm), moderate (2000-5999 cpm), and vigorous (>6000 cpm) PA [42]. Participants will be categorized as reaching the international recommendations for PA if they engage in at least 60 minutes of moderate to vigorous PA (≥2000 cpm) per day on average over valid accelerometer measurement days.

Questionnaires

Overview

All questionnaires will be completed electronically using tablet computers. Different questionnaires are available for survivors of childhood cancer aged 9 to 12 years and their parents (proxy reports) and survivors of childhood cancer aged 13 to 16 years and their parents. The questionnaires comprise mainly validated and standardized questions (Table 1). Questionnaires will be translated into local language if validated translations do not exist, following recommended forward–backward translation procedures. Questionnaire data will be directly uploaded and stored on the secured server for sensitive data at the University of Oslo's Center for Information Technology.

Self-reported PA and Exercise

Items pertaining to exercise in leisure time will be measured by questions extracted from studies of age- and sex-stratified reference material (the Health in Adolescents study, Physical Activity among Norwegian Children study wave 2 [35], and the Relevance of Physical Activity Contexts in the Everyday Life of Adolescents [REPAC]; the REPAC study described below [34-36]) to enable comparisons. The questions target PA and include frequency, duration, and type of PA or exercise and active transportation (questionnaires are available upon request).

Physical Education in School

Questions regarding attitudes and participation toward physical education before and after the cancer experience are based on the questionnaire of an ongoing study of approximately 3000 Norwegian adolescents (REPAC study) [36].

Perceived PA Barriers

Perceived barriers to being physically active will be assessed using a self-constructed questionnaire. Survivors of childhood cancer will rate degree of agreement with 15 statements in response to the question "I am less physically active than I want to because;...." The 15 statements reflect potential physical, psychosocial, and mental barriers to PA and were designed based on clinical experience, input from user representatives, and pilot testing. Examples of statements are "I don't have time for it; I don't know how active I can or should be; My body feels too heavy; I am too fatigued; I have pain; My parents tell me to be careful." Each statement is scored on a 5-point Likert-type scale ranging from *totally disagree* to *totally agree*.

Screen Time

Screen time (time spent in front of a computer, watching television, playing computer games, surfing on the internet, and chatting) will be assessed by 2 questions from a questionnaire used in a large international study on children and adolescents (the Health Behavior in School-aged Children study) [37].

Diet and Nutrition

A total of 3 questions regarding diet were extracted from a nationally representative study on children and adolescents' dietary habits, called the Norwegian Ungkost study [38].

Sleep

Participants will report when they usually get out of bed and when they go to bed on school days [43].

QoL Questionnaire

QoL will be measured using the Pediatric QoL Inventory (PedsQL) core module [39]. The PedsQL includes four subscales—physical, emotional, social, and school functioning—and has 2 different modules validated for children aged 8 to 12 years and adolescents aged 13 to 18 years. The PedsQL has been used in children with cancer [40]. In addition, parents will be asked to complete the parent module (proxy report) of the PedsQL.

Fatigue

Fatigue will be measured using the PedsQL Multidimensional Fatigue Scale, which measures general fatigue, sleep/rest fatigue, and cognitive fatigue [40]. Parents will complete the respective parent modules.

Psychological Adjustment

The Strength and Difficulties Questionnaire (SDQ) will be used to measure psychological adjustment and emotional problems. The SDQ is widely used in children aged 8 to 17 years [41].

Other Variables

Parents will provide sociodemographic variables in their questionnaire (child's age, child's sex, parents' education, and parents' income).

Information Extracted From Medical Records

The following variables will be collected from medical records and stored on Services for Sensitive Data (TSD): sex, age at diagnosis, cancer diagnosis, month and year of diagnosis, month and year when completed treatment, type of treatment (chemotherapy, surgery, radiation [dose—in Gy—and area], and stem cell transplantation), recurrence, diagnosis and treatment before or after puberty, weight, and height at the time of the study.

Reference Material

Age- and sex-stratified reference material for the primary and most of the secondary outcomes are available for comparison of marginal means and will be provided from several large-scale studies among healthy children and adolescents. References on PA and sedentary time will be available from two Norwegian studies—the Physical Activity among Norwegian Children study wave 2 (N=3538; age 9 and 15 years) [35], Health in Adolescents study (N=1528; age 11 and 13 years) [34]—and one German study—the German Health study (N=1500-2000; age 6 to 17 years) [44]. Age- and sex-stratified national reference materials are also available for attitudes toward physical education (REPAC) [36], screen time (Health Behavior in School-aged Children study) [37], and psychosocial measures (PedsQL [45], PedsQL Multidimensional Fatigue Scale [46], and SDQ [47]).

Power Calculation and Statistical Analyses WP 1

From studies in the general Norwegian population, we know that children aged 9 to 15 years engage in an average of 549 (SD 180) cpm (measured by accelerometer) [35]. We assumed reduced PA levels in survivors of childhood cancer. To estimate a difference of 10% in objectively measured PA between survivors of childhood cancer and healthy controls (494 vs 549 cpm, SD 180 for both; power=80%; significance level $P<.05$, one-sided), we need to include 250 survivors of childhood cancer. Given the heterogeneity in the expected sample (age and diagnoses) and to facilitate subgroup comparisons, we aim to include at least n=500 in WP 1 to facilitate subgroup comparisons.

Descriptive statistics will be used to present the levels of PA (including different intensity measures and activity patterns) and sedentary time of survivors of childhood cancer compared with the reference materials and PA recommendations. Multilevel models (mixed models) will be used to adjust for the different study sites as random intercepts. We will perform univariable and multivariable multilevel regression models (linear, logistic, or ordinal depending on the outcome) to (1) assess the clinical and sociodemographic factors associated with meeting the PA recommendations in survivors of childhood cancer; (2) assess clinical and sociodemographic factors associated with different patterns or levels of PA and sedentary time; and (3) assess the associations among PA patterns or levels (exposure) and QoL, fatigue, and other health behaviors.

WP 2

Overview

Measures

Outcomes of WP 2 include measurements of cardiorespiratory fitness, muscle strength, power, and muscular endurance. Norway (Oslo and Bergen) and Switzerland (Basel) will participate in WP 2 (Table 1 provides an overview of tests and self-reported measures performed at each site specific to WP 2). In Norway, age- and sex-matched controls will be the same-age friend or sibling of the survivor of childhood cancer, chosen and invited by the survivors of childhood cancer themselves. In Switzerland, age- and sex-matched controls will be recruited from siblings or friends of survivors or within the network of the Swiss PACCS study team.

Cardiorespiratory Fitness

Before testing, survivors of childhood cancer will undergo a cardiological screening. Peak oxygen uptake (VO_2peak) will be assessed using a cardiopulmonary exercise test (CPET). In Norway, the CPET will be performed by walking and running on a treadmill using a child-friendly continuous graded treadmill protocol until exhaustion [48], whereas in Switzerland, the CPET will be performed on a cycle ergometer using an incremental bike protocol from Godfrey [49]. The starting workload after habituation to the treadmill will be 3.0 km/hour at 0% inclination. The workload will then be increased every minute by alternately increasing the speed by 1 km/hour and the inclination by 2% every other minute. For the bike protocol, the initial workload will be set to 10, 15, or 20 W depending on the child's height and increased by 10, 15, or 20 W every

minute, respectively. The participants will be encouraged to keep a pedaling frequency of 60 to 90 revolutions per minute. Gas exchange and ventilatory variables will be measured continuously, breath by breath, using a Hans Rudolph 2-way breathing mask (Hans Rudolph). Peak heart rate will be measured through a 12-lead electrocardiogram, and blood pressure and transcutaneous oxygen saturation will be measured after 5 minutes of rest in the sitting position, during the test, and shortly after termination. The Borg rating [50] of perceived exertion will be assessed at the end of the test. The VO_2peak is defined as the highest oxygen consumption value sampled over a 30-second interval. The ventilatory threshold will be calculated using the ventilatory equivalent method, expressed as oxygen uptake (VO_2). The minute ventilation (VE) and carbon dioxide responses during exercise will be used to calculate the VE/carbon dioxide uptake slope. The breathing reserve will be calculated using measured maximal voluntary ventilation (MVV) and the maximal VE using the following equation:

$$([MVV - VE]/MVV) \times 100 \quad \textbf{(1)}$$

The oxygen pulse will be calculated by dividing VO_2peak (in mL) by the peak heart rate [51].

Isometric Strength

Isometric bench press and isometric leg extension will be performed using a custom-built strength ergometer (Gym 2000) with a strain gauge (US2A100 kg, Holtinger) and a custom-made amplifier. The ergometer is specially designed for testing children. In the bench press test, participants will be instructed to push an (upward) static bar with maximal effort for 5 seconds. The participants will lie on the ergometer with the bar height adjusted so that the elbow angle is 90° to get the upper arm in a horizontal position and the lower arm in a vertical position. The participants will be required to maintain the elbow angle throughout the duration of the trial. Each participant will perform at least three trials with 2 minutes of rest between each trial but will be allowed to continue with further trials as long as improvements are observed. The highest value of the trials will be used for later analysis. In the knee extension test, the participant will be seated at the short end of the ergometer bench with his or her back straight, arms down, and hands lightly gripping the seat. During contraction, both hip and knee joint angles will be flexed at 90° (180° will be fully extended), respectively. For the knee extension isometric test, the participants will also perform at least three trials with 2-minute rest between trials but will be allowed to continue with further trials as long as improvement is observed. The best result in each test will be used for later analysis. Hand-grip strength will be measured in the left and right hands (Baseline Life Hand Dynamometer) while the participant will be standing and with the arm extended and pointing down. Each hand will be measured 3 times, with alternating sides, starting with the right hand and approximately 30-second breaks between measurements. The best result for each hand will be used for later analysis.

Muscular Endurance

Muscular endurance will be tested using the 1-minute sit-to-stand (STS) test [52]. The participants will perform 1 test trial at least 20 minutes before the final test. The number of repetitions of standing up and sitting down from a chair within 1 minute in the final test will be recorded. In addition, the Borg rating [50] of perceived exertion will be assessed at the end of the test. The test will be performed on a height-adjustable chair to ensure a 90° knee angle. The 1-minute STS test showed high reliability and good criterion-related validity with other exercise capacity tests such as the 6-minute walk test or stair climbing [53-56].

Countermovement Jump

All jumps will be performed on a portable force plate (FP4; HUR-Laboratories Oy). Hands will be placed on the hips throughout the test, and the angular displacement of the knees will be standardized so that the participants are instructed to squat down until the knees are bent at approximately 90° and then immediately jump vertically as high as possible, landing back on the force plate on both feet at the same time. The jump height and power will be measured [57]. The best result of at least three attempts will be recorded; however, the participants will be allowed to continue to jump as long as improvements are made. There will be 1 minute of rest between the trials.

Blood Volume and Hemoglobin Mass

A known carbon monoxide (CO) dose of approximately 1.2 mL/kg body mass will be administered and rebreathed for 2 minutes. Capillary fingertip blood samples will be collected before the start of the test and 7 minutes after the administration of the CO dose. Both blood samples will be measured a minimum of 5 times to determine the percentage of carboxyhemoglobin using an OSM3 Hemoximeter (Radiometer). Hemoglobin mass will be calculated from the mean change in percentage of carboxyhemoglobin before and after rebreathing CO [58]. Blood volume, plasma volume, and red cell volume will be calculated from hemoglobin mass using venous hemoglobin and venous hematocrit according to Burge and Skinner [59] and Heinicke et al [60].

Lung Function

Lung function will be measured using the maximal expiratory flow volume loops. Forced expiratory volume in 1 second, forced vital capacity (FVC), maximal expiratory flow at 50% of FVC (MEF_{50}), peak expiratory flow, and the ratio of forced expiratory volume in 1 second to FVC will be recorded. Lung volumes (total lung capacity and residual volume) and airway resistance will be measured using standardized whole-body plethysmography. Diffusion capacity will be measured using the single-breath method, and the result will be adjusted for serum hemoglobin levels. All measurements are according to the American Thoracic Society/European Respiratory Societies guidelines. The Global Lung Function Initiative reference equations will be used, and the outcomes will be expressed as absolute values, percent predicted, and *z* scores and the lower limit of normal.

Cardiac Examination

A full functional echocardiography protocol, including 2D, color, pulsed, continuous, and tissue Doppler registrations, will be performed using a GE Vivid E95 ultrasound machine (GE Healthcare) with 4- and 5 MHz 2D probes (4Vc-D or M5Sc).

3D volumes will be acquired as multi-beat fusions with breath hold. The frame rate for 2D images will be kept between 50 and 90 frames per second for 2D strain measurements. Automatic image optimization will be used as the default setting, supplied by manual optimization, as required. Parasternal, apical, selective right ventricular apical, and subcostal views will be applied. For clinical use, left ventricular ejection fraction, as calculated by the Simpson biplane method, fractional shortening by m-mode (MediMatic Compacs software), and peak systolic left ventricular global longitudinal strain (GE EchoPac software), will be reported in addition to any other positive findings of clinical significance. Additional research-based measurements will be performed as the mean value of 3 consecutive cardiac cycles. Left ventricular 4D volumes will be measured using a semiautomatic method (EchoPac Auto-LVQ).

Neurological Evaluation

Detailed patient history and clinical neurological examination will be performed. The symptoms and findings relevant for polyneuropathy and neuropathic pain will be scored according to the Toronto Neuropathy Symptom Score, Douleur Neuropathique 4, and Neupsig criteria for neuropathic pain [61-63]. Standard nerve conduction studies on three motor nerves (ulnar, peroneal, and tibial), four sensory nerves (ulnar, radial, superficial peroneal, and sural), and one mixed motor or sensory nerve (tibial plantar nerve) will be performed. Assessment of small fiber function will be conducted by quantitative thermal testing (cold perception threshold and warmth perception threshold) on the hands and feet. All measurements will be performed on the right extremities. The temperature will be maintained at >30 °C in feet and >32 °C in hand. Polyneuropathy will be diagnosed and classified according to the criteria suggested by Tesfaye et al [64] as possible, probable, or confirmed polyneuropathy (clinical grading) and according to the Tesfaye neurophysiological criteria (stadium 0, 1a, 1b, 2a, and 2b), which includes subclinical polyneuropathy.

Anthropometry and Maturity

Body weight will be measured using a digital scale (SECA 876 GmbH and Co), with participants wearing light clothing and no shoes. Height will be measured using a stadiometer (SECA 217 GmbH and Co). Body composition will be assessed by dual-energy x-ray absorptiometry (Lunar iDXA GE Healthcare using the enCORE software Version 14.10.022). Participants will be scanned from head to toe in a supine position, providing values for total lean tissue (kg), fat mass (kg), bone mineral content, and bone mineral density. In WP 2, the puberty stage of survivors of childhood cancer will be assessed clinically by a medical physician at the first visit in Norway. In addition, all participants (survivor of childhood cancer and controls) will self-assess their puberty stage using a modified Tanner method [65].

Muscle Thickness

Muscle thickness and pennation angle of the musculus vastus lateralis and thickness of the musculus biceps brachii will be measured by B-mode ultrasound using a linear-array transducer (50 mm, 5-12 MHz, HD11XE, Philips Ultrasound, revision 2.0.3). The sites for the measurements will be on one-third of the humerus length measured from the lateral epicondyle for the musculus biceps brachii and on 40% of the femur length measured from the lateral epicondyle for the musculus vastus lateralis.

PA Measures

Whenever the time between participation in WP 1 and recruitment in WP 2 exceeds 3 months, PA will be remeasured by accelerometers as described above (WP 1).

Data Management

Study data will be collected and managed using REDCap (Research Electronic Data Capture) electronic data capture tools hosted at the Norwegian School of Sport Sciences [66,67]. REDCap is a secure, web-based software platform designed to support data capture for research studies. Data from all sites will be entered by the study sites independently in REDCap.

Power Calculation and Statistical Analyses WP 2

Sample size calculations and power assessment were based on VO_2peak. On the basis of results from international studies, we expect to find a difference of 10% between survivors of childhood cancer and controls [22]. To detect a difference of 10% in VO_2peak between survivors and healthy controls, with a statistical strength of 80%, the study needs to include 23 survivors of childhood cancer. The same number of patients are needed to get the same statistical power for the strength tests assuming a difference of 10%. Owing to heterogeneity between diagnostic groups and the need for subgroup analysis (diagnosis, age, and gender), we plan to include 150 survivors as we expect the group size to be as small as 30 participants.

We will use descriptive statistics to present the current PF, physiological determinants of PF, and PA of survivors of childhood cancer compared with controls. We will use multilevel (mixed) linear models to compare PF and PA between survivors of childhood cancer and controls and include an interaction term (for instance, age group, gender, and diagnosis) to identify differences in physiological determinants between survivors of childhood cancer and controls. All analyses will be adjusted for the study site by including sites as random intercepts in the mixed models. Where available and depending on the research question, the fitness outcomes will be transformed into *z* scores according to the norm data. Depending on the outcome, we will perform univariable and multivariable mixed logistic, linear, or ordinal regression models to identify clinical and physiological factors associated with PF in survivors.

WP 3

The outcomes of WP 3 include facilitators of and barriers to PA. Norway (Oslo and Bergen) and Denmark (Copenhagen) will participate in WP 3.

Theoretical Framework and Procedures

The World Health Organization's International Classification of Functioning, Disability, and Health for Children and Youth (ICF-CY) illustrates the interrelationships between body functions and structural impairments (eg, late effects), activity limitations, and restricted ability to participate within an

individual's context of environmental and personal factors to predict outcomes (eg, functioning and health) [68]. Therefore, the ICF-CY is a suitable biopsychosocial framework to systematically identify potential barriers to and facilitators of PA among young survivors of childhood cancer. Inspired by the ICF-CY model, we will explore facilitators of and barriers to PA through in-depth interviews with young survivors of childhood cancer and their parents. A semistructured interview guide will be developed based on current knowledge of barriers to and facilitators of PA and inspired by the ICF-CY framework (Multimedia Appendix 1) to be used at all 3 sites. The interviews will follow the guide and last approximately 15 to 60 minutes. All interviews will be audio recorded and transcribed ad verbatim.

Data Analysis WP 3

NVivo software will be used for data management and analysis [69]. Analysis will be conducted in parallel with data collection to ensure that any new issues of interest arising are explored in subsequent interviews. Interviews from survivors and parents will be analyzed separately. All data will be analyzed according to the principles of thematic analysis [70,71]. After familiarization with the initial interviews, at least two researchers will analyze the data line by line, creating codes reflecting the contents of each utterance to ensure objectivity [72,73]. Any disagreements regarding the codes will be discussed with the project group and settled. For analytical rigor, codes will be used and discussed by all members of the research team to develop a codebook that will be used in all interviews for all sites. We will follow the COREQ (Consolidated Criteria for Reporting Qualitative Studies) guidelines for conducting and reporting qualitative research [74].

WP 4

Overview

WP 4 aims to develop and feasibility test a sustainable and safe PA intervention for survivors of childhood cancer based on the results obtained in WPs 1 to 3. Norway, Finland, and Germany will participate in WP 4. The study will comprise two study visits (baseline and after 6 months), including self-reported questionnaires, interviews, measurement of PA, and PF tests (Table 1; Table S1 in Multimedia Appendix 1). In between, the survivors will participate in a personalized PA intervention, followed by central and local coaches.

In specific, we aim to do the following:

1. Describe the adherence to and acceptability, satisfaction, and safety of the personalized PA intervention
2. Test the involvement of local structures (parents, school, teachers, peers, community and school nurses, and sports clubs) to improve the PA of survivors
3. Assess the effect of the intervention on PA, attitude toward PA, QoL, fatigue, and PF
4. Explore the experiences of survivors of childhood cancer and their parents of participating in the PA intervention
5. Explore the central coaches' experiences providing the intervention and collaborating with local structures

Intervention

The exercise intervention will last 6 months and starts with a personal motivational interview (MI) session at the baseline study visit. For each survivor, a central study coach is appointed who performs the MI and, together with the participant, defines a personal target of duration and intensity of PAs per week. The MI technique is a communication method that involves enhancing a patient's internal motivation to change [75]. During this session, local structures (local coaches) will be identified that can help and support the survivor to reach their PA goals and maintain a more active lifestyle after the intervention. On the basis of this assessment, individualized PAs are defined and implemented into a survivor's daily life. Each survivor will get a written personalized plan at the end of the session and a Polar watch (Polar Global) to track their behavior and provide immediate feedback on their achievements.

The central study coach will follow up with the survivor and local coaches with a phone call based on a predefined schedule of decreasing frequency over time (Table S1 in Multimedia Appendix 1). During the follow-up phone calls, the coach uses MI-inspired techniques and a standardized guideline to discuss adherence, motivation, progress, and possible problems. The training program can be adapted, and solutions for solving problems will be searched. The central coach will also follow the PA behavior in real time based on the Polar watch and set up an additional follow-up contact if the survivor struggles to reach his or her activity goals. The central coaches will be trained in the use of MI techniques to be used in the initial and follow-up conversations with survivors.

Assessments

Table S1 in Multimedia Appendix 1 gives an overview of all assessments in WP 4.

Feasibility Parameters

Recruitment

The number of contacted and recruited survivors with reasons for nonrecruitment will be recorded by the study nurses at each site.

PA Parameters

PA will be measured by an accelerometer as described for WP 1 at baseline and after 6 months. In addition, the PA will be tracked continuously with the Polar watch (Polar Global). At each follow-up discussion with the central coaches, the survivors will be asked about their participation in the agreed PA program and, if applicable, reasons for nonparticipation.

Compliance (Adherence and Retention)

Compliance with the PA intervention will be assessed based on the PA measurements described above. In addition, we will record the number of contacts of each survivor with the central and local coaches and the number of missing information at each study time point. Dropouts and reasons for dropout will be assessed by the study nurses.

Safety

Safety and adverse events will be assessed by the central coaches during follow-up calls.

Subjective Rating of the Intervention and Its Motivational Features

At the end of the intervention, the survivors will give their subjective rating of satisfaction with participation in and intended sustainability of the PA intervention. The questions will be developed specifically for this study.

Attitude Toward PA (Barriers, Perceived Physical Competence, Perceived Enjoyment, and Motivation)

Attitudes toward PA will be assessed at baseline and at 2 and 6 months. The same questions as in WP 1 will be used to assess *barriers to PA*. *Perceived physical competence* (3 items) and *perceived enjoyment* (5 items) will be measured by items from the *Children's Attraction to PA* scale [76-78]. *Motivation toward PA* will be measured using the Behavioral Regulation in Exercise Questionnaire 2 [79].

QoL and Fatigue

At baseline and 6 months, survivors will fill in the same questionnaires as in WP 1 to assess QoL and fatigue (only self-reported version and no parent module).

PF Assessment

PF will be assessed during the study visits at baseline and after 6 months. All sites will perform a CPET and 1-minute STS test according to the same protocol as described for WP 2. In addition, Oslo and Bergen will perform the same isometric strength tests (knee extension and chest press) as in WP 2. Turku will perform the same knee extension test as WP2 but on its own device as well as a dynamic upper extremity lifting test (standing with 5 kg dumbbells for women and 10 kg for men; number of repetitions are reported). Essen will perform the same isometric knee extension and chest press as in WP2 but on their own devices.

Qualitative Interviews With Survivors, Parents, and Central Coaches

Feasibility, acceptability, sustainability, and satisfaction with the PA intervention will also be assessed by qualitative in-depth interviews in a subsample of survivors and their parents from the Oslo site at the 6-month visit (8-15 survivors and their parents until information saturation is reached). Furthermore, the central coaches will participate in a focus group interview to explore their views on feasibility, acceptability, and satisfaction. The interview guides are provided in Multimedia Appendix 2.

Data Management

All self-reported questionnaires will be answered electronically either on the study tablet PCs during the study visits (baseline and 6 months) or from home via a personal link (2 months). The questionnaires will be linked to the TSD secure data server at the University of Oslo, where the data will be stored. Device-measured data (Actigraph and Polar watch [Polar Global]) will be downloaded and added to the folder on the TSD. Information assessed by the study nurses, central coaches, and assessors of the physical testing will be collected in Microsoft Excel files and stored on TSD. All information will then be merged into a master data file using the statistical software STATA (version 17 or newer; StataCorp LLC).

Data Analyses WP 4

Quantitative Data

As this is a feasibility study, we will mainly use descriptive statistics to present the study results. We will use numbers and proportions to describe categorical variables and means, SDs (or median and IQR if the data are skewed), and ranges to describe continuous variables. *P* values will be calculated based on appropriate univariable and multivariable models depending on the outcome and adjusted for the study site by introducing a random intercept (mixed model). *P* values are interpreted as flagging trends and not as hypothesis tests.

Qualitative Data

Qualitative content analyses of interview transcripts will be performed [70].

Ethics Approval

The PACCS study has been approved by the Norwegian Regional Committee for Medical Research Ethics (WP 1,3,4: 2016/953 and WP 2:2018/739), the Data Protection Officer at Oslo University Hospital, and by the equivalent institutions in Finland, Denmark (file. H-19032270), Germany, and Switzerland (Ethics Committee of Northwestern and Central Switzerland; project ID: 2019-00410). Approvals from the regional ethical committees for medical and health research at all sites and the project owner institution are the assumptions for the implementation of the project. As most of the participants will be too young to provide valid informed consent, they will receive written and verbal information about the studies adapted to their developmental stage. Parents or legal guardians will provide written consent on their behalf. School nurses, physical education teachers, and other relevant parties participating in WP 3 will provide informed consent in writing. This protocol is written according to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [80].

Results

Research activity for this study commenced in November 2016 with the piloting of questionnaires and recruitment to WP 1 at the Oslo, Norway site. The project received funding from the Research Council of Norway for the period from June 2018 to June 2022. International collaboration was established after June 2018. Data collection at all sites was initiated by October 2020 for WP 1 and concluded on January 2021, with 517 survivors of childhood cancer included. Data cleaning and analysis are currently underway. A total of articles are under preparation for publication. Data collection for WP 2 was initiated in January 2019 and concluded in December 2020, with 157 survivors of childhood cancer and 113 controls included. Data cleaning and analysis are currently being conducted. A total of 6 articles are planned based on the WP 2 data set. Data collection for WP 3 was initiated in January 2018. Patient and parent interviews were concluded in July 2020, with 63 survivors of childhood cancer and 68 parents. Interviews with local stakeholders were initiated in January 2020 and concluded in October 2020, with a total of 18 participants interviewed (teachers, special needs teachers, and school nurses). Data from the patient and parent interviews have been analyzed

and are currently being written for publication. A total of 2 articles have been published as of January 2022 [81,82], and 2 more are under preparation for publication. Data from interviews with local stakeholders are currently being analyzed, and 1 article is planned. WP 4 recruitment was initiated in May 2021 and is still ongoing as of January 2022. The study is planned to be completed by July 2022. As of January 2022, 23 survivors of childhood cancer have been included, of which 11 have completed the intervention. Detailed plans for data analyses for WPs 1 to 4 are described in the "Assessments, Power Calculations, and Statistical Analyses" section.

Discussion

Principal Findings

Survivors of childhood cancer represent a constantly growing population in society with high morbidity [4], premature mortality [83], and increased uptake of social benefits [84]. Owing to their disease and treatment exposure, they are at risk of chronic medical conditions such as cardiovascular and pulmonary diseases, secondary cancers, metabolic syndrome and obesity, osteoporosis, fatigue, neurocognitive complaints, and psychological distress, all of which can negatively affect PF and function and, thereby, the survivors' ability to participate in everyday and social activities [3,4,8]. Options for preventing late effects are currently limited; however, a few small-scale intervention studies indicate promising effects of PA on promoting physical functioning, health, and QoL [10-13]. However, important knowledge gaps and methodological shortcomings remain, including heterogeneity of the samples, lack of controls, quality of assessment methods, and outcome measures used, limiting the generalizability of the results [17]. This applies to survivors of childhood cancer in general, especially for young survivors of childhood cancer aged <18 years.

The PACCS study will generate high-quality, generalizable, objective, and robust data on PA and PF in young survivors of childhood cancer compared with controls, as well as on physiological, psychological, and social factors associated with PA and PF in young survivors of childhood cancer. These data are crucial to establish the extent of the need for PA interventions for young survivors of childhood cancer and for the development and testing of evidence-based, safe, and efficacious PA interventions. Moreover, the data generating WPs and the intervention development will follow a behavioral science approach, supported by the theoretical framework of the ICF-CY, to strengthen their methodological rigor [85].

The resulting intervention will make use of and test behavioral change strategies (MI) for feasibility as a means of tailoring PAs and goals to the individual survivors, with adherence monitored using activity-tracking watches. This will contribute to overcoming barriers to PA and low adherence rates, thereby potentially increasing the sustainability of such interventions. Crucially, sustainability of the intervention is boosted by anchoring it locally in the survivor's home environment, in collaboration with key school, health, and sports activity personnel to maintain focus on PA beyond the intervention period.

PA is currently not a part of routine follow-up care in many countries, and non–evidence-based PA programs are only sporadically offered to survivors of childhood cancer. Thus, the international, multidisciplinary nature of the PACCS consortium, representing 6 European pediatric oncology institutions, will aid the harmonization of PA programs and counseling. Therefore, the project results may improve the current clinical practice.

Promoting regular moderate to vigorous levels of PA in survivors of childhood cancer is likely to reduce cardiovascular morbidity [86] and mortality [14,87-89], osteoporosis, and fatigue and increase psychological well-being. Moreover, as PA is an important determinant of PF, both PA and PF are essential for gaining developmental skills in children and adolescents. Ensuring sufficient levels of PA in young survivors of childhood cancer is likely to have long-reaching benefits for their functioning and, therefore, the opportunity to be active and participating members of society, including maintaining friendships and participating in sports and recreational activities. As such, systematic PA rehabilitation programs have the potential to reduce the burden of survival on individual survivors, as well as reduce health care and social costs for the survivors of childhood cancer.

Conclusions

Adolescence provides a *golden period* of life for the establishment of healthy behaviors and is, therefore, the prime target for intervention [27,42,90]. Our intention is to design a rehabilitation program that uses already existing local expertise and infrastructure. Moreover, if successful, the results of the observational study and feasibility intervention may be extrapolated to other groups of adolescents with chronic health problems; for example, adolescents with congenital heart defects or chronic pulmonary diseases. Thus, the project may benefit larger patient groups.

Acknowledgments

The authors thank all members of the Physical Activity and Fitness in Childhood Cancer Survivors (PACCS) study group who are not listed as coauthors of this paper. Research assistants Jannicke Kirkøen and Anastasia Isotova provided invaluable help in editing the manuscript.

The study is funded by the Research Council of Norway (ID number: 273500). Lene Thorsen is supported by AKTIV Against Cancer.

Authors' Contributions

ER, SA, HCL, M Grydeland, TR, CSR, M Götte, PL, SK, and HBL conceived the original study idea. ER, SA, HCL, M Grydeland, TR, CSR, M Götte, PML, SK, and HBL designed the study, the study materials, the data collection tools, and methods and drafted the manuscript. The principal investigators are ER and SA for Norway, PML for Finland, M Götte for Germany, SK for Switzerland, and HBL for Denmark. M Grydeland is the head of work package (WP) 1, TR is the head of WP 2, HCL is the head of WP 3, and CSR is the head of WP 4. All authors reviewed and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Schedule of enrollment, interventions, and assessments within Physical Activity and Fitness in Childhood Cancer Survivors work package 4.
[DOCX File , 43 KB - resprot_v11i3e35838_app1.docx]

Multimedia Appendix 2
Interview guides work packages 3 and 4.
[DOCX File , 29 KB - resprot_v11i3e35838_app2.docx]

Multimedia Appendix 3
Peer review report from project application process by the Research Council Norway.
[PDF File (Adobe PDF File), 101 KB - resprot_v11i3e35838_app3.pdf]

References

1. Gatta G, Botta L, Rossi S, Aareleid T, Bielska-Lasota M, Clavel J, EUROCARE Working Group. Childhood cancer survival in Europe 1999-2007: results of EUROCARE-5--a population-based study. Lancet Oncol 2014;15(1):35-47. [doi: 10.1016/S1470-2045(13)70548-5] [Medline: 24314616]
2. Vassal G, Schrappe M, Pritchard-Jones K, Arnold F, Basset L, Biondi A, et al. The SIOPE strategic plan: a European cancer plan for children and adolescents. J Cancer Policy 2016;8:17-32. [doi: 10.1016/j.jcpo.2016.03.007]
3. Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, Childhood Cancer Survivor Study. Chronic health conditions in adult survivors of childhood cancer. N Engl J Med 2006;355(15):1572-1582. [doi: 10.1056/NEJMsa060185] [Medline: 17035650]
4. Hudson MM, Ness KK, Gurney JG, Mulrooney DA, Chemaitilly W, Krull KR, et al. Clinical ascertainment of health outcomes among adults treated for childhood cancer. JAMA 2013;309(22):2371-2381 [FREE Full text] [doi: 10.1001/jama.2013.6296] [Medline: 23757085]
5. Erdmann F, Frederiksen LE, Bonaventure A, Mader L, Hasle H, Robison LL, et al. Childhood cancer: survival, treatment modalities, late effects and improvements over time. Cancer Epidemiol 2021;71(Pt B):101733. [doi: 10.1016/j.canep.2020.101733] [Medline: 32461035]
6. Yeh JM, Ward ZJ, Chaudhry A, Liu Q, Yasui Y, Armstrong GT, et al. Life expectancy of adult survivors of childhood cancer over 3 decades. JAMA Oncol 2020;6(3):350-357 [FREE Full text] [doi: 10.1001/jamaoncol.2019.5582] [Medline: 31895405]
7. Fidler MM, Reulen RC, Winter DL, Kelly J, Jenkinson HC, Skinner R, British Childhood Cancer Survivor Study Steering Group. Long term cause specific mortality among 34 489 five year survivors of childhood cancer in Great Britain: population based cohort study. BMJ 2016;354:i4351 [FREE Full text] [doi: 10.1136/bmj.i4351] [Medline: 27586237]
8. Bhakta N, Liu Q, Ness KK, Baassiri M, Eissa H, Yeo F, et al. The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). Lancet 2017;390(10112):2569-2582 [FREE Full text] [doi: 10.1016/S0140-6736(17)31610-0] [Medline: 28890157]
9. Antwi GO, Jayawardene W, Lohrmann DK, Mueller EL. Physical activity and fitness among pediatric cancer survivors: a meta-analysis of observational studies. Support Care Cancer 2019;27(9):3183-3194. [doi: 10.1007/s00520-019-04788-z] [Medline: 30993451]
10. Cheung AT, Li WH, Ho LL, Ho KY, Chan GC, Chung JO. Physical activity for pediatric cancer survivors: a systematic review of randomized controlled trials. J Cancer Surviv 2021;15(6):876-889 [FREE Full text] [doi: 10.1007/s11764-020-00981-w] [Medline: 33389553]
11. Esbenshade AJ, Ness KK. Dietary and exercise interventions for pediatric oncology patients: the way forward. J Natl Cancer Inst Monogr 2019;2019(54):157-162. [doi: 10.1093/jncimonographs/lgz021] [Medline: 31532528]

12. Morales JS, Valenzuela PL, Herrera-Olivares AM, Baño-Rodrigo A, Castillo-García A, Rincón-Castanedo C, et al. Exercise interventions and cardiovascular health in childhood cancer: a meta-analysis. Int J Sports Med 2020;41(3):141-153. [doi: 10.1055/a-1073-8104] [Medline: 31935777]
13. Morales JS, Valenzuela PL, Herrera-Olivares AM, Rincón-Castanedo C, Martín-Ruiz A, Castillo-García A, et al. What are the effects of exercise training in childhood cancer survivors? A systematic review. Cancer Metastasis Rev 2020;39(1):115-125. [doi: 10.1007/s10555-020-09852-3] [Medline: 31970589]
14. Scott JM, Li N, Liu Q, Yasui Y, Leisenring W, Nathan PC, et al. Association of exercise with mortality in adult survivors of childhood cancer. JAMA Oncol 2018;4(10):1352-1358 [FREE Full text] [doi: 10.1001/jamaoncol.2018.2254] [Medline: 29862412]
15. Tonorezos ES, Ford JS, Wang L, Ness KK, Yasui Y, Leisenring W, et al. Impact of exercise on psychological burden in adult survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. Cancer 2019;125(17):3059-3067 [FREE Full text] [doi: 10.1002/cncr.32173] [Medline: 31067357]
16. Wurz A, McLaughlin E, Lategan C, Ellis K, Culos-Reed SN. Synthesizing the literature on physical activity among children and adolescents affected by cancer: evidence for the international Pediatric Oncology Exercise Guidelines (iPOEG). Transl Behav Med 2021;11(3):699-708 [FREE Full text] [doi: 10.1093/tbm/ibaa136] [Medline: 33538309]
17. Braam KI, van der Torre P, Takken T, Veening MA, van Dulmen-den Broeder E, Kaspers GJ. Physical exercise training interventions for children and young adults during and after treatment for childhood cancer. Cochrane Database Syst Rev 2016;3(3):CD008796 [FREE Full text] [doi: 10.1002/14651858.CD008796.pub3] [Medline: 27030386]
18. Winter C, Müller C, Hoffmann C, Boos J, Rosenbaum D. Physical activity and childhood cancer. Pediatr Blood Cancer 2010;54(4):501-510. [doi: 10.1002/pbc.22271] [Medline: 19743298]
19. Götte M, Seidel CC, Kesting SV, Rosenbaum D, Boos J. Objectively measured versus self-reported physical activity in children and adolescents with cancer. PLoS One 2017;12(2):e0172216 [FREE Full text] [doi: 10.1371/journal.pone.0172216] [Medline: 28207820]
20. Wilson CL, Stratton K, Leisenring WL, Oeffinger KC, Nathan PC, Wasilewski-Masker K, et al. Decline in physical activity level in the Childhood Cancer Survivor Study cohort. Cancer Epidemiol Biomarkers Prev 2014;23(8):1619-1627 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-0213] [Medline: 24842624]
21. Yildiz Kabak V, Calders P, Duger T, Mohammed J, van Breda E. Short and long-term impairments of cardiopulmonary fitness level in previous childhood cancer cases: a systematic review. Support Care Cancer 2019;27(1):69-86. [doi: 10.1007/s00520-018-4483-8] [Medline: 30251066]
22. Berkman AM, Lakoski SG. A review of cardiorespiratory fitness in adolescent and young adult survivors of childhood cancer: factors that affect its decline and opportunities for intervention. J Adolesc Young Adult Oncol 2016;5(1):8-15. [doi: 10.1089/jayao.2015.0031] [Medline: 26812453]
23. Fridh MK, Simonsen C, Schmidt-Andersen P, Nissen AA, Christensen JF, Larsen A, et al. Cardiorespiratory fitness and physical performance after childhood hematopoietic stem cell transplantation: a systematic review and meta-analysis. Bone Marrow Transplant 2021;56(9):2063-2078. [doi: 10.1038/s41409-021-01370-2] [Medline: 34155356]
24. Järvelä LS, Niinikoski H, Lähteenmäki PM, Heinonen OJ, Kapanen J, Arola M, et al. Physical activity and fitness in adolescent and young adult long-term survivors of childhood acute lymphoblastic leukaemia. J Cancer Surviv 2010;4(4):339-345. [doi: 10.1007/s11764-010-0131-0] [Medline: 20552291]
25. Christiansen JR, Kanellopoulos A, Lund MB, Massey R, Dalen H, Kiserud CE, et al. Impaired exercise capacity and left ventricular function in long-term adult survivors of childhood acute lymphoblastic leukemia. Pediatr Blood Cancer 2015;62(8):1437-1443. [doi: 10.1002/pbc.25492] [Medline: 25832752]
26. Kanellopoulos A, Hamre HM, Dahl AA, Fosså SD, Ruud E. Factors associated with poor quality of life in survivors of childhood acute lymphoblastic leukemia and lymphoma. Pediatr Blood Cancer 2013;60(5):849-855. [doi: 10.1002/pbc.24375] [Medline: 23335116]
27. Ness KK, Hudson MM, Ginsberg JP, Nagarajan R, Kaste SC, Marina N, et al. Physical performance limitations in the Childhood Cancer Survivor Study cohort. J Clin Oncol 2009;27(14):2382-2389 [FREE Full text] [doi: 10.1200/JCO.2008.21.1482] [Medline: 19332713]
28. Rueegg CS, Gianinazzi ME, Michel G, von der Weid NX, Bergstraesser E, Kuehni CE, Swiss Paediatric Oncology Group (SPOG). Do childhood cancer survivors with physical performance limitations reach healthy activity levels? Pediatr Blood Cancer 2013;60(10):1714-1720. [doi: 10.1002/pbc.24595] [Medline: 23681516]
29. Mizrahi D, Wakefield CE, Simar D, Ha L, McBride J, Field P, et al. Barriers and enablers to physical activity and aerobic fitness deficits among childhood cancer survivors. Pediatr Blood Cancer 2020;67(7):e28339. [doi: 10.1002/pbc.28339] [Medline: 32386117]
30. Ross WL, Le A, Zheng DJ, Mitchell H, Rotatori J, Li F, et al. Physical activity barriers, preferences, and beliefs in childhood cancer patients. Support Care Cancer 2018;26(7):2177-2184. [doi: 10.1007/s00520-017-4041-9] [Medline: 29383508]
31. Stolley MR, Restrepo J, Sharp LK. Diet and physical activity in childhood cancer survivors: a review of the literature. Ann Behav Med 2010;39(3):232-249 [FREE Full text] [doi: 10.1007/s12160-010-9192-6] [Medline: 20559768]
32. Huang TT, Ness KK. Exercise interventions in children with cancer: a review. Int J Pediatr 2011;2011:461512 [FREE Full text] [doi: 10.1155/2011/461512] [Medline: 22121378]

33. Physical activity in childhood cancer survivors. URL: https://www.paccs.no/ [accessed 2022-02-24]
34. Lien N, Bjelland M, Bergh IH, Grydeland M, Anderssen SA, Ommundsen Y, et al. Design of a 20-month comprehensive, multicomponent school-based randomised trial to promote healthy weight development among 11-13 year olds: the HEalth in Adolescents study. Scand J Public Health 2010;38(5 Suppl):38-51. [doi: 10.1177/1403494810379894] [Medline: 21062838]
35. Dalene KE, Anderssen SA, Andersen LB, Steene-Johannessen J, Ekelund U, Hansen BH, et al. Cross-sectional and prospective associations between physical activity, body mass index and waist circumference in children and adolescents. Obes Sci Pract 2017;3(3):249-257 [FREE Full text] [doi: 10.1002/osp4.114] [Medline: 29071101]
36. Erdvik IB, Haugen T, Ivarsson A, Säfvenbom R. The temporal relations of adolescents' basic need satisfaction in physical education and global self-worth. J Sport Exerc Psychol 2020:1-10. [doi: 10.1123/jsep.2019-0030] [Medline: 33186899]
37. Samdal O, Mathisen FK, Torsheim T, Diseth AR, Fismen AS, Larsen TM, et al. Helse og trivsel blant barn og unge. Resultater fra den landsrepresentative spørreundersøkelsen «Helsevaner blant skoleelever. En WHO-undersøkelse i flere land». Bergen Open Research Archive. 2016. URL: https://bora.uib.no/bora-xmlui/handle/1956/13072 [accessed 2021-09-10]
38. Ungkost 2000 – Landsomfattende kostholdsundersøkelse. Norwegian Directorate of Health. 2020. URL: https://www.helsedirektoratet.no/rapporter/ungkost-2000-landsomfattende-kostholdsundersokelse [accessed 2021-09-10]
39. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care 1999;37(2):126-139. [doi: 10.1097/00005650-199902000-00003] [Medline: 10024117]
40. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in pediatric cancer: reliability and validity of the pediatric quality of life inventory generic core scales, multidimensional fatigue scale, and cancer module. Cancer 2002;94(7):2090-2106 [FREE Full text] [doi: 10.1002/cncr.10428] [Medline: 11932914]
41. Goodman R. The strengths and difficulties questionnaire: a research note. J Child Psychol Psychiatry 1997;38(5):581-586. [doi: 10.1111/j.1469-7610.1997.tb01545.x] [Medline: 9255702]
42. Kolle E, Steene-Johannessen J, Andersen LB, Anderssen SA. Objectively assessed physical activity and aerobic fitness in a population-based sample of Norwegian 9- and 15-year-olds. Scand J Med Sci Sports 2010;20(1):e41-e47. [doi: 10.1111/j.1600-0838.2009.00892.x] [Medline: 19422647]
43. Dalene KE, Anderssen SA, Andersen LB, Steene-Johannessen J, Ekelund U, Hansen BH, et al. Cross-sectional and prospective associations between sleep, screen time, active school travel, sports/exercise participation and physical activity in children and adolescents. BMC Public Health 2018;18(1):705 [FREE Full text] [doi: 10.1186/s12889-018-5610-7] [Medline: 29879929]
44. Burchartz A, Oriwol D, Kolb S, Schmidt SC, Wunsch K, Manz K, et al. BMC Public Health 2021;21(1):1081 [FREE Full text] [doi: 10.1186/s12889-021-11114-y] [Medline: 34090411]
45. Reinfjell T, Diseth TH, Veenstra M, Vikan A. Measuring health-related quality of life in young adolescents: reliability and validity in the Norwegian version of the Pediatric Quality of Life Inventory 4.0 (PedsQL) generic core scales. Health Qual Life Outcomes 2006;4:61 [FREE Full text] [doi: 10.1186/1477-7525-4-61] [Medline: 16972987]
46. Hypher R, Andersson S, Finnanger TG, Brandt AE, Hoorelbeke K, Lie HC, et al. Fatigue following pediatric acquired brain injury: interplay with associated factors in a clinical trial population compared to healthy controls. Neuropsychology 2021;35(6):609-621. [doi: 10.1037/neu0000753] [Medline: 34197132]
47. Lien L, Green K, Welander-Vatn A, Bjertness E. Mental and somatic health complaints associated with school bullying between 10th and 12th grade students; results from cross sectional studies in Oslo, Norway. Clin Pract Epidemiol Ment Health 2009;5:6 [FREE Full text] [doi: 10.1186/1745-0179-5-6] [Medline: 19309497]
48. Risum K, Edvardsen E, Godang K, Selvaag AM, Hansen BH, Molberg Ø, et al. Physical fitness in patients with oligoarticular and polyarticular juvenile idiopathic arthritis diagnosed in the era of biologics: a controlled cross-sectional study. Arthritis Care Res (Hoboken) 2019;71(12):1611-1620. [doi: 10.1002/acr.23818] [Medline: 30474929]
49. Godfrey S. Physiological response to exercise in children with lung or heart disease. Arch Dis Child 1970;45(242):534-538 [FREE Full text] [doi: 10.1136/adc.45.242.534] [Medline: 5506939]
50. Borg G. Perceived exertion as an indicator of somatic stress. Scand J Rehabil Med 1970;2(2):92-98. [Medline: 5523831]
51. Wasserman KH, Hansen JE, Sue DY, Stringer WW, Sietseman KE, Sun XG, et al. Principles of exercise testing and interpretation: including pathophysiology and clinical applications. 5th edition. Philadelphia, PA: Lippincot Williams & Wilkins; 2012.
52. Strassmann A, Steurer-Stey C, Lana KD, Zoller M, Turk AJ, Suter P, et al. Population-based reference values for the 1-min sit-to-stand test. Int J Public Health 2013;58(6):949-953. [doi: 10.1007/s00038-013-0504-z] [Medline: 23974352]
53. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999;70(2):113-119. [doi: 10.1080/02701367.1999.10608028] [Medline: 10380242]
54. Ozalevli S, Ozden A, Itil O, Akkoclu A. Comparison of the sit-to-stand test with 6 min walk test in patients with chronic obstructive pulmonary disease. Respir Med 2007;101(2):286-293 [FREE Full text] [doi: 10.1016/j.rmed.2006.05.007] [Medline: 16806873]
55. Bohannon RW, Bubela DJ, Magasi SR, Wang YC, Gershon RC. Sit-to-stand test: performance and determinants across the age-span. Isokinet Exerc Sci 2010;18(4):235-240 [FREE Full text] [doi: 10.3233/IES-2010-0389] [Medline: 25598584]

56. Bohannon RW, Magasi SR, Bubela DJ, Wang YC, Gershon RC. Grip and knee extension muscle strength reflect a common construct among adults. Muscle Nerve 2012;46(4):555-558 [FREE Full text] [doi: 10.1002/mus.23350] [Medline: 22987697]
57. Norum M, Risvang LC, Bjørnsen T, Dimitriou L, Rønning PO, Bjørgen M, et al. Caffeine increases strength and power performance in resistance-trained females during early follicular phase. Scand J Med Sci Sports 2020;30(11):2116-2129. [doi: 10.1111/sms.13776] [Medline: 32681596]
58. Schmidt W, Prommer N. The optimised CO-rebreathing method: a new tool to determine total haemoglobin mass routinely. Eur J Appl Physiol 2005;95(5-6):486-495. [doi: 10.1007/s00421-005-0050-3] [Medline: 16222540]
59. Burge CM, Skinner SL. Determination of hemoglobin mass and blood volume with CO: evaluation and application of a method. J Appl Physiol (1985) 1995;79(2):623-631. [doi: 10.1152/jappl.1995.79.2.623] [Medline: 7592227]
60. Heinicke K, Wolfarth B, Winchenbach P, Biermann B, Schmid A, Huber G, et al. Blood volume and hemoglobin mass in elite athletes of different disciplines. Int J Sports Med 2001;22(7):504-512. [doi: 10.1055/s-2001-17613] [Medline: 11590477]
61. Bril V, Perkins BA. Validation of the Toronto clinical scoring system for diabetic polyneuropathy. Diabetes Care 2002;25(11):2048-2052. [doi: 10.2337/diacare.25.11.2048] [Medline: 12401755]
62. Bouhassira D, Attal N, Alchaar H, Boureau F, Brochet B, Bruxelle J, et al. Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4). Pain 2005;114(1-2):29-36. [doi: 10.1016/j.pain.2004.12.010] [Medline: 15733628]
63. Finnerup NB, Haroutounian S, Kamerman P, Baron R, Bennett DL, Bouhassira D, et al. Neuropathic pain: an updated grading system for research and clinical practice. Pain 2016;157(8):1599-1606 [FREE Full text] [doi: 10.1097/j.pain.0000000000000492] [Medline: 27115670]
64. Tesfaye S, Boulton AJ, Dyck PJ, Freeman R, Horowitz M, Kempler P, Toronto Diabetic Neuropathy Expert Group. Diabetic neuropathies: update on definitions, diagnostic criteria, estimation of severity, and treatments. Diabetes Care 2010;33(10):2285-2293 [FREE Full text] [doi: 10.2337/dc10-1303] [Medline: 20876709]
65. Carskadon MA, Acebo C. A self-administered rating scale for pubertal development. J Adolesc Health 1993;14(3):190-195. [doi: 10.1016/1054-139x(93)90004-9] [Medline: 8323929]
66. 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;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
67. 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;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
68. World Health Organization. International classification of functioning, disability and health: children and youth version: ICF-CY. Geneva, Switzerland: World Health Organization; 2007.
69. NVivo. QSR International. 2020. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2001-05-17]
70. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
71. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
72. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
73. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A, "Psychological Theory" Group. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
74. 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;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
75. Miller WR, Rollnick S. Motivational interviewing: helping people change. 3rd edition. New York, NY: Guilford press; 2012.
76. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. Med Sci Sports Exerc 2000;32(5):963-975. [doi: 10.1097/00005768-200005000-00014] [Medline: 10795788]
77. Harter S. Effectance motivation reconsidered: toward a developmental model. Hum Dev 1978;21(1):34-64. [doi: 10.1159/000271574]
78. Brustad RJ. Who will go out and play? Parental and psychological influences on children's attraction to physical activity. Pediatr Exer Sci 1993;5(3):210-223 [FREE Full text] [doi: 10.1123/pes.5.3.210]
79. Mullan E, Markland D, Ingledew DK. A graded conceptualisation of self-determination in the regulation of exercise behaviour: development of a measure using confirmatory factor analytic procedures. Pers Individ Differ 1997;23(5):745-752. [doi: 10.1016/s0191-8869(97)00107-4]
80. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013;158(3):200-207 [FREE Full text] [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]

81. Larsen MH, Larsen EH, Ruud E, Mellblom A, Helland S, Lie HC. "I have to do things differently now, but I make it work"-young childhood cancer survivors' experiences of self-management in everyday living. J Cancer Surviv 2021. [doi: 10.1007/s11764-021-01066-y] [Medline: 34097249]
82. Petersen NN, Hansson H, Lie HC, Brinkkjaer M, Graungaard B, Hamilton Larsen M, et al. A qualitative study of young childhood cancer survivors and their parents' experiences with treatment-related late effects in everyday life post-treatment. J Adv Nurs 2022;78(3):858-868. [doi: 10.1111/jan.15073] [Medline: 34636441]
83. Reulen RC, Winter DL, Frobisher C, Lancashire ER, Stiller CA, Jenney ME, British Childhood Cancer Survivor Study Steering Group. Long-term cause-specific mortality among survivors of childhood cancer. JAMA 2010;304(2):172-179. [doi: 10.1001/jama.2010.923] [Medline: 20628130]
84. Ghaderi S, Engeland A, Moster D, Ruud E, Syse A, Wesenberg F, et al. Increased uptake of social security benefits among long-term survivors of cancer in childhood, adolescence and young adulthood: a Norwegian population-based cohort study. Br J Cancer 2013;108(7):1525-1533 [FREE Full text] [doi: 10.1038/bjc.2013.107] [Medline: 23481179]
85. Brown MC, Sharp L, Sniehotta FF, Skinner R, Araújo-Soares V. The development of health behaviour change interventions for childhood cancer survivors: the need for a behavioural science approach. Pediatr Blood Cancer 2020;67(9):e28500. [doi: 10.1002/pbc.28500] [Medline: 32614142]
86. Chen JJ, Wu PT, Middlekauff HR, Nguyen KL. Aerobic exercise in anthracycline-induced cardiotoxicity: a systematic review of current evidence and future directions. Am J Physiol Heart Circ Physiol 2017;312(2):H213-H222 [FREE Full text] [doi: 10.1152/ajpheart.00646.2016] [Medline: 27923793]
87. Blaauwbroek R, Bouma MJ, Tuinier W, Groenier KH, de Greef MH, Meyboom-de Jong B, et al. The effect of exercise counselling with feedback from a pedometer on fatigue in adult survivors of childhood cancer: a pilot study. Support Care Cancer 2009;17(8):1041-1048 [FREE Full text] [doi: 10.1007/s00520-008-0533-y] [Medline: 19015892]
88. Van Dijk-Lokkart EM, Steur LM, Braam KI, Veening MA, Huisman J, Takken T, et al. Longitudinal development of cancer-related fatigue and physical activity in childhood cancer patients. Pediatr Blood Cancer 2019;66(12):e27949. [doi: 10.1002/pbc.27949] [Medline: 31436372]
89. Zürcher SJ, Jung R, Monnerat S, Schindera C, Eser P, Meier C, et al. High impact physical activity and bone health of lower extremities in childhood cancer survivors: a cross-sectional study of SURfit. Int J Cancer 2020;147(7):1845-1854 [FREE Full text] [doi: 10.1002/ijc.32963] [Medline: 32167159]
90. Ness KK, Leisenring WM, Huang S, Hudson MM, Gurney JG, Whelan K, et al. Predictors of inactive lifestyle among adult survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. Cancer 2009;115(9):1984-1994 [FREE Full text] [doi: 10.1002/cncr.24209] [Medline: 19224548]

Abbreviations

CO: carbon monoxide
COREQ: Consolidated Criteria for Reporting Qualitative Studies
CPET: cardiopulmonary exercise test
cpm: counts per minute
FVC: forced vital capacity
ICF-CY: International Classification of Functioning, Disability, and Health for Children and Youth
MI: motivational interview
MVV: maximal voluntary ventilation
PA: physical activity
PACCS: Physical Activity and Fitness in Childhood Cancer Survivors
PedsQL: Pediatric Quality of Life Inventory
PF: physical fitness
QoL: quality of life
REDCap: Research Electronic Data Capture
REPAC: Relevance of Physical Activity Contexts in the Everyday Life of Adolescents
SDQ: Strength and Difficulties Questionnaire
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
STS: sit-to-stand
VE: minute ventilation
VO_2: oxygen uptake
VO_2peak: peak oxygen uptake
WP: work package
TSD: Services for Sensitive Data

Edited by T Derrick; submitted 20.12.21; this is a non–peer-reviewed article;accepted 11.01.22; published 08.03.22.

Please cite as:
Lie HC, Anderssen S, Rueegg CS, Raastad T, Grydeland M, Thorsen L, Stensrud T, Edvardsen E, Larsen MH, Torsvik IK, Bovim LP, Götte M, Lähteenmäki PM, Kriemler S, Larsen HB, Fridh MK, Ørstavik K, Brun H, Matthews I, Hornset E, Ruud E
The Physical Activity and Fitness in Childhood Cancer Survivors (PACCS) Study: Protocol for an International Mixed Methods Study
JMIR Res Protoc 2022;11(3):e35838
URL: https://www.researchprotocols.org/2022/3/e35838
doi:10.2196/35838
PMID:35258456

Protocol

Researching the Links Between Smartphone Behavior and Adolescent Well-being With the FUTURE-WP4 (Modeling the Future: Understanding the Impact of Technology on Adolescent's Well-being Work Package 4) Project: Protocol for an Ecological Momentary Assessment Study

Steriani Elavsky[1*], PhD; Jana Blahošová[1*], Mgr; Michaela Lebedíková[1*], Mgr; Michał Tkaczyk[1*], PhD; Martin Tancos[1], Mgr; Jaromír Plhák[1*], PhD; Ondřej Sotolář[1*], Mgr; David Smahel[1*], PhD

Faculty of Informatics, Masaryk University, Brno, Czech Republic

[*]these authors contributed equally

Corresponding Author:
Steriani Elavsky, PhD
Faculty of Informatics
Masaryk University
Botanická 68A
Brno, 60200
Czech Republic
Phone: 420 549491814
Email: elavsky@fi.muni.cz

Abstract

Background: Smartphone ownership has increased among teens within the last decade, with up to 89% of adolescents owning a smartphone and engaging daily with the online world through it. Although the results of recent meta-analyses suggest that engaging digital technology plays only a small role in adolescent well-being, parents, professionals, and policymakers remain concerned about the impact that the instant connectivity of smartphones has on adolescent well-being.

Objective: Herein, we introduce the protocol of a research study investigating the associations between adolescent smartphone use and different facets of well-being (social, physical, and psychological), with the aim to apply innovative methods to address the limitations of existing empirical studies.

Methods: This 12-month prospective study of adolescents uses a repeated measurement-burst design with the ecological momentary assessment methodology. Adolescents (N=203; age range 13-17 years) complete baseline assessments through online questionnaires, four 14-day intensive data collection bursts, and an online questionnaire at the end of the study. As part of the 4 measurement bursts, adolescent smartphone behavior is assessed objectively by passive data collection of smartphone data logs and through self-reports in short questionnaires administered via a custom-built Android app.

Results: The protocol describes the study objectives, research tools (including the development of the Android app and specialized software), and process (including pilot studies, the main study, and targets for machine learning approaches). Two of the 203 enrolled participants provided no data during the first data collection burst of the main study. Preliminary analyses of the data from the first data collection burst indicated an acceptable level of compliance (72.25%) with the daily questionnaires. The design of the study will allow for the assessment of both within- and between-person variabilities in smartphone behavior, as well as short-term variation and long-term change in smartphone behavior and how it impacts the indicators of social, physical, and psychological well-being.

Conclusions: The innovative methods applied in this study (objective smartphone logs, ecological momentary assessment, and machine learning) will allow for a more nuanced assessment of the links between smartphone use and well-being, informing strategies to help adolescents navigate the online world more constructively in terms of the development of their physical, social, and psychological well-being.

International Registered Report Identifier (IRRID): DERR1-10.2196/35984

(JMIR Res Protoc 2022;11(3):e35984) doi:10.2196/35984

KEYWORDS

well-being; adolescents; smartphones; intensive data; ecological momentary assessment

Introduction

Background

Both socialization and leisure-time experiences of adolescents have largely shifted to the digital domain in the past decade. Smartphone ownership has increased among teens over the past 6 years, going from 41% in 2012 up to 89% today among 13- to 17-year-olds [1].

Over 80% of children aged 9 to 16 years report accessing the internet from their phones at least once daily, and only a minority report accessing the internet less often than daily or almost daily, ranging between 11% among Lithuanian children and 35% in France (average 20%) [2]. These European data are in line with data from the United States where close to 90% of adolescents now own smartphones and over 90% access the internet from it at least occasionally [3], suggesting that adolescents' digital experiences occur mostly on their smartphones.

Smartphones, compared to other electronic media devices, enable instant connectivity to friends as well as the "online" world. Consequently, concerns have been raised about the impact this constant connectivity that smartphones afford could have on children's mental health and well-being. These concerns have been stirred largely by findings from cross-sectional studies pointing to the links between time spent online (or intensity of mobile use) and various indicators of well-being. For example, studies have found links between being a heavy (versus light) user of digital technologies and lower well-being as well as having suicide risk factors such as depression [4-6]. However, the evidence from a study applying specification curve analysis across 3 large-scale data sets (total n=355,358) concluded that the association between digital technology use and adolescent well-being is negative but small, explaining at the most 0.4% of the variation in well-being [7]. On the other hand, a systematic review of 43 original research papers on adolescents concluded that information and communication technology usage can have benefits, such as in the form of higher self-esteem and higher perceived social support, although harmful effects were also reported, such as increased exposure to social isolation and depression [8]. The authors concluded that the majority of studies reported either mixed or no effects of technologies on adolescent well-being.

Still, a mostly negative discourse is prominent in the domain of physical well-being, where negative associations with technology have been demonstrated for sleep, physical activity, or overweight/obesity in children and adolescents. Namely, concerns over excessive screen time have been linked with low physical activity, poor sleep [9-12], and a higher risk of overweight/obesity rates [13]. This combined with the facts that levels of physical activity are already low among children and the proportion of overweight/obese children has been increasing in most Western populations has led to calls for public health guidelines and limits on children's screen time, although the evidence to support them remains limited [14].

Regarding impacts on psychological or social well-being, both negative and positive associations with smartphone use have been identified. On the one hand, social network sites, which are now more accessible through smartphones, provide more opportunities for being in contact with other people, self-disclosure, and building intimate relationships, all of which are crucial for well-being [15]. On the other hand, using social network sites could be connected with negative outcomes when using these sites does not fulfill social needs such as the need for belonging or need for acceptance [16]. For example, a meta-analysis of 61 studies found a significant but weak relationship between time spent on social networking sites and depression and loneliness [17]. For lonely and socially anxious people, the use of social media is an easier way to be in touch with others [18]. However, they tend to use them excessively and more passively, which can paradoxically lead to declined well-being [19,20]. As such, it is clearly an issue in need of further exploration in order to properly educate young and vulnerable groups of people for promulgating beneficial patterns for social media use.

It is clear that the effects of mobile technologies are not uniform, with benefits conferred among some adolescents (eg, skill building among shy adolescents) and risks exacerbated among others (eg, worsening existing mental health problems). An increasing number of researchers are calling for studies that would be designed to capture the online experience more holistically and create a more nuanced picture of adolescent online experiences and their impacts [8,21]. Among the key limitations to existing studies are the relatively short duration of the studies and their reliance on self-reporting of smartphone behavior, despite related studies showing that people are poor judges of their online or smartphone use [22,23]. Longitudinal, experimental, and quasiexperimental studies that go beyond a reliance on self-reported information are required to understand how, for whom, and under what conditions adolescents' interactions with mobile technologies influence their crucially developing social relationships, brains, and bodies. State-of-the-art approaches to managing online behaviors in children and adolescents will thus increasingly rely on methodologies incorporating objective data collection and artificial intelligence tools for the automatic detection of online risks and subsequent real-time interventions toward their mitigation.

While there are a number of opportunities associated with the deployment of technology-based data collection and innovative methodologies (eg, ecological momentary assessment [EMA]), there are also considerable risks associated with such research protocols. Among the advantages are the ability to separate between- and within-person variability, and closely examine the unfolding of temporal relationships between variables, making this approach good for theory testing. Other advantages are the ability to collect data from smartphones unobtrusively (eg, through mobile apps collecting data passively in the background), supplementing such data with sensor-based or self-reporting–based data without any recall bias allowing for

a detailed granularity of data. However, such an approach is intrusive in terms of privacy for not only participants but also any third party that communicates with them, which means that special steps to ensure anonymity must be applied [24,25].

Herein, we introduce the protocol of the "Adolescents and Smartphone Use Study" that aims to investigate the associations between adolescent smartphone use and different facets of well-being (social, physical, and psychological) and, in doing so, implement innovations related to data-collection protocols so as to address some of the aforementioned empirical concerns. The study is part of the larger research project "Modelling the future: Understanding the impact of technology on adolescent's well-being" (FUTURE [26]) that aims to develop a complex evidence-based theory depicting the impacts of technology usage on the physical, psychological, and social well-being of adolescents aged 11 to 18 years. This protocol describes the key elements of the research contained in one work package utilizing intense longitudinal data collection methods and innovative research tools with artificial intelligence.

The FUTURE-WP4 Project

The overall aim of this project is to better assess the short-term and long-term impacts of smartphone use on well-being using innovative data collection approaches and the automatic recognition of adolescents' online activities in real-time. Specifically, we planned to conduct a 12-month study of adolescent smartphone behavior that would utilize a repeated measurement-burst design [27] with 4 intensive 14-day data collection periods (ie, bursts) during which a specialized software (a custom-built mobile app) would capture numerous smartphone metrics and screenshots, and would distribute short questionnaires several times per day to assess aspects of subjective well-being and self-reported smartphone behavior. The project unfolded in several stages as presented below.

Stage 1: Evaluating the Ethical Aspects of the Study (2019)

During the first year of the study, we carefully evaluated the ethical and legal implications of the proposed work. An interdisciplinary team of experts was created to evaluate every aspect of the planned activities and create a risk/benefit analysis. Since the planned data collection was to include detailed records of adolescents' online activities captured by a custom-made mobile app, we anticipated encountering numerous challenges associated with the ethics of such data collection, legal obstacles, and subsequent procedures for secure and careful data management. Therefore, the first phase of the project included a detailed examination of the possibilities for data collection and management, with the goal of creating a detailed study protocol that would comply with all ethical and legal standards. Team members who were experts in informatics, ethics, the social sciences, developmental psychology, and law met weekly for a year to discuss technology law and then the newly enacted General Data Protection Regulation policies, and to scrutinize the planned procedures, identify all problematic aspects, and help propose needed adjustments. During this phase, we also consulted with the Research Ethics Committee of Masaryk University. The output of these discussions was a final refined protocol for the subsequent processing of data collection and management deemed acceptable from ethical and legal perspectives (including the recommendation for developing a data anonymization tool, as described below). As a result of this process, the project received ethical approval from the university ethics board.

Stage 2: Software Development (2019-2020)

A fundamental aspect of the project was the development of a customized smartphone app for data collection. Designed to run in the background of the participants' devices, the app passively captures key smartphone logs and screenshots during the active collection period. It also allows for the delivery of self-report questionnaires in a flexible manner (ie, based on different schemas, eg, timed, context-based, and self-initiated triggers). The collection parameters, such as start dates and participant groups, are organized with a companion web application for researchers. The software is backed by a dedicated server and a relational database. Additionally, as per the recommendation from the legal analysis process, we developed customized anonymization software.

Optical Character Recognition and the Anonymization Software

Optical character recognition (OCR) will be used to collect unstructured text from user screenshots, especially from instant messaging apps and web browsers. This way, we circumvent the inaccessibility of the data through the Android application programming interface (API) and potential ethical and legal issues by accessing the private space of the individual app's storage. Additionally, it solves the extremely complicated task of anonymizing the screenshot images because only the extracted text will be saved, with the images themselves discarded.

Since such an automated data collection would also capture data from private (or at least nonpublic) conversations or profiles of people who had not provided consent with the study, it was necessary to develop a solution for automatic anonymization of the data during the data collection. Thus, we created software that would automatically anonymize the data during the collection and store only filtered anonymized data (ie, suppressing or masking the names, nicknames, addresses, and any other identifying information) [28].

Smartphone App for Data Collection and Questionnaires

We subsequently created an Android app (Interdisciplinary Research Team on Internet and Society [IRTIS] app) that became the primary building block for acquiring objective data from the participants. It collects various logs, such as screen activation, foreground apps, battery state, and connected Wi-Fi (for a complete list see Multimedia Appendix 1). The app also enables the delivery of questionnaires to get feedback from the participant or collect self-reported data on behavior and well-being. The app also allows the respondents to stop or pause the data collection at any time they wish. We enhanced it in the final stage of development by implementing a game-based reward system to improve the questionnaire compliance of the participants. Participants are familiarized with all features of the app (including the capturing of smartphone logs and screenshots) prior to the start of data collection.

We first released the app in November 2019. This release was focused on testing the overall function of the app on selected devices of volunteer testers. After the initial "in-house" testing phase, we tested the app more broadly through 3 pilot studies in the first half of 2020. The main focus of the pilot studies was to find and fix possible problems (especially related to screen data collection and the background running of apps) on various devices. We also introduced new features, such as messaging with our participants, new logs and their optimization to save battery life, dissemination of multiple data collection periods (referred to as "bursts"), and advanced management of questionnaires.

For the final pilot study at the end of 2020, we introduced the game-based reward system for enhancing questionnaire compliance and automated communication via messages with noncomplying participants. Participants could collect coins for completed questionnaires, which were then linked with lottery drawings for a number of prizes (eg, online vouchers). In this pilot, we also, for the first time, included questions intended for the main study, in order to test the general comprehension, reliability, and validity of the scales. The summary overview of the pilot tests and the details of the pilot testing process are described in Multimedia Appendix 2.

Stage 3: Machine Learning

One of the study goals is the development of predictive machine learning models for the automated detection, classification, and explanation of communicative behavior in the data collected from adolescents' smartphones. Specifically, we focus on 2 types of communicative behaviors that can be associated with adolescents' well-being (supportive online interactions and discussing risky behaviors online). We plan to use 2 supervised learning approaches. The first one uses structured data from the collected logs, and the other uses unstructured text from instant messenger (IM) conversations [29]. This work is in progress. We will use the data set with collected logs for the former, which is labeled by the time-matched answers in the self-reported questionnaires. The latter (IM conversations) requires schema development and manual annotation. The machine learning aspects of the project and both processes will be described in a separate manuscript.

Methods

Ethics Approval

This study was approved by the Research Ethics Committee of Masaryk University (EKV-2018-068).

Design

The main study includes a 12-month prospective study of adolescents using EMA and a repeated measurement burst design. In EMA, participants are usually prompted several times a day to answer questions, and they may be asked to self-initiate a report when an event occurs, so as to capture phenomena as they unfold in natural environments in real life. This approach reduces the risk of retrospective recall bias associated with self-reporting or recollection of behavior [30]. EMA performed on smartphones has a number of advantages, including the automated recording of the timestamp of answers in real-time, the possibility of tracking of compliance and response patterns, and the possibility of combining survey responses with other data such as metrics from the smartphone device, from other online sources, or from a connected third device (eg, a sensor). This methodology also appears highly suitable for the study of adolescent online/smartphone behavior as it allows for a minimally obtrusive repeated assessment of authentic smartphone usage in the context of daily experiences, moods, and behaviors [31-33].

The study started in May 2021. At the beginning and end of the 12-month study, participants complete online surveys assessing key susceptibility variables and long-term well-being. Across the 12 months, participants complete 4 bursts of 2-week (10 weekdays and 4 weekend days) intensive data collection of smartphone behavior assessed passively through the custom-built Android mobile app (IRTIS app) installed on participants' own smartphones, which should enhance the ecological validity of the collected data (as compared to relying on phones provided by a researcher). The app also administers short surveys 4 times per day, which assess real-time smartphone behavior and short-term changes (moment to moment, daily) in well-being. The surveys are administered in 4 predetermined time windows (6 AM-10 AM, 10 AM-3 PM, 3 PM-8 PM, and 8 PM-12 AM) on a semirandom schedule, with the exception of the morning survey. The morning survey had a default trigger time preset at 7 AM (to ensure it occurs before school), but the participants were encouraged to personalize this time with the possibility to select a different time between 6 AM and 10 AM that best fits their schedule (a different time could be set for each day of the week). All other surveys were triggered once at random within the respective time windows but always at least 1 hour apart. Upon notification, participants had 45 minutes to complete the morning and evening surveys and 90 minutes to complete the 2 daily surveys. The longer completion time for the daily surveys was chosen to ensure participants had a chance to complete the survey at the class break during school hours (a typical class period in the Czech Republic lasts 45 minutes, with 10- to 20-minute breaks in between) and to accommodate after-school activities. The surveys assess affective states, self-reported screen time, acute stressors, sexual content exposure, online vigilance, sporting and walking behaviors, and perceived social support (a list of items, item sources, and the EMA protocol can be found in Multimedia Appendix 3). The morning survey in addition has questions about sleep during the previous night. The evening survey also has retrospective questions about the day as a whole.

Additionally, participants are asked to complete a self-initiated report when something happened on the internet that left them bothered or upset (ie, uncomfortable discussions, news, pictures, or videos that left them feeling frightened or with an uneasiness afterward) through a self-initiated open-ended questionnaire. After each burst (on the 15th day), a summary postburst survey was administered. The postburst survey triggers at 7 AM, and participants have 12 hours to complete this survey.

The morning questionnaire and both daily questionnaires were designed to take under 2 to 3 minutes to complete, and the evening questionnaire was designed to take 4 to 5 minutes to complete. The actual duration of survey completion during the

first burst was as follows: morning, median 77 seconds; daily I, median 24 seconds; daily II, median 23 seconds; evening, median 110 seconds; postburst, median 291 seconds.

Participant Recruitment

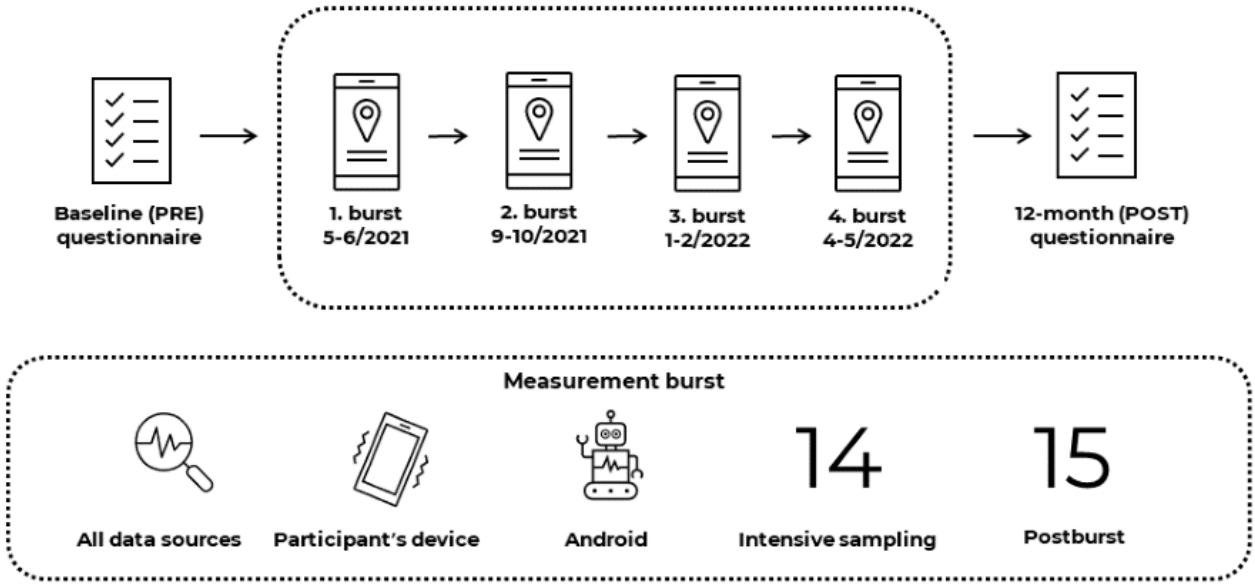
The study sample was recruited from the Czech Republic with the help of a professional social science research and marketing company selected after a market review from multiple solicited bids. The specification was to find 300 adolescents aged 13 to 17 years, who have a smartphone with Android (at least version 5 “Lollipop”), and the sample was supposed to have an equal distribution across age and gender. Due to the lack of data from other comparable studies, the sample size was selected based on the pragmatic recommendation to recruit as many participants as we had resources for [34], while taking into account the rate of missing or problematic data from our pilot studies.

An independent social research and marketing company was commissioned to assist with study compliance maintenance. The research team managed most day-to-day responsibilities with real-time compliance monitoring during active data collection bursts. The company enforced compliance between bursts and handled problematic participants. The decision to involve a professional company in this way was partially motivated by the difficulty in offering incentives to participants from a university budget within the Czech legal framework.

Materials and Procedures

Upon recruitment, participants were asked to complete an online baseline questionnaire (administered via the Qualtrics platform). Subsequently, they were given instructions to install the study mobile app from the Google Play store. Participants were provided with a written manual with step-by-step instructions on downloading and operating the app and the study procedures. Short instructional videos were also created to facilitate the learning process. This strategy was chosen based on participant feedback from prior pilot studies. Personal demonstration or training was not feasible due to the ongoing COVID-19 pandemic. Each participant received their unique credentials to sign in to the app. The app automatically navigates the participant through permissions to collect different types of data upon the first sign-in. Researchers set up the survey assignment schedule for each of the 4 data collection bursts across the 12-month study in the researcher web application interface. The questionnaire schedule and 14-day bursts across the 12-month period are depicted in Figure 1.

Figure 1. Depiction of the study design and protocol including the ecological momentary assessment bursts.



Participant Incentives

The participants receive no monetary compensation; however, they may win prizes in multiple lotteries, and the number of entries is determined by their questionnaire completion rate during the measurement burst. For each completed (at least 70% completed) questionnaire on a given day, the participant receives a bronze coin. When they complete 3 questionnaires (out of 4, or 75%), they receive a silver coin, and if they complete all 4 questionnaires, they receive a gold coin. A bronze coin is worth 1 lottery entry, a silver coin is worth 5 lottery entries, and a gold coin is worth 10 lottery entries. In total, 4 bronze coins, 1 silver coin, and 1 gold coin may be acquired in 1 day of a measurement burst, resulting in 19 entries. There is a lottery at the end of each measurement burst and at the end of the whole study. After each burst, participants may win vouchers to an electronics store in the amount ranging from 500 CZK to 2000 CZK (22 USD to 88 USD). In the final lottery at the end of the study, participants may win 1 of 2 smartphones (worth 7000 CZK and 10,000 CZK [209 USD and 442 USD, respectively]) or a Play Station 5 (13490 CZK [597 USD]).

Research Hypotheses and Analytical Approach

Our main research questions center on the associations between smartphone use and its impact on different domains of well-being. We are especially interested in how different aspects of online behavior (assessed as patterns of smartphone use from the objective data logs) impact psychological, social, and physical well-being. Our design enables tracking of daily and momentary short-term fluctuations in well-being indicators as

well as long-term change (across four bursts) and how it is influenced by patterns of smartphone app use (eg, time spent in social networking apps, mobile games, communication apps, or browsers). We are also interested in the reciprocal relationship between smartphone usage and well-being indicators so as to obtain a more nuanced view of the temporal associations between the 2 (eg, through examination of cross-lagged effects). Given the hierarchical nature of the data (moments nested within days, bursts, and persons), we plan to utilize multilevel approaches including vector autoregressive models to capture associations among change over time and to disentangle the between-person and within-person associations among outcomes of interest.

Results

Sample Description

In total, 203 adolescents were enrolled in the study. The social science research and marketing company recruited participants through their network of adult and adolescent respondents. However, given the unique type of research that requires intensive participation and poses a serious privacy intrusion, the agency was not able to find a sufficient number of participants (see CONSORT [Consolidated Standards of Reporting Trials] diagram in Figure 2). The agency invited almost 12,000 children or parents with children; however, only 180 participants provided informed consent. Thus, we supplemented the sample with participants from our own recruitment efforts through paid advertisements on social media and through chain referral, which yielded an additional 63 participants. In the end, overall, 243 participants provided informed consent, but only 203 completed all requirements (ie, provided consent [both the participant and their parent], completed the baseline questionnaire, and installed the app) and entered the study. Two participants did not provide any data (daily questionnaire data and metrics data) during the first measurement burst. A descriptive overview of the final study sample can be seen in Table 1.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram describing participant recruitment. FB: Facebook; IG: Instagram.

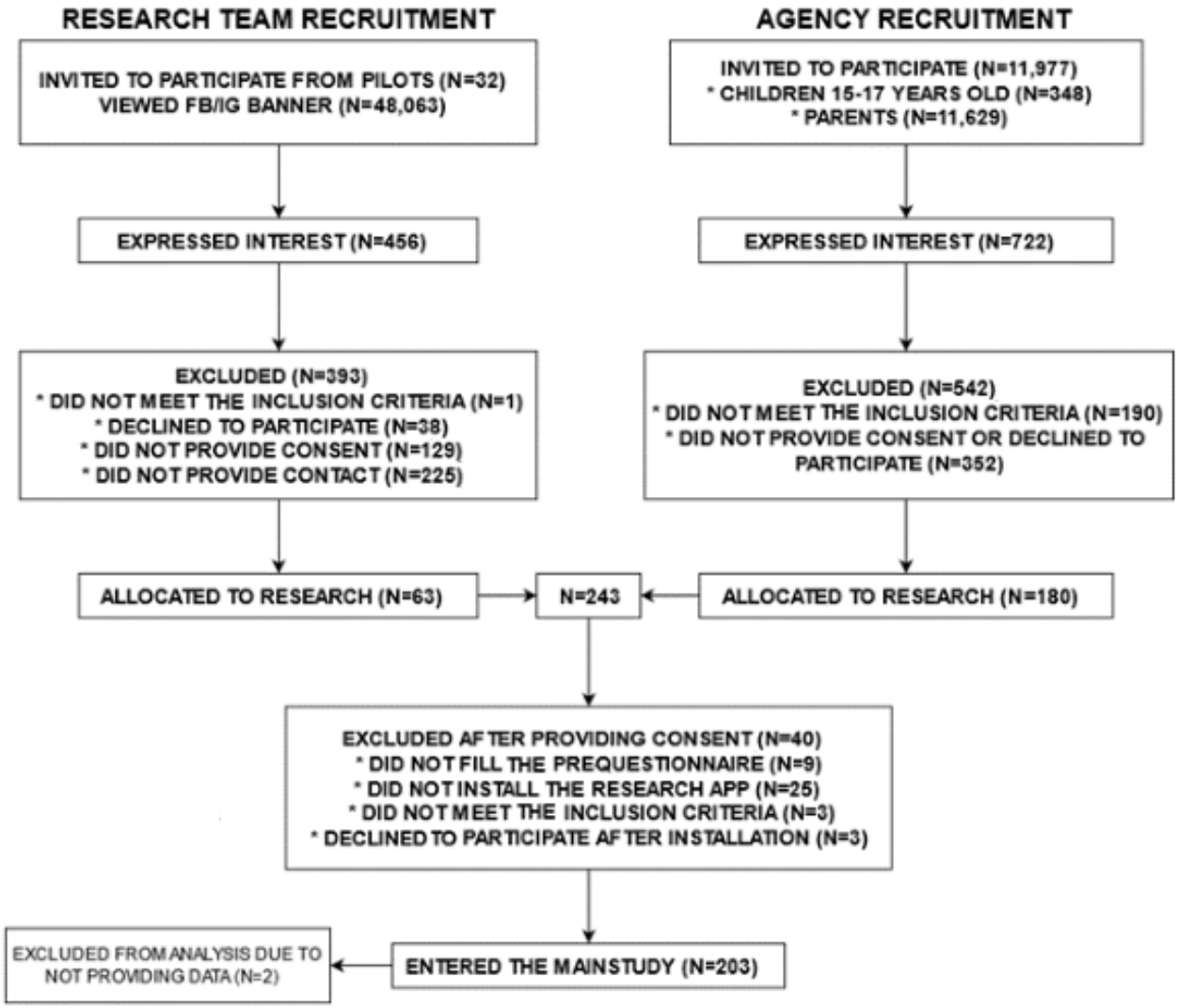


Table 1. Study sample.

Age group (years)	MEDIAN agency		IRTIS[a] team	
	Girls (n=54), n	Boys (n=89), n	Girls (n=30), n	Boys (n=30), n
13	15	25	2	0
14	13	24	0	4
15	11	21	2	5
16	12	8	9	7
17	3	11	17	14

[a]IRTIS: Interdisciplinary Research Team on Internet and Society.

Preliminary Compliance

Data collection for the first burst took place in 2 waves (2 weeks apart). The second wave was added to accommodate the schedules of some participants and the stragglers (n=21) who submitted their informed consent and baseline questionnaire after the deadline. Each wave installed the app on Thursday and started data collection on Friday. The overall completion rate of the timed (morning, daily I, daily II, and evening) questionnaires was 72.3% (6690 completed surveys [with at least one answered question] out of 9260 notified surveys). The completion rate was higher for school days (80.4%; 5356 completed surveys out of 6662 notified surveys) than for weekend days (68.13%; 1770 completed surveys out of 2598 notified surveys). There was some failure in delivering surveys (related to notifications). Despite planning for each burst to present 56 surveys (4 surveys × 14 days) for an expected collection total of 11,256 surveys, only 9260 (82.3%) were successfully notified on the participants' smartphones. The completion rate for each type of survey (as well as the corresponding "success" rate for the actually delivered surveys) is included in Table 2.

Table 2. Survey completion rates for the first burst.

Survey	Number of completed surveys (at least one answered question)	Number of notified surveys	Completion rate (number of completed surveys/number of notified surveys)	Successful delivery rate (number of notified surveys/number of planned surveys)
Morning	1613	2319	69.56% (1613/2319)	82.41% (2319/2814)
Daily I	1696	2320	73.10% (1696/2320)	82.44% (2320/2814)
Daily II	1802	2336	77.14% (1802/2336)	83.01% (2336/2814)
Evening	1579	2285	69.10% (1579/2285)	81.20% (2285/2814)
Postburst[a]	149	165	90.30% (149/165)	82.09% (165/201)
Self-initiated	21	N/A[b]	N/A	N/A

[a]Values calculated based on n=201 (ie, without 2 participants who did not provide any data).

[b]N/A: not applicable.

Discussion

This paper introduces the protocol of a study focusing on the smartphone behavior of adolescents and its impact on physical, psychological, and social well-being. In our study, we combine passive data collection of smartphone logs with the intensive assessment of self-reported states and behaviors through a mobile app. The design of our study (a prospective measurement-burst design) will allow us to capture both short-term variability in outcomes of interest as well as long-term change. We apply the EMA methodology, which has been shown to be feasible and productive in studies of adolescents [31,35]. When supplemented with our app, this methodology allows us to link self-reported data with data from objective smartphone logs, affording the analysis of the temporality of effects at both the within- and between-person levels. Concretely, this means being able to compare the impact of different levels and patterns of smartphone use (eg, heavy versus light usage) on well-being, as well as assessing how changes in smartphone behavior over time reflect on well-being (eg, whether spending more time online on one's smartphone than usual leads to increased or decreased well-being).

Collecting smartphone log data is a strength of our approach. This allows for the collection of objective data with minimum demands on study participants (ie, data are collected automatically and passively in the background). However, studies from other fields suggest that when it comes to behavior prediction, there may be a trade-off between accuracy or data details and participant burden. For example, in studies of dietary behavior, prediction models using collected sensor-based data (posing minimum participant burden) have resulted in lower accuracy than models using self-reported data through EMA prompts when predicting dietary lapses [36]. The extent to which smartphone log data will successfully predict behavior or well-being will likely depend on the operationalization of the objective log data. For example, when assessing smartphone

behavior, one must take into account an entire spectrum of behaviors that may have different impacts. A user may passively scroll the news feed or engage actively with other people in the comments section. In each case, we may expect a different impact on user psychological outcomes [37]. Thus, different "metrics" must be generated from the smartphone log data to capture these different aspects of behavior.

Some behaviors may also be more difficult to operationalize using smartphone log data. Consider, for example, the question of how to accurately capture what exactly adolescents are doing when spending time in a specific social networking app. While we may be able to effectively quantify the time spent in social networking apps, it is more challenging to capture what happens during the social interactions while on the app. Moreover, this may be crucial when evaluating the association between app use and psychological and social well-being [38]. For example, a study using momentary sampling techniques showed that individuals who are involved in a greater amount of supportive interactions with others feel more positive emotions after these interactions and report more perceived social support [39]. While we are unable to directly evaluate the "quality" of online interactions, the OCR tool we are developing for the analyses of screenshots from mobile phones along with the application of machine learning algorithms could lead to a more nuanced assessment of adolescent smartphone behavior in future research. Other limitations associated with the smartphone log data include constraints applied directly by Google Play policies, which preclude the collection of some data (eg, GPS and web-browsing histories). The use of web browsers may result in exposure to very different types of content (eg, educational and harmful), thus leading to different types of psychological outcomes.

A critical component to interpreting the data from our study will be the consideration of selection bias and how the resulting sample differs from the general adolescent population of smartphone users. While the study was not planned to be representative, it will be key to understand in what ways our participants differ from the general population. The recruitment process posed a big challenge in spite of the extensive experience that the IRTIS team has with large-scale studies of children and adolescents. We attribute this primarily to the sensitive nature of the data collected and the perceived intrusiveness into the privacy of the adolescents. Despite carefully considering the ethical implications of the data collection, creating detailed descriptions of data handling and safety procedures (for both parents and children), developing anonymization software, and having received the approval of relevant ethical bodies, some participants reported distrust and intrusion of privacy among the reasons for nonparticipation. The social research and marketing agency tasked with the recruitment of participants in our study provided feedback and recommendations for the future, including offering financial incentives to all participants (not only through lottery drawings), extending the inclusion criteria to iOS users, and further alleviating privacy concerns through modifications in the research design.

The development of the custom mobile app spanned a period of 2 years, and in spite of extensive pilot testing, we were unable to fully eliminate technical issues associated with data collection. For example, there was a 17.74% failure rate of surveys not being properly delivered to the users' devices (only 9425 surveys out of 11,457 planned surveys were properly notified). This is a problem primarily from the standpoint of inflating missing data. Even when a survey notification is activated, it is common for users to not notice the notification or simply disregard it. While certain strategies can help enhance response rates (eg, incentives and gamification [40]), the lack of responses due to technical problems/failure should be minimized. We found that problems of failed notifications were more common in certain types of devices (especially the smartphones of some manufacturers, such as Huawei and Xiaomi). Additionally, ongoing optimization of the Android system and its updates necessitate ongoing technical and programming support throughout all phases of the study. Our app was downloadable through Google Play, where we also faced privacy policy limitations on data collection and therefore had to modify our app throughout the pilot testing phase (eg, deleting the GPS location component).

Although there were some challenges with the application of EMA (eg, adapting the EMA protocol to fit youths' school schedules and complications related to the ongoing COVID-19 pandemic such that we had to incorporate additional questions in the postburst survey regarding school attendance/distance learning), data from the first burst indicated acceptable levels of compliance with the study protocol. The average compliance rate in the first burst in our study was in line with other EMA studies of adolescents where compliance varies widely from 51.56% to 92.00% in studies of psychological outcomes [41] or 43.8% to 95.9% in studies of health behaviors [42]. In a meta-analysis of EMA studies targeting children and adolescents [43], the compliance rate in studies prompting participants 4 to 5 times per day was 77.4%, which is only slightly higher than in our study (72.25%). The data are however not directly comparable, since the majority of studies included in the meta-analysis were short-term (did not plan to involve more than one data collection burst) or used research devices that were novel for users, perhaps inflating compliance estimates. In our study, we relied on participants' own devices. As is recommended in EMA studies [31,44], we continually monitored study adherence with a system of both automatic notifications and reminders (in-app), as well as telephone, SMS, and email check-ins and follow-up contacts when necessary. We also provided compliance-based incentives to support participant engagement. Additional strategies are being considered to boost adherence in follow-up bursts (eg, bolstering contact with participants in-between bursts through activities such as a Christmas competition).

The challenges surrounding accurate data collection related to real-time smartphone usage remain, especially in light of continual technological or software innovation and the ever-morphing proclivities of user engagement. As such, it demands increased vigilance relative to the processes by which research in this realm is conducted. Studies that seek to employ innovative, nuanced, and more comprehensive approaches to the protocol of collecting the said data are, therefore, of paramount importance. Further refinement of research

instruments, protocols, and methodologies is needed to obtain a more accurate portrait of how adolescents are actually engaging with their online worlds. This was and remains our goal, as we seek to develop, produce, and implement procedures that will more effectively assist adolescents to navigate these worlds more constructively in terms of the development of their physical, social, and psychological well-being.

Acknowledgments

This work has received funding from the Czech Science Foundation (project number 19-27828X).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Objective smartphone data codebook.
[DOCX File , 22 KB - resprot_v11i3e35984_app1.docx]

Multimedia Appendix 2
Pilot studies in the Adolescents and Smartphone Use Study.
[DOCX File , 24 KB - resprot_v11i3e35984_app2.docx]

Multimedia Appendix 3
Ecological momentary assessment study protocol for the Adolescents and Smartphone Use Study.
[DOCX File , 982 KB - resprot_v11i3e35984_app3.docx]

Multimedia Appendix 4
Peer-review report by the Czech Science Foundation.
[PDF File (Adobe PDF File), 187 KB - resprot_v11i3e35984_app4.pdf]

References

1. Social Media, Social Life: Teens Reveal Their Experiences, 2018. Common Sense Media. URL: https://www.commonsensemedia.org/research/social-media-social-life-2018 [accessed 2022-02-09]
2. EU Kids Online 2020: survey results from 19 countries. LSE Research Online. URL: https://eprints.lse.ac.uk/103294/ [accessed 2022-02-09]
3. George MJ, Odgers CL. Seven fears and the science of how mobile technologies may be influencing adolescents in the digital age. Perspect Psychol Sci 2015 Nov 17;10(6):832-851 [FREE Full text] [doi: 10.1177/1745691615596788] [Medline: 26581738]
4. Twenge JM, Campbell WK. Media use is linked to lower psychological well-being: evidence from three datasets. Psychiatr Q 2019 Jun 11;90(2):311-331. [doi: 10.1007/s11126-019-09630-7] [Medline: 30859387]
5. Twenge JM, Martin GN, Campbell WK. Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of smartphone technology. Emotion 2018 Sep;18(6):765-780. [doi: 10.1037/emo0000403] [Medline: 29355336]
6. Liu J, Liu C, Wu T, Liu B, Jia C, Liu X. Prolonged mobile phone use is associated with depressive symptoms in Chinese adolescents. J Affect Disord 2019 Dec 01;259:128-134. [doi: 10.1016/j.jad.2019.08.017] [Medline: 31450134]
7. Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. Nat Hum Behav 2019 Feb 14;3(2):173-182. [doi: 10.1038/s41562-018-0506-1] [Medline: 30944443]
8. Best P, Manktelow R, Taylor B. Online communication, social media and adolescent wellbeing: A systematic narrative review. Children and Youth Services Review 2014 Jun;41:27-36. [doi: 10.1016/j.childyouth.2014.03.001]
9. Munezawa T, Kaneita Y, Osaki Y, Kanda H, Minowa M, Suzuki K, et al. The association between use of mobile phones after lights out and sleep disturbances among Japanese adolescents: a nationwide cross-sectional survey. Sleep 2011 Aug 01;34(8):1013-1020 [FREE Full text] [doi: 10.5665/SLEEP.1152] [Medline: 21804663]
10. Lemola S, Perkinson-Gloor N, Brand S, Dewald-Kaufmann JF, Grob A. Adolescents' electronic media use at night, sleep disturbance, and depressive symptoms in the smartphone age. J Youth Adolesc 2015 Feb 10;44(2):405-418. [doi: 10.1007/s10964-014-0176-x] [Medline: 25204836]
11. Hale L, Guan S. Screen time and sleep among school-aged children and adolescents: a systematic literature review. Sleep Med Rev 2015 Jun;21:50-58 [FREE Full text] [doi: 10.1016/j.smrv.2014.07.007] [Medline: 25193149]

12. Lund L, Sølvhøj IN, Danielsen D, Andersen S. Electronic media use and sleep in children and adolescents in western countries: a systematic review. BMC Public Health 2021 Sep 30;21(1):1598 [FREE Full text] [doi: 10.1186/s12889-021-11640-9] [Medline: 34587944]
13. Fang K, Mu M, Liu K, He Y. Screen time and childhood overweight/obesity: A systematic review and meta-analysis. Child Care Health Dev 2019 Sep;45(5):744-753. [doi: 10.1111/cch.12701] [Medline: 31270831]
14. Ashton J, Beattie R. Screen time in children and adolescents: is there evidence to guide parents and policy? Lancet Child Adolesc Health 2019 May;3(5):292-294. [doi: 10.1016/S2352-4642(19)30062-8] [Medline: 30853301]
15. Chen H, Li X. The contribution of mobile social media to social capital and psychological well-being: Examining the role of communicative use, friending and self-disclosure. Computers in Human Behavior 2017 Oct;75:958-965. [doi: 10.1016/j.chb.2017.06.011]
16. Clark J, Algoe S, Green M. Social network sites and well-being: the role of social connection. Curr Dir Psychol Sci 2017 Dec 19;27(1):32-37 [FREE Full text] [doi: 10.1177/0963721417730833]
17. Huang C. Time spent on social network sites and psychological well-being: a meta-analysis. Cyberpsychol Behav Soc Netw 2017 Jun;20(6):346-354. [doi: 10.1089/cyber.2016.0758] [Medline: 28622031]
18. Song H, Zmyslinski-Seelig A, Kim J, Drent A, Victor A, Omori K, et al. Does Facebook make you lonely?: A meta analysis. Computers in Human Behavior 2014 Jul;36:446-452. [doi: 10.1016/j.chb.2014.04.011]
19. Aalbers G, McNally R, Heeren A, de Wit S, Fried E. Social media and depression symptoms: A network perspective. J Exp Psychol Gen 2019 Aug;148(8):1454-1462. [doi: 10.1037/xge0000528] [Medline: 30507215]
20. Verduyn P, Lee DS, Park J, Shablack H, Orvell A, Bayer J, et al. Passive Facebook usage undermines affective well-being: Experimental and longitudinal evidence. J Exp Psychol Gen 2015 Apr;144(2):480-488. [doi: 10.1037/xge0000057] [Medline: 25706656]
21. Meier A, Reinecke L. Computer-mediated communication, social media, and mental health: a conceptual and empirical meta-review. Communication Research 2020 Oct 21;48(8):1182-1209 [FREE Full text] [doi: 10.1177/0093650220958224]
22. Felisoni D, Godoi A. Cell phone usage and academic performance: An experiment. Computers & Education 2018 Feb;117:175-187. [doi: 10.1016/j.compedu.2017.10.006]
23. Hodes L, Thomas K. Smartphone screen time: inaccuracy of self-reports and influence of psychological and contextual factors. Computers in Human Behavior 2021 Feb;115:106616. [doi: 10.1016/J.CHB.2020.106616]
24. Tsavli M, Efraimidis PS, Katos V. Reengineering the user: privacy concerns about personal data on smartphones. Information and Computer Security 2015;23(4):394-405. [doi: 10.1108/ICS-10-2014-0071]
25. Keusch F, Struminskaya B, Antoun C, Couper M, Kreuter F. Willingness to participate in passive mobile data collection. Public Opin Q 2019 Jul;83(Suppl 1):210-235 [FREE Full text] [doi: 10.1093/poq/nfz007] [Medline: 31337924]
26. Modeling the future: Understanding the impact of technology on adolescent's well-being (FUTURE). IRTIS. URL: https://irtis.muni.cz/research/projects/future [accessed 2022-02-11]
27. Sliwinski M. Measurement-burst designs for social health research. Social and Personality Psychology Compass 2008;2(1):245-261. [doi: 10.1111/j.1751-9004.2007.00043.x]
28. Sotolář O, Plhák J, Šmahel D. Towards Personal Data Anonymization for Social Messaging. In: Ekštein K, Pártl F, Konopík M, editors. Text, Speech, and Dialogue. TSD 2021. Lecture Notes in Computer Science, vol 12848. Cham: Springer; 2021:281-292.
29. Sotolár O, Plhák J, Tkaczyk M, Lebedíková M, Smahel D. Detecting Online Risks and Supportive Interaction in Instant Messenger Conversations using Czech Transformers. RASLAN 2021. URL: https://is.muni.cz/auth/publication/1809721/cs?lang=en;setlang=en [accessed 2022-02-09]
30. Stone A, Shiffman S, Atienza A, Nebeling L. The Science of Real-Time Data Capture: Self-Reports in Health Research. New York, NY: Oxford University Press; 2007.
31. Heron KE, Everhart RS, McHale SM, Smyth JM. Using Mobile-Technology-Based Ecological Momentary Assessment (EMA) Methods With Youth: A Systematic Review and Recommendations. J Pediatr Psychol 2017 Nov 01;42(10):1087-1107. [doi: 10.1093/jpepsy/jsx078] [Medline: 28475765]
32. Reeves B, Ram N, Robinson TN, Cummings JJ, Giles CL, Pan J, et al. Screenomics: a framework to capture and analyze personal life experiences and the ways that technology shapes them. Hum Comput Interact 2021 Mar 13;36(2):150-201 [FREE Full text] [doi: 10.1080/07370024.2019.1578652] [Medline: 33867652]
33. Domoff S, Banga C, Borgen A, Foley R, Robinson C, Avery K, et al. Use of passive sensing to quantify adolescent mobile device usage: Feasibility, acceptability, and preliminary validation of the application. Human Behav and Emerg Tech 2021 Jan 05;3(1):63-74. [doi: 10.1002/hbe2.247]
34. Albers C, Lakens D. When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. Journal of Experimental Social Psychology 2018 Jan;74:187-195. [doi: 10.1016/J.JESP.2017.09.004]
35. Russell MA, Gajos JM. Annual Research Review: ecological momentary assessment studies in child psychology and psychiatry. J Child Psychol Psychiatry 2020 Mar 29;61(3):376-394 [FREE Full text] [doi: 10.1111/jcpp.13204] [Medline: 31997358]

36. Crochiere R, Zhang F, Juarascio A, Goldstein S, Thomas J, Forman E. Comparing ecological momentary assessment to sensor-based approaches in predicting dietary lapse. Transl Behav Med 2021 Dec 14;11(12):2099-2109. [doi: 10.1093/tbm/ibab123] [Medline: 34529044]
37. Bayer J, Ellison N, Schoenebeck S, Brady E, Falk E. Facebook in context(s): Measuring emotional responses across time and space. New Media & Society 2016 Dec 27;20(3):1047-1067 [FREE Full text] [doi: 10.1177/1461444816681522]
38. Sun H, Harris K, Vazire S. Is well-being associated with the quantity and quality of social interactions? J Pers Soc Psychol 2020 Dec 11;119(6):1478-1496. [doi: 10.1037/pspp0000272] [Medline: 31647273]
39. Oh HJ, Ozkaya E, LaRose R. How does online social networking enhance life satisfaction? The relationships among online supportive interaction, affect, perceived social support, sense of community, and life satisfaction. Computers in Human Behavior 2014 Jan;30:69-78. [doi: 10.1016/j.chb.2013.07.053]
40. van Berkel N, Goncalves J, Hosio S, Kostakos V. Gamification of mobile experience sampling improves data quality and quantity. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol 2017 Sep 11;1(3):1-21. [doi: 10.1145/3130972]
41. van Roekel E, Keijsers L, Chung JM. A review of current ambulatory assessment studies in adolescent samples and practical recommendations. J Res Adolesc 2019 Sep 17;29(3):560-577 [FREE Full text] [doi: 10.1111/jora.12471] [Medline: 31573762]
42. Liao Y, Skelton K, Dunton G, Bruening M. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted STROBE Checklist for Reporting EMA Studies (CREMAS). J Med Internet Res 2016 Jun 21;18(6):e151 [FREE Full text] [doi: 10.2196/jmir.4954] [Medline: 27328833]
43. Wen C, Schneider S, Stone A, Spruijt-Metz D. Compliance with mobile ecological momentary assessment apotocols in children and adolescents: a systematic review and meta-analysis. J Med Internet Res 2017 Apr 26;19(4):e132 [FREE Full text] [doi: 10.2196/jmir.6641] [Medline: 28446418]
44. Degroote L, DeSmet A, De Bourdeaudhuij I, Van Dyck D, Crombez G. Content validity and methodological considerations in ecological momentary assessment studies on physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2020 Mar 10;17(1):35 [FREE Full text] [doi: 10.1186/s12966-020-00932-9] [Medline: 32151251]

Abbreviations

API: application programming interface
EMA: ecological momentary assessment
IM: instant messenger
IRTIS: Interdisciplinary Research Team on Internet and Society
OCR: optical character recognition

Edited by T Derrick; submitted 26.12.21; this is a non–peer-reviewed article;accepted 02.01.22; published 08.03.22.

Please cite as:
Elavsky S, Blahošová J, Lebedíková M, Tkaczyk M, Tancos M, Plhák J, Sotolář O, Smahel D
Researching the Links Between Smartphone Behavior and Adolescent Well-being With the FUTURE-WP4 (Modeling the Future: Understanding the Impact of Technology on Adolescent's Well-being Work Package 4) Project: Protocol for an Ecological Momentary Assessment Study
JMIR Res Protoc 2022;11(3):e35984
URL: https://www.researchprotocols.org/2022/3/e35984
doi:10.2196/35984
PMID:35258467

Protocol

HIV Surveillance and Research for Migrant Populations: Protocol Integrating Respondent-Driven Sampling, Case Finding, and Medicolegal Services for Venezuelans Living in Colombia

Andrea L Wirtz[1,2], MHS, PhD; Kathleen R Page[1,2,3], MD; Megan Stevenson[1], MSPH; José Rafael Guillén[4], MSc; Jennifer Ortíz[4], MSc, MD; Jhon Jairo López[4], MSc; Jhon Fredy Ramírez[4], BSN; Cindy Quijano[4], MSc; Alejandra Vela[4], BSW; Yessenia Moreno[4], BSW; Francisco Rigual[4], JD; James Case[5], MBI; Avi J Hakim[6], MA, MPH; Wolfgang Hladik[6], MD, PhD; Paul B Spiegel[2], MD, MPH

[1]Center for Public Health and Human Rights, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[2]Center for Humanitarian Health, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[3]Division of Infectious Diseases, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[4]Red Somos, Bogotá, Colombia

[5]Johns Hopkins School of Nursing, Baltimore, MD, United States

[6]Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Andrea L Wirtz, MHS, PhD
Center for Public Health and Human Rights
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe St
Baltimore, MD, 21205
United States
Phone: 1 4109292881
Email: awirtz1@jhu.edu

Abstract

Background: Epidemiologic research among migrant populations is limited by logistical, methodological, and ethical challenges, but it is necessary for informing public health and humanitarian programming.

Objective: We describe a methodology to estimate HIV prevalence among Venezuelan migrants in Colombia.

Methods: Respondent-driven sampling, a nonprobability sampling method, was selected for attributes of reaching highly networked populations without sampling frames and analytic methods that permit estimation of population parameters. Respondent-driven sampling was modified to permit electronic referral of peers via SMS text messaging and WhatsApp. Participants complete sociobehavioral surveys and rapid HIV and syphilis screening tests with confirmatory testing. HIV treatment is not available for migrants who have entered Colombia through irregular pathways; thus, medicolegal services integrated into posttest counseling provide staff lawyers and legal assistance to participants diagnosed with HIV or syphilis for sustained access to treatment through the national health system. Case finding is integrated into respondent-driven sampling to allow partner referral. This study is implemented by a local community-based organization providing HIV support services and related legal services for Venezuelans in Colombia.

Results: Data collection was launched in 4 cities in July and August 2021. As of November 2021, 3105 of the target 6100 participants were enrolled, with enrollment expected to end by February/March 2022.

Conclusions: Tailored methods that combine community-led efforts with innovations in sampling and linkage to care can aid in advancing health research for migrant and displaced populations. Worldwide trends in displacement and migration underscore the value of improved methods for translation to humanitarian and public health programming.

International Registered Report Identifier (IRRID): DERR1-10.2196/36026

(JMIR Res Protoc 2022;11(3):e36026) doi:10.2196/36026

KEYWORDS

HIV; epidemiology; migrant; Venezuela; Colombia; respondent-driven sampling; case finding; HIV treatment; HIV surveillance; research

Introduction

The economic crisis and political instability in the Bolivarian Republic of Venezuela has led to mass migration in the Americas, displacing approximately 5.4 million Venezuelans as of September 2020, according to the most recent estimates [1]. This humanitarian emergency has been associated with deteriorating health care infrastructure and worsening health outcomes among Venezuelans living in the country, as well as among those displaced to neighboring countries [2]. The re-emergence of previously controlled infectious diseases and resurgence of endemic diseases have been documented and raised concerns of a spillover effect to neighboring countries [2-8]. The COVID-19 pandemic has exacerbated public health concerns and strained the capacity of the receiving countries to meet the health care needs of Venezuelan migrants [9,10].

Gaps in HIV diagnostics and treatment in Venezuela since 2015 have limited the availability of reliable estimates of HIV burden. In 2018, the Pan American Health Organization (PAHO) estimated that 69,308 people living with HIV (PLHIV), 87% of whom were registered to receive antiretroviral therapy (ART), were not receiving them owing to nationwide drug shortages [11]. A coordinated response led by PAHO has improved ART coverage [12], although diagnosis, treatment, and suppression remain suboptimal. The Joint United Nations Program on HIV/AIDS estimates that 100,000 people were living with HIV in Venezuela in 2020, with 71% of PLHIV diagnosed and 55% of those diagnosed receiving ART [13]. No data on virologic suppression rates are available [13]. Less than one-third (30%) of pregnant women living with HIV received ART for prevention of maternal-to-child transmission [13]. Access to HIV treatment for displaced Venezuelans in receiving countries is variable and depends on the national health programs and policies of the host country. Data from other studies show that migrant populations, regardless of the situation or motivation for migration, often face delays to care and have higher risk of AIDS-defining events than nonmigrant populations [14]. Treatment interruptions, including partial or intermittent treatment, can lead to virologic rebound and increase the risk of onward transmission and acquired resistance [2]. Diagnostic delays due to lack of HIV-testing capacity, including in pregnant women, can also lead to ongoing transmission. These concerns, coupled with an estimated 25,000 Venezuelans crossing the Colombian border per day at the peak of the exodus [15,16], underscore the importance of implementing appropriate surveillance methods coupled with access to HIV diagnosis, treatment, and care for migrants.

Colombia currently receives the largest number of displaced Venezuelans in the region. As of February 2021, approximately 1.7 million were living in Colombia [17-19]. Treatment for Venezuelans with irregular migrant status, that is, those who have entered the country outside of official or regular migration channels, is not available through the health system with the exception of prenatal care. Drug donations have made treatment available in Cúcuta, a border city in Colombia, with many Venezuelans living there or crossing the border temporarily to access treatment [20,21]. In other areas of the country, treatment options for migrants with irregular status are limited although several organizations provide HIV testing, support services, and prevention for Venezuelan migrants. Population-based estimates of HIV are absent, but they are needed to inform treatment distribution plans for future drug donations [20] and national health programming. Migrant populations are often excluded or not classified by public health and disease surveillance methods. Traditional epidemiologic surveillance efforts among displaced populations are challenged by lack of sampling frames, mobility, and ethical concerns [22,23]. Migrants do not always reside in well-defined geographic spaces, are frequently dispersed within host communities, and may move multiple times before settling in an area, all of which limits implementation of traditional probability sampling approaches for surveillance. Finally, ethical considerations to protect participants, mitigate stigma, and ensure linkage to HIV care in settings where treatment is not regularly available add further logistical challenges to such surveillance methods [22,23].

This paper describes a protocol for community-led HIV surveillance among Venezuelan migrants residing in Colombia. The protocol expands upon a network-based sampling method by integrating case finding and linkage to care through medicolegal partnerships. The findings aim to inform local treatment distribution plans [20] and country-level and regional HIV programming for migrants.

Methods

Design

The BIENVENIR Project (Bienestar de Venezolanos quienes son Inmigrantes y Refugiados) is a cross-sectional design that uses a hybrid sampling and case finding approach, coupled with medicolegal services to link individuals with HIV diagnosis to HIV treatment and care, regardless of migration status. *Red Somos*, a community-based organization, leads this implementation. Staff members are nationals of Venezuela, Colombia, or both, and have expertise in HIV testing, ancillary services, and linkage to care; legal services related to migration; psychology and social work; and community strengthening. This study is conducted among newly arrived Venezuelans living in 4 cities in Colombia. Study findings will generate estimates of HIV prevalence among adults who have arrived in Colombia since 2015, as well as qualitative estimates of engagement along the HIV care continuum among Venezuelan PLHIV. A qualitative, formative research phase was conducted to assess barriers to HIV case and health services in Colombia and to inform the quantitative research methods.

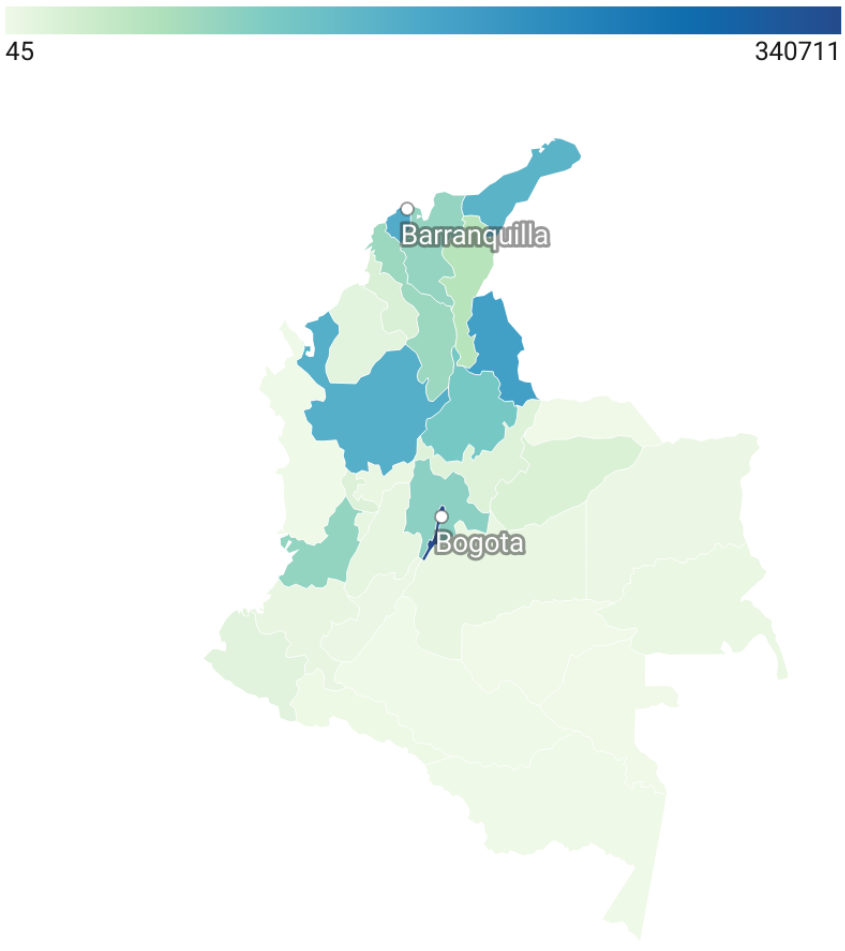
Formative Research

Key informant interviews with stakeholders (n=29), including humanitarian and health providers, government officials, and medical providers, were conducted in English and Spanish between June and October 2020. In-depth interviews (n=31) and 1 focus group discussion (n=9) with Venezuelans living in Colombia were conducted in Spanish between April 2021 to June 2021. Data collection was conducted remotely by phone or video teleconference to reduce COVID-19 transmission risks. Formative research served to provide contextual information about the humanitarian situation and programming; availability of HIV prevention and care for Venezuelans in Colombia; the impact of COVID-19 pandemic on these issues and research; and to inform decisions related to the incentives, sampling, and development of survey measures. Qualitative findings will also be used to guide subsequent interpretation of surveillance findings. Finally, formative interviews helped to ensure the study is culturally relevant and appropriate.

Sample and Setting

Data collection activities are conducted in 2 territories, encompassing the neighboring cities of (1) Bogotá and Soacha and (2) Barranquilla and Soledad (Figure 1) [24]. Locations were selected for the distribution and heterogeneous profiles of Venezuelan migrants, accessibility to humanitarian and health programs, plans for treatment distribution, and lower presence of pendulares (Venezuelans who live in Venezuela but who cross to Colombia regularly to access services) and caminantes (Venezuelans who transit through Colombia to another country). One office was established in each of the 4 cities. All adult Venezuelan nationals (age ≥18 years) who recently migrated to Colombia are eligible to participate. To ensure recruitment depth in the network of Venezuelans, only 1 member in an immediate family is eligible to participate. Inclusion criteria are as follows: Venezuelan national based on self-report (proof/documentation of nationality is not requested), born in Venezuela based on self-report, age ≥18 years, migrated to Colombia as of 2015 or later, currently residing (ie, spends most of their nights) in the study city, and has a valid study coupon to enrollment (except seeds). Participants with any of the following characteristics are excluded from participation: previous participation, have an immediate family member in the same household who participated, currently resides outside of Colombia, reports being in transit through Colombia (ie, reports an immediate destination outside of Colombia), or lacks capacity to consent. Enrolled participants are asked to provide their names and mobile phone or WhatsApp numbers for recontact and for identification of duplicate participants.

Figure 1. Distribution of Venezuelan migrants in Colombia by department. Source: Colombian Department of Migration, 2021.



Sampling

Respondent-driven sampling (RDS), a chain referral sampling method that employs limited referrals within peer networks to achieve target sample sizes, is used to accrue the study sample. RDS is widely used across international settings to sample populations that lack a sampling frame. Research has shown that with sufficient recruitment depth, biases associated with initial peer referrals are minimized [25-27], and there is an increased ability to identify previously undiagnosed HIV infections and PLHIV who have fallen out of care [27,28]. RDS thus provides an opportunity to generate unbiased estimates as well as to link surveillance with HIV clinical services. Estimation methods generate survey weights based on participant network size, and network features have been developed to calculate population-based prevalence estimates [25]. RDS has been previously used among migrants and displaced populations to a limited degree in international settings [29-33] as well as among key populations affected by HIV in Colombia [34-36].

Sampling commenced at the end of July and August 2021 in Bogotá/Soacha and Barranquilla/Soledad, respectively, and is ongoing. Recruitment started with 19 "seeds" (9-10 per territory)—well-networked individuals who were selected from the target population. Seeds were purposively selected on the basis of being well-respected and influential among peers, socially networked (know at least 10 Venezuelans outside of their household), and diverse in characteristics (eg, age, gender, geographic residence within each city). To minimize cluster effects, we identified and enrolled seeds who did not know each other and who likely did not have overlapping networks. Additional seeds may be initiated at a later date if prior seeds fail to produce peer referrals or if recruitment slows.

Seeds participate in all study activities and are asked to invite 4 adult Venezuelan peers (recruits) to participate in study activities, which is the first sampling wave. Eligible and participating recruits are then asked to refer up to 4 more peer Venezuelans. At the end of each study visit, participants undergo a brief training on how to distribute coupons and refer peers to the study. Participants have the option to use paper or electronic coupons via SMS text messages or WhatsApp to refer peers. Coupons contain study contact information and unique codes that anonymously link seeds/recruiters to recruits for analysis. Although documentation of regular migration status is necessary to acquire a phone in Colombia, 70% of Venezuelan migrants in Colombia report using mobile phones [37], and anecdotal reports support the use of mobile technology. Participants receive automated SMS text messages or WhatsApp notifications to remind them to distribute coupons or to notify them when coupons have been used and they can retrieve their incentive. As is typical in RDS research, participants are provided with a secondary incentive (COP 10,000 or USD 2.60) for each eligible and participating referred peer. Participants return within days to weeks to obtain their secondary incentive and are asked at that time to complete a brief survey about their experience of referring peers and potential biases in the referral process.

RDS is monitored in real time by using the RDS-Analyst platform [38] to ensure that sampling has reached appropriate recruitment depth (waves) to provide unbiased population-based estimates. The data management team monitors for convergence, bottlenecks, homophily, and population proportions for key indicators, including HIV infection and socially salient variables [25]. Monitoring is also used to identify whether any RDS networks may oversample certain key populations. Chains that continue to sample the same group (including men who have sex with men, people who inject drugs or engaged in sex work, and transgender populations) for more than 2 waves without the entry of other populations will be stopped. We aim to achieve a recruitment depth of at least 8-10 waves to minimize bias associated with the initial seed selection.

Data Collection

Participation consists of a sociobehavioral survey and dual rapid testing for HIV and syphilis infection, followed by confirmatory testing. At entry, participants undergo screening for eligibility and written consent procedures in a private office space. Participants completed a literacy screener using the Spanish language version of the Rapid Estimate of Adult Literacy in Medicine–Short Form [39] prior to completing a self-administered electronic survey questionnaire or an interviewer-administered option depending on the literacy score. Participants with less frequent technology use can also request an interviewer-administered survey. Following best practices in survey research, we used a mode-enhancement construction, which develops the data collection instruments to be optimized for the main mode (here, self-administered surveys given high literacy) with the auxiliary mode (interviewer-administered) designed to be equivalent [40,41]. Participants who complete the survey and dual rapid test are provided an incentive of COP 30,000 (USD 7.85).

We use a secure system for managing participant tracking and data within the study. This system was developed internally for observational (including RDS) and clinical trial research and was customized to this study [42,43]. The system is a web application and accessible in English and Spanish languages. User access to the system is location- and role-based and protected with a username and password. Entered data are encrypted in transit and stored in a secure server at Johns Hopkins. Study staff can register a participant and follow the status of the participant with respect to the completion of study activities, generate RDS coupons, send notifications, and track receipt of primary and secondary incentives. The system has built-in algorithms to check for duplicate participant registrations and to validate returned coupons for eligibility. All exported data are automatically stripped of identifiers and are linked to a participant via a unique study identifier.

Survey Measures

Survey measures included individual, social, and structural domains, drawing upon previously developed measures, as applicable (Table 1) [26,44-58]. Other health indicators beyond HIV prevention and care measures are included for assessing overall health status, identifying other health concerns that may particularly affect PLHIV (eg, malnutrition), and identifying correlates of HIV infection. Inclusion of other health measures

also helps to minimize the stigma associated with an HIV survey questionnaire. All study measures and participant materials are translated into Spanish and back translated for quality.

Table 1. Domains and measures included in the survey questionnaire.

Domain	Measures
Demographics	• Basic demographics adapted from the Colombia Demographic Health Survey [52] • Food (in)security as measured by the US Food and Drug Administration food security scale [44]
Migration and displacement	• Displacement history (timing, location of residence in Colombia, and migration status, eg, regular or irregular)
Health	• Recent health history • Self-rated health [53] • Body mass index (self-reported height and weight) • Depression symptoms measured by the Patient Health Questionnaire for Depression and Anxiety [54,55] • Alcohol measured by Alcohol Use Disorders Identification Test-Concise [49,50] and drug use [56] • COVID-19 symptoms and testing history
HIV: behavioral risks and uptake of HIV prevention and care	• HIV acquisition risk behaviors adapted from World Health Organization biobehavioral survey guidelines for populations at risk for HIV [56] • Access to and engagement in HIV services: HIV testing, HIV prevention [56] • HIV care continuum: self-reported diagnosis of HIV, engagement in HIV care including CD4 testing, viral load testing, and suppression [45,46,57], including country(ies) where care and treatment were accessed • Access to, uptake, and adherence to HIV treatment adapted from the Adult AIDS Clinical Trials Group survey measures [58]
Social measures	• Discrimination using the Everyday Discrimination Scale (short version) [51] • Violence victimization using the Assessment Screen to Identify Survivors Toolkit for Gender-Based Violence screen for displaced populations [47,48]
Respondent-driven sampling	• Social network size questions used for respondent-driven sampling weighting procedures [26]

Biological Measures

Biological measures include rapid HIV and syphilis screening using Standard Diagnostics BIOLINE HIV/Syphilis Duo with finger-prick blood specimens. Standard Diagnostics BIOLINE HIV/Syphilis Duo has a reported sensitivity of 99.8% and specificity of 100% for anti-HIV antibody detection and a reported sensitivity of 90% and specificity of 99.9% for anti–*Treponema pallidum* antibody detection [59]. Screening results are available within 20 minutes and provided to the study participants during the study visit. Participants with a reactive result on either or both tests are asked to provide an additional venous specimen for laboratory-based confirmatory testing. Specimens are transported the same day to the local reference laboratory. Confirmatory testing for participants with a positive HIV screen follows national testing algorithms and policies [60] and is conducted using the MP Bio HIV BLOT 2.2, a qualitative enzyme immunoassay for antibody detection of HIV-1 and HIV-2. For clinical purposes, and in keeping with national guidelines [60], CD3, CD4, CD5, CD8, and viral load tests are also performed at the same time by the laboratory. HUMAN Diagnostics Syphilis Rapid Plasma Reagin Test was used for confirmatory testing and identification of active syphilis for participants with positive syphilis screen. All laboratory results are available to the study team within 2 weeks; negative results are communicated by phone to the participants within 1 business day of receipt. Participants with confirmed HIV, syphilis, or both are contacted within 1 business day to arrange a time to deliver the results in person. All laboratory results are provided to the participant. The medicolegal triage process begins immediately at that time. Laboratory results are shared with medical providers upon participant request to assist with treatment decisions that are made by the patient and provider.

Medicolegal Services

Lawyers employed by *Red Somos* support the legal process of registering Venezuelans with irregular migration status to support access to care through the national program. All participants identified to be living with HIV (previously or newly diagnosed) or who have active syphilis will undergo a legal triage in which their legal status in Colombia will be reviewed by the assigned lawyer. For those with irregular migration status, staff lawyers initiate and support the completion of necessary paperwork and processes to acquire *permiso especial de permanencia* (or permit of stay), a *salvoconducto* (a paper demonstrating the regularization process has been initiated), or the forthcoming *estatuto temporal de protección* (or temporary statute of protection) documentation. Access to medical care and treatment under the national health system generally requires possession of one of these forms of documentation or other documents in special circumstances.

The process to obtain a *salvoconducto* takes 15 days maximum, while the process to access ART through the national health system (for any individual, inclusive of Colombian citizens) takes up to 30 days. During the period, while participants wait for their permit of stay or *salvoconducto* request to be processed, *Red Somos* will link the participants to local clinics and agencies that currently provide interim HIV care and treatment to the Venezuelan population who they regularly serve. At this point, the providers and patients will make treatment decisions with

consideration of viral load, CD4 counts, and any critical comorbid conditions. Those agencies can initiate stopgap ART in as little as 8 days.

Case Finding

To support efforts to identify new or undiagnosed infections, we employed a hybrid RDS–case finding approach. Case finding follows World Health Organization and Centers for Disease Control and Prevention guidelines for partner notification services and were adapted to reflect community recommendations to mitigate risk of violations of privacy, breaches in confidentiality, and coercive medical practices [61-63]. Participants with laboratory-confirmed HIV are invited to participate in partner notifications services to identify and support HIV and syphilis testing of sexual or injecting partners. Although encouraged to invite contacts to get tested, participants have the option to decline or to request anonymous partner notification through study staff. Case finding is open to all adult partners (age ≥18 years) of participants with laboratory-confirmed HIV, regardless of country of origin or citizenship status. Children are not eligible, but clinical referrals are offered to parents of children who may have been infected perinatally. During posttest counseling and linkage to care, participants who opt-in to case finding activities are trained on how to disclose their HIV status and invite partners to participate in HIV and syphilis testing. Participants are provided with a uniquely identified coupon that contains study contact information, which enables anonymous linkage of participants and contacts. The case-finding coupon is distinct in appearance from the RDS coupon but functions in a similar way. Any persons found to be living with HIV or syphilis through the case finding are connected to care through the same pathways. Brief service-focused interview questionnaires are administered among case-finding contacts to identify transmission risk behaviors and access/use of HIV prevention and testing.

Sample Size

Assuming a 1% HIV prevalence among general population, based on reports from local providers that suggest a range of 0.5% prevalence among adults to 1.5% prevalence in antenatal care surveillance, alpha .05, 0.005 margin of error, and design effect of 2 that has been suggested for RDS [64-66], we estimated that a sample size of 3043 per territory (approximately 6100 overall) is needed to estimate population HIV prevalence. This sample size provides a sufficiently small sampling fraction required by most of the RDS estimators [25], given that the Venezuelan migrant populations are estimated to exceed 115,000 persons in both territories. Individuals who are enrolled in case finding activities will contribute to the total enrolled sample. Assuming an HIV prevalence as high as 1.5% and an average of 2 contacts (case finding referrals) per index participant, we estimate that an additional 190 individuals will be enrolled in case finding (95 per territory) for HIV. Assuming 3% prevalence of syphilis, an additional 360 case finding participants will be enrolled for syphilis.

Analytic Plan

Basic descriptive analysis will be performed to estimate the prevalence of key demographic and health characteristics of the sample population. Primary analysis will focus on estimation of HIV prevalence among the general population of Venezuelans residing in the 2 territories, with estimates separately for each territory. Among participants living with HIV infection (prevalent or new diagnoses), we will assess engagement in the HIV care continuum, including the proportion who report being aware of their infection, engaged in HIV care, currently on ART, completing viral load testing in the last 6 months, and having suppressed viral load [67]. Subgroup analyses will be conducted to estimate HIV prevalence by HIV risk behaviors, gender, and age. All descriptive analyses will include unweighted and RDS-weighted population estimates for the adult Venezuelan population [68]. RDS-weighted analysis will be performed using Stata (StataCorp LLC) [69] or RDS-A (RDS-Analyst) [38] software and selecting the estimator most appropriate based on the sample characteristics. Analyses will incorporate RDS survey weights based on self-reported network size to calculate population prevalence. Bootstrapping procedures will be performed to calculate associated 95% confidence intervals [69]. Descriptive analysis will be conducted among case-finding participants to assess the demographic and social characteristics of contacts. Except for case-finding participants who are recruited via RDS or otherwise eligible for RDS, unweighted estimates will be calculated given that they are not part of the original RDS network chains. Data will be combined with RDS participant data to calculate unweighted HIV and syphilis positivity estimates.

Ethical Considerations and Participant Protection

Study activities were reviewed and approved by the ethical review committee at the Universidad el Bosque in Bogotá, Colombia, and the Institutional Review Board at Johns Hopkins School of Public Health (28223). The protocol was also reviewed in accordance with Centers for Disease Control and Prevention human research protection procedures. Formative research with stakeholders was deemed not human subjects research and commenced prior to other study activities.

This study uses multiple strategies to address unique social risks that underlie research with migrant populations, which go beyond risks typically associated with HIV surveillance. Risks for migrant populations largely encompass concerns for social harms related to stigma and discrimination as well as barriers to access to services, particularly for those with irregular migration status. First, we use a vague study title, *BIENVENIR*, to avoid perceptions of increased risk of HIV among migrants, should others learn about the study. No information that would identify the study focus on HIV or among migrant populations is included in recruitment materials or other outward facing materials. Prior to the implementation of study activities, appropriate referral pathways for HIV, syphilis, ancillary health, and humanitarian services (eg, nutrition, housing, mental health, social support) were identified for participant referral. Our electronic survey is programmed to flag to staff instances when a participant self-reports symptoms of depression or anxiety, food insecurity, hazardous alcohol use, and violence victimization. Although all participants are offered locally tailored resource guides, these individuals will be provided with a more in-depth discussion about resources and services specific to their needs and living situation.

The onset of the COVID-19 pandemic occurred between the funding of this project and the initiation of the study activities. Study launch was delayed during the early peaks of the pandemic, and formative research was conducted through secure remote methods at that time. A separate and extensive COVID-19 biosecurity protocol was developed for in-person data collection, and it aligns with local policies. The biosecurity protocol was submitted to all ethical and protocol review committees, and it underwent additional review and approval by an independent Human Subjects Research Restart Committee at Johns Hopkins University before in-person research commenced.

Results

As of November 8, 2021, 3278 people have been screened and 3105 participants have been enrolled across sites, inclusive of 20 seeds, and we have reached a maximum recruitment depth of 12 waves thus far (Figures 2 and 3). The enrollment is expected to end by February/March 2022. The number of new participant screenings range from 20-27 per weekday in each territory; Saturday data collection sessions experience higher no-show rates with 9-20 screened per day. Table 2 displays the characteristics of the participants enrolled to date.

Figure 2. Respondent-driven sampling network graphs of participants in Bogotá and Soacha. The large red triangular nodes represent seeds, and the small blue circular nodes represent recruits.

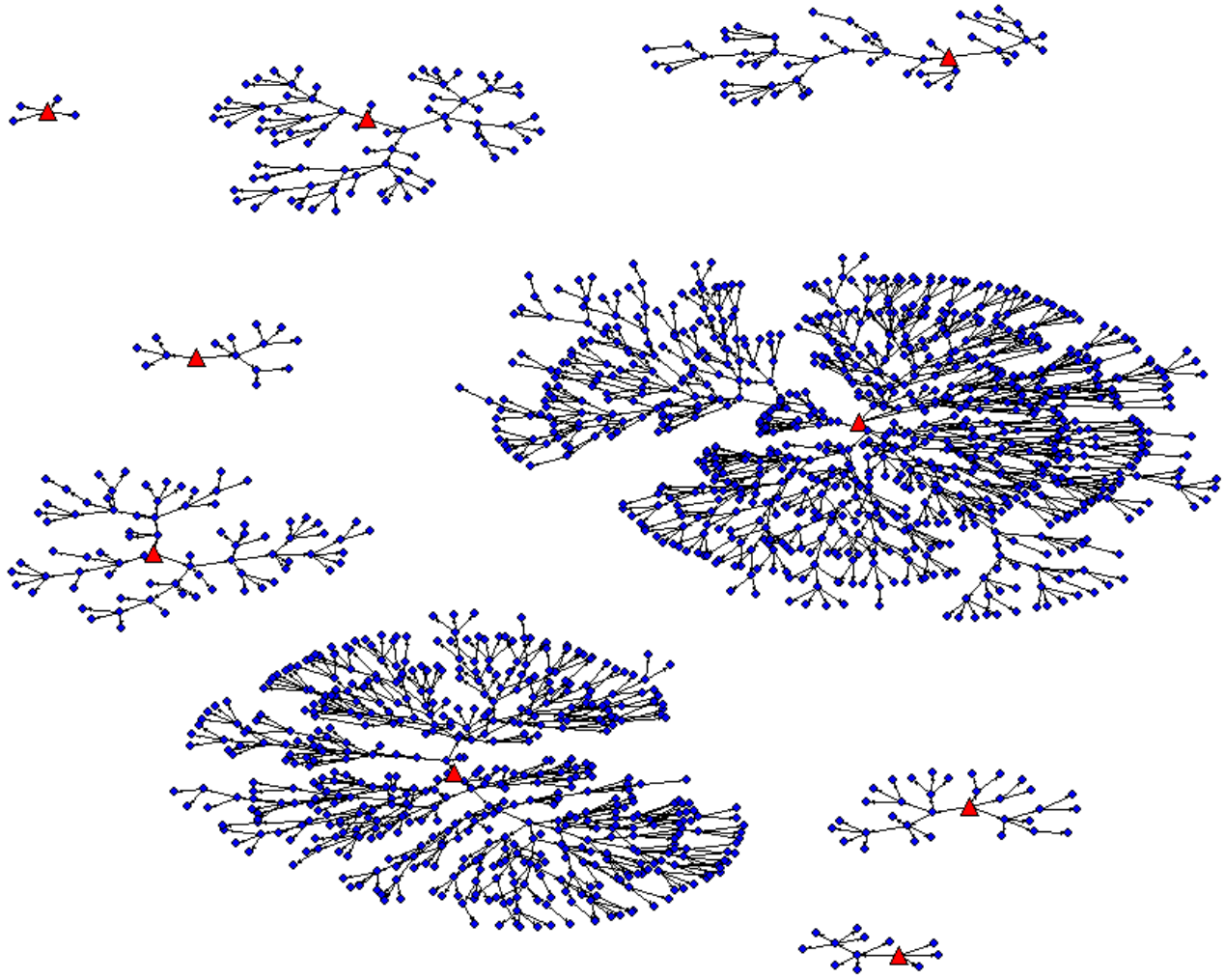


Figure 3. Respondent-driven sampling network graphs of participants in Barranquilla and Soledad. The large red triangular nodes represent seeds, and the small blue circular nodes represent recruits.

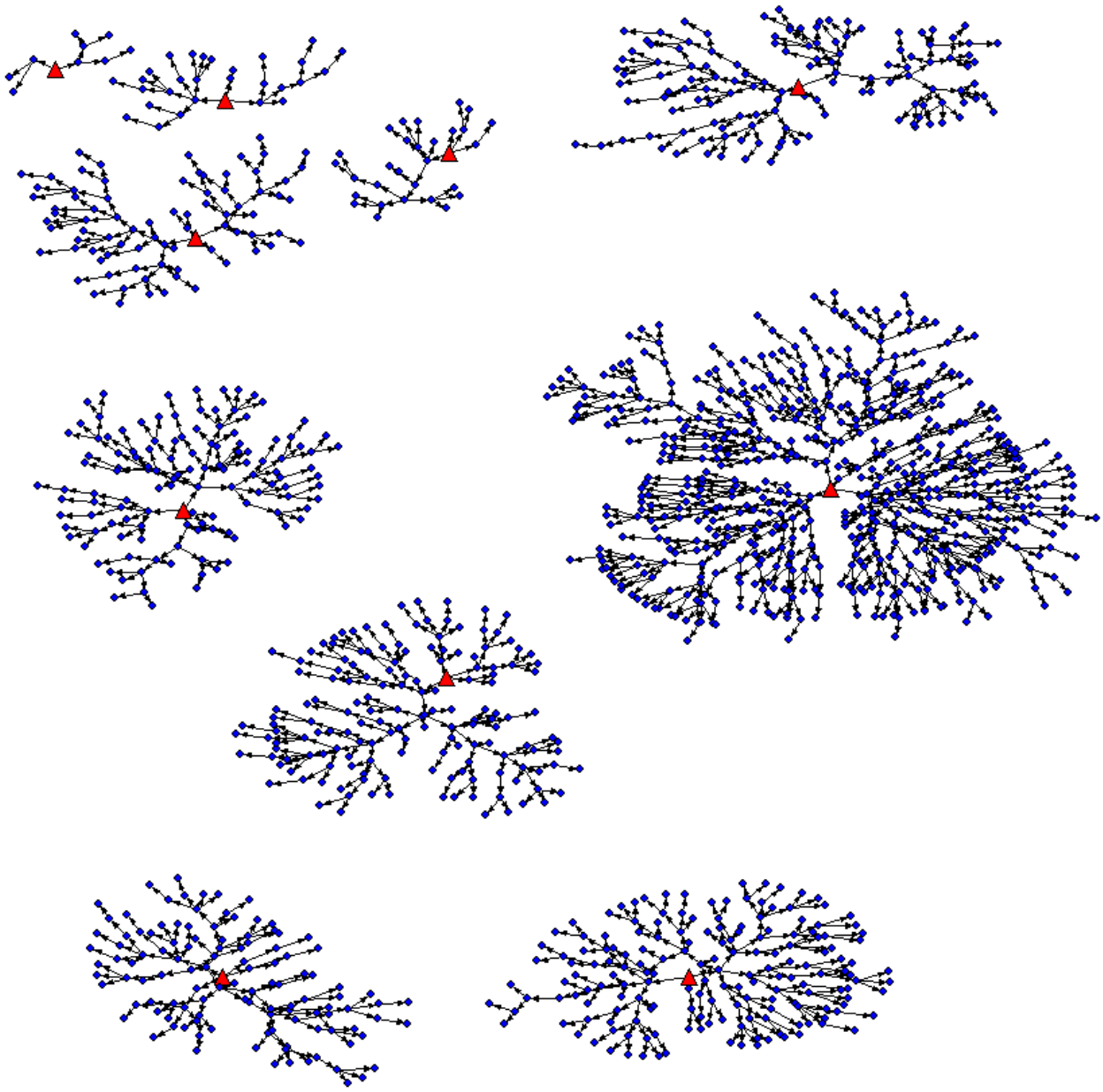


Table 2. Demographic and other characteristics of the study participants as of November 8, 2021.

Characteristics	Territory		
	Bogotá and Soacha (n=1684)	Barranquilla and Soledad (n=1421)	Total (N=3105)
Demographics			
Age (years), median (IQR)	32 (26-41)	33 (26-41)	32 (26-41)
City of residence, n (%)			
Bogotá	861 (51.1)	0 (0)	861 (27.7)
Soacha	822 (48.8)	0 (0)	824 (26.5)
Barranquilla	0 (0)	861 (60.6)	861 (27.7)
Soledad	0 (0)	560 (39.4)	558 (18)
Migration status, n (%)			
Regular	522 (31)	250 (17.6)	772 (24.9)
Irregular	1162 (69)	1171 (82.4)	2333 (75.1)
Gender identity, n (%)			
Male	562 (33.5)	351 (24.7)	913 (29.5)
Female	1091 (65.1)	1041 (73.3)	2132 (68.8)
Transgender or Nonbinary	24 (1.4)	29 (2)	53 (1.7)
High literacy (Rapid Estimate of Adult Literacy in Medicine–Short Form>6, reference<6), n (%)	1485 (89.4)	889 (63)	2374 (77.3)
HIV behavioral risks and testing history, n (%)			
Lifetime injecting drug use (reference: no)	37 (2.2)	22 (1.5)	59 (1.9)
Reports sex with a cis man (reference: no; denominator cis men and trans women, n=941)	53 (9.3)	27 (7.3)	80 (8.5)
Sex work (last 12 months)	32 (1.9)	30 (2.1)	62 (2)
Lifetime HIV test (reference: no)	1027 (61.2)	681 (48)	1708 (55.1)
Past diagnosis of HIV (reference: last test negative or unknown; n=3099)	7 (0.4)	9 (0.6)	16 (0.5)

Discussion

Early evaluation of enrollment and participant data show early signals that the methods described here are both feasible and acceptable for research in this context. The hybrid RDS–case finding approach is an innovation in RDS research, with the goal of increasing our ability to identify new or undiagnosed infections among partners and providing linkage to care. Modification of RDS to permit electronic referral of peers via SMS text messages and WhatsApp enables safe referral of peers while maintaining social distancing in the context of COVID-19. Use of text-message referrals builds on common communication pathways and appears to efficiently support peer referral. Given the tenuous access to HIV treatment of Venezuelans in Colombia with irregular migration status, the integration of medicolegal services in posttest counseling aims to increase access to HIV care, decrease time to ART initiation, as well as reduce untreated syphilis. The high recruitment rate thus far is a testament to the efficiencies of RDS and to the model of community-led research implementation and comprehensive service provision inclusive of HIV prevention and linkage to care, legal services, and other ancillary services. These findings will have direct relevance to Colombia, but methods and lessons learned from this study can be adapted for use across diverse settings with numerous health outcomes. With almost 272 million international migrants globally and over 82 million forcibly displaced persons, of whom 55 million are internally displaced due to conflict and insecurity as of 2020 [70,71], such methods are increasingly valuable for understanding and informing strategies related to migrant, humanitarian, and public health.

Acknowledgments

We acknowledge with gratitude the support and collaboration of Abu S Abdul-Quader, Dante Bugli, Kevin Clarke, Eva Leidman, Horacio Ruiseñor-Escudero, and Paul Young from US Centers for Disease Control and Prevention; Ricardo Luque Núñez from the Ministry of Health and Social Protection in Colombia; and Federico Duarte and Saskia Loochkartt from the United Nations High Commissioner for Refugees. This work was supported by the Centers for Disease Control and Prevention under the terms

of cooperative agreement NU2GGH002000-03-01. The Centers for Disease Control and Prevention investigators do not interact with human participants or have access to identifiable data or specimens for research purposes. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services.

Authors' Contributions

ALW, PS, and KP conceptualized the design of the study; JRG designed the medicolegal triage system and is the site principal investigator in Colombia; MS coordinates the overall study; JO, JJL, JFR, CQ, AV, YM, and FR supervised and coordinated data collection at each site; JC designed and engineered automated respondent-driven sampling processes and data management system; AJH and WH provide technical support; ALW and MS wrote the initial drafts of this manuscript; and all authors reviewed and contributed to this paper. All authors have approved this paper.

Conflicts of Interest

None declared.

References

1. Venezuela situation. UNHCR. URL: https://www.unhcr.org/en-us/venezuela-emergency.html#:~:text=Over%204%20million%20Venezuelans%20have,world's%20biggest%20recent%20displacement%20crises [accessed 2021-09-07]
2. Page KR, Doocy S, Reyna Ganteaume F, Castro JS, Spiegel P, Beyrer C. Venezuela's public health crisis: a regional emergency. Lancet 2019 Mar 23;393(10177):1254-1260. [doi: 10.1016/S0140-6736(19)30344-7] [Medline: 30871722]
3. Grillet ME, Hernández-Villena JV, Llewellyn MS, Paniz-Mondolfi AE, Tami A, Vincenti-Gonzalez MF, et al. Venezuela's humanitarian crisis, resurgence of vector-borne diseases, and implications for spillover in the region. Lancet Infect Dis 2019 May;19(5):e149-e161. [doi: 10.1016/S1473-3099(18)30757-6] [Medline: 30799251]
4. Rodríguez-Morales AJ, Bonilla-Aldana DK, Morales M, Suárez JA, Martínez-Buitrago E. Migration crisis in Venezuela and its impact on HIV in other countries: the case of Colombia. Ann Clin Microbiol Antimicrob 2019 Mar 08;18(1):9 [FREE Full text] [doi: 10.1186/s12941-019-0310-4] [Medline: 30849989]
5. Rodríguez-Morales AJ, Suárez JA, Risquez A, Delgado-Noguera L, Paniz-Mondolfi A. The current syndemic in Venezuela: Measles, malaria and more co-infections coupled with a breakdown of social and healthcare infrastructure. Quo vadis? Travel Med Infect Dis 2019;27:5-8. [doi: 10.1016/j.tmaid.2018.10.010] [Medline: 30326278]
6. Franco L, Pagan I, Serre Del Cor N, Schunk M, Neumayr A, Molero F, et al. Molecular epidemiology suggests Venezuela as the origin of the dengue outbreak in Madeira, Portugal in 2012-2013. Clin Microbiol Infect 2015 Jul;21(7):713.e5-713.e8 [FREE Full text] [doi: 10.1016/j.cmi.2015.03.016] [Medline: 25843502]
7. Paniz-Mondolfi AE, Tami A, Grillet ME, Márquez M, Hernández-Villena J, Escalona-Rodríguez MA, et al. Resurgence of Vaccine-Preventable Diseases in Venezuela as a Regional Public Health Threat in the Americas. Emerg Infect Dis 2019 Apr;25(4):625-632 [FREE Full text] [doi: 10.3201/eid2504.181305] [Medline: 30698523]
8. Epidemiological update: yellow fever. Pan American Health Organization. URL: https://www.paho.org/en/documents/epidemiological-update-yellow-fever-6-october-202110 [accessed 2021-10-06]
9. Paniz-Mondolfi AE, Sordillo EM, Márquez-Colmenarez MC, Delgado-Noguera LA, Rodriguez-Morales AJ. The arrival of SARS-CoV-2 in Venezuela. Lancet 2020 May 16;395(10236):e85-e86 [FREE Full text] [doi: 10.1016/S0140-6736(20)31053-9] [Medline: 32380043]
10. Zambrano-Barragán P, Ramírez Hernández S, Freier LF, Luzes M, Sobczyk R, Rodríguez A, et al. The impact of COVID-19 on Venezuelan migrants' access to health: A qualitative study in Colombian and Peruvian cities. J Migr Health 2021;3:100029 [FREE Full text] [doi: 10.1016/j.jmh.2020.100029] [Medline: 34405183]
11. PAHO, WHO, UNAIDS, Gobierno BDV. Plan maestro para el fortalecimiento de la respuesta al VH, la tuberculosis y la malaria en la República Bolivariana de Venezuela desde una perspectiva de salud pública. ONUSIDA. URL: https://tinyurl.com/48hr3f44 [accessed 2018-07-01]
12. Proposal for increasing health care coverage for Venezuelan refugees migrants living with HIV: Update in response to the COVID-19 pandemic. Pan American Health Organization. URL: https://iris.paho.org/handle/10665.2/54911 [accessed 2021-09-04]
13. Country factsheet: Venezuela. UNAIDS. URL: https://www.unaids.org/en/regionscountries/countries/venezuela [accessed 2020-01-01]
14. Antiretroviral Therapy Cohort Collaboration (ART-CC), Shepherd BE, Jenkins CA, Parrish DD, Glass TR, Cescon A, et al. Higher rates of AIDS during the first year of antiretroviral therapy among migrants: the importance of tuberculosis. AIDS 2013 May 15;27(8):1321-1329 [FREE Full text] [doi: 10.1097/QAD.0b013e32835faa95] [Medline: 23925379]
15. Displaced Colombians open their doors to Venezuelan refugees and migrants. UNHCR. URL: https://www.unhcr.org/en-us/news/stories/2018/12/5c12bbae4/displaced-colombians-open-doors-venezuelan-refugees-migrants.html [accessed 2022-06-14]

16. 25,000 Venezuelans cross the Colombian border every day. Relief Web International. URL: https://reliefweb.int/report/colombia/25000-venezuelans-cross-colombian-border-every-day [accessed 2020-01-01]
17. Operational data portal: Venezuela situation. UNHCR. URL: https://data2.unhcr.org/en/situations/vensit [accessed 2019-06-12]
18. UNHCR and IOM welcome Colombia's decision to regularize Venezuelan refugees and migrants. UNHCR. URL: https://www.unhcr.org/en-us/news/press/2021/2/60214cf74/unhcr-iom-welcome-colombias-decision-regularize-venezuelan-refugees-migrants.html [accessed 2021-03-21]
19. Distribución de Venezolanos en Colombia - Migración Colombia. URL: https://www.migracioncolombia.gov.co/infografias/distribucion-de-venezolanos-en-colombia-corte-31-de-enero-de-2021 [accessed 2022-02-16]
20. US to pay for thousands of doses of HIV drugs for Venezuelan migrants. Reuters. 2019 Aug 28. URL: https://www.reuters.com/article/us-colombia-usa-health-idUSKCN1VI2LQ [accessed 2022-02-16]
21. AIDS Healthcare Foundation Colombia. URL: https://www.aidshealth.org/global/colombia/ [accessed 2021-03-01]
22. Deren S, Shedlin M, Decena CU, Mino M. Research challenges to the study of HIV/AIDS among migrant and immigrant Hispanic populations in the United States. J Urban Health 2005 Jun;82(2 Suppl 3):iii13-iii25 [FREE Full text] [doi: 10.1093/jurban/jti060] [Medline: 15933327]
23. Goldenberg SM, Strathdee SA, Perez-Rosales MD, Sued O. Mobility and HIV in Central America and Mexico: a critical review. J Immigr Minor Health 2012 Feb;14(1):48-64. [doi: 10.1007/s10903-011-9505-2] [Medline: 21789558]
24. Distribución de Venezolanos en Colombia - corte 31 de enero de 2021. Ministerio DRE. URL: https://www.migracioncolombia.gov.co/infografias/distribucion-de-venezolanos-en-colombia-corte-31-de-enero-de-2021 [accessed 2021-03-01]
25. Gile KJ, Handcock MS. Respondent-Driven Sampling: An Assessment of Current Methodology. Sociol Methodol 2010 Aug;40(1):285-327 [FREE Full text] [doi: 10.1111/j.1467-9531.2010.01223.x] [Medline: 22969167]
26. Heckathorn DD. Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. Social Problems 1997 May;44(2):174-199. [doi: 10.2307/3096941]
27. Wirtz AL, Mehta SH, Latkin C, Zelaya CE, Galai N, Peryshkina A, et al. Comparison of Respondent Driven Sampling Estimators to Determine HIV Prevalence and Population Characteristics among Men Who Have Sex with Men in Moscow, Russia. PLoS One 2016;11(6):e0155519 [FREE Full text] [doi: 10.1371/journal.pone.0155519] [Medline: 27248818]
28. Solomon SS, McFall AM, Lucas GM, Srikrishnan AK, Kumar MS, Anand S, et al. Respondent-driven sampling for identification of HIV- and HCV-infected people who inject drugs and men who have sex with men in India: A cross-sectional, community-based analysis. PLoS Med 2017 Nov;14(11):e1002460 [FREE Full text] [doi: 10.1371/journal.pmed.1002460] [Medline: 29182638]
29. Wangroongsarb P, Hwang J, Thwing J, Karuchit S, Kumpetch S, Rand A, et al. Using Respondent Driven Sampling to Identify Malaria Risks and Occupational Networks among Migrant Workers in Ranong, Thailand. PLoS One 2016;11(12):e0168371 [FREE Full text] [doi: 10.1371/journal.pone.0168371] [Medline: 28033322]
30. Johnston L, Oumzil H, El Rhilani H, Latifi A, Bennani A, Alami K. Sex Differences in HIV Prevalence, Behavioral Risks and Prevention Needs Among Anglophone and Francophone Sub-Saharan African Migrants Living in Rabat, Morocco. AIDS Behav 2016 Apr;20(4):746-753. [doi: 10.1007/s10461-015-1115-x] [Medline: 26122648]
31. Roberton T, Weiss W, Jordan Health Access Study Team, Lebanon Health Access Study Team, Doocy S. Challenges in Estimating Vaccine Coverage in Refugee and Displaced Populations: Results From Household Surveys in Jordan and Lebanon. Vaccines (Basel) 2017 Aug 12;5(3):22 [FREE Full text] [doi: 10.3390/vaccines5030022] [Medline: 28805672]
32. Montealegre JR, Risser JM, Selwyn BJ, McCurdy SA, Sabin K. Prevalence of HIV risk behaviors among undocumented Central American immigrant women in Houston, Texas. AIDS Behav 2012 Aug;16(6):1641-1648. [doi: 10.1007/s10461-011-0130-9] [Medline: 22249955]
33. Weinmann T, AlZahmi A, Schneck A, Mancera Charry JF, Fröschl G, Radon K. Population-based assessment of health, healthcare utilisation, and specific needs of Syrian migrants in Germany: what is the best sampling method? BMC Med Res Methodol 2019 Jan 07;19(1):5 [FREE Full text] [doi: 10.1186/s12874-018-0652-1] [Medline: 30616507]
34. Berbesi-Fernández DY, Segura-Cardona AM, Montoya-Velez LP, Santamaría AÁ. Use of injected heroin and risk of hepatitis C in three cities in Colombia. Rev Bras Epidemiol 2017;20(3):435-444 [FREE Full text] [doi: 10.1590/1980-5497201700030007] [Medline: 29160436]
35. Zea MC, Reisen CA, del Río-González AM, Bianchi FT, Ramirez-Valles J, Poppen PJ. HIV Prevalence and Awareness of Positive Serostatus Among Men Who Have Sex With Men and Transgender Women in Bogotá, Colombia. Am J Public Health 2015 Aug;105(8):1588-1595. [doi: 10.2105/AJPH.2014.302307] [Medline: 25602899]
36. Rubio Mendoza ML, Jacobson JO, Morales-Miranda S, Sierra Alarcón CÁ, Luque Núñez R. High HIV Burden in Men Who Have Sex with Men across Colombia's Largest Cities: Findings from an Integrated Biological and Behavioral Surveillance Study. PLoS One 2015;10(8):e0131040 [FREE Full text] [doi: 10.1371/journal.pone.0131040] [Medline: 26252496]
37. Only half of refugees and migrants from Venezuela feel informed, survey finds. Relief Web International. URL: https://reliefweb.int/report/ecuador/only-half-refugees-and-migrants-venezuela-feel-informed-survey-finds [accessed 2021-03-01]

38. Handcock M, Fellows I, Gile K. DeducerRDSAnalyst: graphical user interface to the RDS package for respondent-driven sampling. GitHub. URL: https://github.com/ifellows/DeducerRDSAnalyst [accessed 2022-02-16]
39. Arozullah AM, Yarnold PR, Bennett CL, Soltysik RC, Wolf MS, Ferreira RM, et al. Development and validation of a short-form, rapid estimate of adult literacy in medicine. Med Care 2007 Nov;45(11):1026-1033. [doi: 10.1097/MLR.0b013e3180616c1b] [Medline: 18049342]
40. Beimer P, Lyberg L. Data Collection Modes and Associated Errors. Introduction to Survey Quality 2003 Jan 01:188-214. [doi: 10.1002/0471458740.ch6]
41. Dillman D, Smyth J, Christian L. Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 3rd ed. USA: John Wiley & Sons Inc; Jan 01, 2009.
42. Wirtz AL, Naing S, Clouse E, Thu KH, Mon SHH, Tun ZM, et al. The Parasol Protocol: An Implementation Science Study of HIV Continuum of Care Interventions for Gay Men and Transgender Women in Burma/Myanmar. JMIR Res Protoc 2017 May 17;6(5):e90 [FREE Full text] [doi: 10.2196/resprot.7642] [Medline: 28526661]
43. 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 03;8(10):e14704 [FREE Full text] [doi: 10.2196/14704] [Medline: 31584005]
44. Ballard T, Coates J, Swindale A, Deitchler M. Household hunger scale: indicator definition and measurement guide. Food and Nutrition Technical Assistance. URL: https://www.fantaproject.org/sites/default/files/resources/HHS-Indicator-Guide-Aug2011.pdf [accessed 2019-09-01]
45. Kozak M, Zinski A, Leeper C, Willig JH, Mugavero MJ. Late diagnosis, delayed presentation and late presentation in HIV: proposed definitions, methodological considerations and health implications. Antivir Ther 2013;18(1):17-23. [doi: 10.3851/IMP2534] [Medline: 23341432]
46. Mugavero MJ, Amico KR, Horn T, Thompson MA. The state of engagement in HIV care in the United States: from cascade to continuum to control. Clin Infect Dis 2013 Oct;57(8):1164-1171. [doi: 10.1093/cid/cit420] [Medline: 23797289]
47. Vu A, Wirtz A, Pham K, Singh S, Rubenstein L, Glass N, et al. Psychometric properties and reliability of the Assessment Screen to Identify Survivors Toolkit for Gender Based Violence (ASIST-GBV): results from humanitarian settings in Ethiopia and Colombia. Confl Health 2016;10:1 [FREE Full text] [doi: 10.1186/s13031-016-0068-7] [Medline: 26865857]
48. Wirtz AL, Glass N, Pham K, Perrin N, Rubenstein LS, Singh S, et al. Comprehensive development and testing of the ASIST-GBV, a screening tool for responding to gender-based violence among women in humanitarian settings. Confl Health 2016;10:7 [FREE Full text] [doi: 10.1186/s13031-016-0071-z] [Medline: 27099617]
49. de la Espriella Guerrero RA, Rodriguez V, Rincón CJ, Morales DC, Rodríguez SJP, Gómez-Restrepo C. [Alcohol Consumption in the Colombian Population, 2015 National Mental Health Survey]. Rev Colomb Psiquiatr 2016 Dec;45 Suppl 1:76-88. [doi: 10.1016/j.rcp.2016.05.002] [Medline: 27993259]
50. 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]
51. Williams DR, Yu Y, Yan Yu, Jackson JS, Anderson NB. Racial Differences in Physical and Mental Health: Socio-economic Status, Stress and Discrimination. J Health Psychol 1997 Jul;2(3):335-351. [doi: 10.1177/135910539700200305] [Medline: 22013026]
52. Colombia: standard DHS. Demographic Health Surveys. URL: https://dhsprogram.com/methodology/survey/survey-display-476.cfm [accessed 2021-09-22]
53. Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I Epidemiologic Follow-up Study. Am J Public Health 1990 Apr;80(4):446-452. [doi: 10.2105/ajph.80.4.446] [Medline: 2316767]
54. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
55. Kocalevent R, Finck C, Jimenez-Leal W, Sautier L, Hinz A. Standardization of the Colombian version of the PHQ-4 in the general population. BMC Psychiatry 2014 Jul 19;14:205 [FREE Full text] [doi: 10.1186/1471-244X-14-205] [Medline: 25037706]
56. Biobehavioral survey guidelines for populations at risk for HIV. WHO. 2017 Jan 01. URL: https://apps.who.int/iris/bitstream/handle/10665/258924/9789241513012-eng.pdf [accessed 2022-02-16]
57. Monitoring, evaluation, and reporting indicator reference guide. PEPFAR. URL: https://www.state.gov/wp-content/uploads/2019/10/PEPFAR-MER-Indicator-Reference-Guide-Version-2.4-FY20.pdf [accessed 2019-01-01]
58. Chesney MA, Ickovics JR, Chambers DB, Gifford AL, Neidig J, Zwickl B, et al. AIDS Care 2000 Jun;12(3):255-266. [doi: 10.1080/09540120050042891] [Medline: 10928201]
59. HIV syphilis duo package insert. WHO. URL: https://www.who.int/diagnostics_laboratory/evaluations/151028_final_report_0179-012-00_sd_bioline_hiv_syphilis2.pdf [accessed 2022-02-16]

60. Guía de práctica clínica (GPC) basada en la evidencia científica para la atención de la infección por VIH/Sida en adolescentes (con 13 años de edad o más) y adultos. Minsalud. URL: https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/CA/gpc-completa-evidencia-cientifica-vih-sida-adolescentes-adultos.pdf [accessed 2019-01-01]
61. Guidelines on HIV Self-testing and Partner Notification: Supplement to Consolidated Guidelines on HIV Testing Services. Geneva: World Health Organization; Jan 01, 2016.
62. Recommendations for partner services programs for HIV infection, syphilis, gonorrhea, and chlamydial infection. Centers for Disease Control and Prevention. 2012. URL: https://www.cdc.gov/nchhstp/partners/docs/08_124108_Stuckey_QuickGuideInsides_121508_Update_WithCover-508C.pdf [accessed 2019-09-01]
63. Ayala G, Bahati M, Balan E, Chang J, Do TD, Fokeerbux NA, et al. Partner Notification: A Community Viewpoint. J Int AIDS Soc 2019 Jul;22 Suppl 3:e25291 [FREE Full text] [doi: 10.1002/jia2.25291] [Medline: 31321916]
64. Salganik MJ. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. J Urban Health 2006 Nov;83(6 Suppl):i98-112 [FREE Full text] [doi: 10.1007/s11524-006-9106-x] [Medline: 16937083]
65. Wejnert C, Pham H, Krishna N, Le B, DiNenno E. Estimating design effect and calculating sample size for respondent-driven sampling studies of injection drug users in the United States. AIDS Behav 2012 May;16(4):797-806 [FREE Full text] [doi: 10.1007/s10461-012-0147-8] [Medline: 22350828]
66. Johnston LG, Hakim AJ, Dittrich S, Burnett J, Kim E, White RG. A Systematic Review of Published Respondent-Driven Sampling Surveys Collecting Behavioral and Biologic Data. AIDS Behav 2016 Aug;20(8):1754-1776 [FREE Full text] [doi: 10.1007/s10461-016-1346-5] [Medline: 26992395]
67. Piñeirúa A, Sierra-Madero J, Cahn P, Guevara Palmero RN, Martínez Buitrago E, Young B, et al. The HIV care continuum in Latin America: challenges and opportunities. Lancet Infect Dis 2015 Jul;15(7):833-839. [doi: 10.1016/S1473-3099(15)00108-5] [Medline: 26122456]
68. Volz E, Heckathorn D. Probability-based estimation theory for respondent driven sampling. Journal of Official Statistics. 2008 Jan 01. URL: http://www.sverigeisiffror.scb.se/contentassets/ff271eeeca694f47ae99b942de61df83/probability-based-estimation-theory-for-respondent-driven-sampling.pdf [accessed 2022-02-16]
69. Schonlau M, Liebau E. Respondent-Driven Sampling. The Stata Journal 2012 Mar 01;12(1):72-93. [doi: 10.1177/1536867x1201200106]
70. International Organization for Migration. World Migration Report 2020. Geneva: UN Migration; Jan 01, 2020.
71. 2020 internal displacement. International Displacement Monitoring Center. URL: https://www.internal-displacement.org/database/displacement-data [accessed 2021-09-17]

Abbreviations

ART: antiretroviral therapy
BIENVENIR: Bienestar de Venezolanos quienes son Inmigrantes y Refugiados
PAHO: Pan American Health Organization
PLHIV: people living with HIV
RDS: respondent-driven sampling

Edited by T Derrick; submitted 29.12.21; this is a non–peer-reviewed article;accepted 05.01.22; published 08.03.22.

Please cite as:
Wirtz AL, Page KR, Stevenson M, Guillén JR, Ortíz J, López JJ, Ramírez JF, Quijano C, Vela A, Moreno Y, Rigual F, Case J, Hakim AJ, Hladik W, Spiegel PB
HIV Surveillance and Research for Migrant Populations: Protocol Integrating Respondent-Driven Sampling, Case Finding, and Medicolegal Services for Venezuelans Living in Colombia
JMIR Res Protoc 2022;11(3):e36026
URL: https://www.researchprotocols.org/2022/3/e36026
doi:10.2196/36026
PMID:35258458

Protocol

Linking Electronic Health Records and In-Depth Interviews to Inform Efforts to Integrate Social Determinants of Health Into Health Care Delivery: Protocol for a Qualitative Research Study

Annemarie Hirsch[1], MPH, PhD; T Elizabeth Durden[2], PhD; Jennifer Silva[3], PhD

[1]Department of Population Health Sciences, Geisinger, Danville, PA, United States

[2]Department of Sociology and Anthropology, Bucknell University, Lewisburg, PA, United States

[3]Paul H. O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington, IN, United States

Corresponding Author:
Annemarie Hirsch, MPH, PhD
Department of Population Health Sciences
Geisinger
100 N. Academy Avenue
Danville, PA, 17822
United States
Phone: 1 267 626 8110
Email: aghirsch@geisinger.edu

Abstract

Background: Health systems are attempting to capture social determinants of health (SDoH) in electronic health records (EHR) and use these data to adjust care plans. To date, however, methods for identifying social needs, which are the SDoH prioritized by patients, have been underexplored, and there is little guidance as to how clinicians should act on SDoH data when caring for patients. Moreover, the unintended consequences of collecting and responding to SDoH are poorly understood.

Objective: The objective of this study is to use two data sources, EHR data and patient interviews, to describe divergences between the EHR and patient experiences that could help identify gaps in the documentation of SDoH in the EHR; highlight potential missed opportunities for addressing social needs, and identify unintended consequences of efforts to integrate SDoH into clinical care.

Methods: We are conducting a qualitative study that merges discrete and free-text data from EHRs with in-depth interviews with women residing in rural, socioeconomically deprived communities in the Mid-Atlantic region of the United States. Participants had to confirm that they had at least one visit with the large health system that serves the region. Interviews with the women included questions regarding health, interaction with the health system, and social needs. Next, with consent, we extracted discrete data (eg, diagnoses and medication orders) for each participant and free-text clinician notes from this health system's EHRs between 1996 and the year of the interview. We used a standardized protocol to create an EHR narrative, a free-text summary of the EHR data. We used NVivo to identify themes in the interviews and the EHR narratives.

Results: To date, we have interviewed 88 women, including 51 White women, 19 Black women, 14 Latina women, 2 mixed Black and Latina women, and 2 Asian Pacific women. We have completed the EHR narratives on 66 women. The women range in age from 18 to 90 years. We found corresponding EHR data on all but 4 of the interview participants. Participants had contact with a wide range of clinical departments (eg, psychiatry, neurology, and infectious disease) and received care in various clinical settings (eg, primary care clinics, emergency departments, and inpatient hospitalizations). A preliminary review of the EHR narratives revealed that the clinician notes were a source of data on a range of SDoH but did not always reflect the social needs that participants described in the interviews.

Conclusions: This study will provide unique insight into the demands and consequences of integrating SDoH into clinical care. This work comes at a pivotal point in time, as health systems, payors, and policymakers accelerate attempts to deliver care within the context of social needs.

International Registered Report Identifier (IRRID): DERR1-10.2196/36201

(JMIR Res Protoc 2022;11(3):e36201) doi:10.2196/36201

KEYWORDS

electronic health records; social determinants of health; poverty; rural; qualitative; health system

Introduction

Health systems are under increasing pressure to assess and act on social determinants of health (SDoH), with widespread support from professional organizations including the American Academy of Pediatrics, the American Academy of Family Physicians, and the American Medical Association. In 2014, the Institute of Medicine (currently the National Academy of Medicine) recommended the documentation of core measures of social and behavioral domains in electronic health records (EHRs), including education, financial resource strain, and stress [1]. More recently, the National Academies of Sciences, Engineering, and Medicine (NASEM) presented guidance for how clinicians and health systems should use SDoH data when managing patients [2]. These calls to action have prompted health systems to accelerate efforts to integrate SDoH into care delivery. It is imperative that these efforts are informed by the voice of key stakeholders, including clinicians documenting in the EHR and the patients receiving care.

Data on SDoH are entered into EHR systems through multiple pathways, but data are not routinely or systematically collected [3]. Medical vocabularies, including International Classification of Disease (ICD) codes and Logical Observation Identifiers, Names and Codes capture SDoH (eg, ICD-10 code Z55.0: problems related to education and literacy), but the codes are not comprehensive, and they are underused [4,5]. Medicare reported that the ICD codes used to capture SDoH were used for only 1.6% of Medicare beneficiaries in 2019 [6]. EHR vendors are working with health systems to develop patient-reported tools to improve the capture of SDoH in structured fields, but there is little standardization regarding what data to collect [7]. Clinician notes, while more challenging to access and use than discrete EHR data, have recently emerged as a potentially important source of SDoH data [7,8], but this data source remains relatively underexplored.

Potentially more challenging than capturing SDoH is the ability to identify the social needs and chronic stresses that patients prioritize. As described by Alderwick and Gottlieb [9], while screening tools can reveal multiple SDoH, such as food and housing instability, a patient may perceive that her most pressing social need is to escape from a violent partner. Thus, it is not sufficient for health systems to only collect SDoH data; they must also adequately capture the patient's perspective on these data [10].

In addition to identifying social needs, the NASEM report called for health systems to respond to social needs through adjustment (altering clinical care to accommodate identified social barriers), assistance (providing support in connecting patients with relevant social care resources), alignment (activities to work with existing social care assets in the community), and advocacy [2]. Health systems have already started implementing these activities, but research regarding the impact of these activities on health outcomes is limited. Importantly, the unintended consequences of these activities, including the potential for perpetuating bias and further marginalizing vulnerable populations, are poorly understood [2,11,12]. EHR data, particularly clinician notes, maybe a rich source of secondary data that provides some important insight into how clinicians are currently responding to social needs and the consequences of their responses.

This paper describes the protocol for a qualitative research study that merges discrete and free-text data from EHRs with in-depth interviews from among women living in rural and socioeconomically deprived communities in the Mid-Atlantic region of the United States. The objective of this study is to describe discrepancies between the medical record and patient narratives that could help identify gaps in documenting SDoH in the EHR, highlight potential missed opportunities for addressing social needs, and observe unintended negative consequences of documenting or responding to SDoH. This study leverages multiple perspectives to provide unique insight into the demands and consequences of integrating SDoH into clinical care.

Methods

Overview

This study is being conducted by an interdisciplinary team across three institutions, a health system (here forward referred to as Central Health), Bucknell University, and Indiana University. It is an ongoing qualitative study that combines EHR data from the health system with in-depth interviews of patients recruited by the university research teams. This study was approved by the institutional review boards (IRB) of all participating institutions (Bucknell University IRB approval number: 1920-123; Geisinger IRB approval number: 2017-0440).

Study Participants and Recruitment

To be eligible for this study, individuals had to have received care, per self-report, from Central Health and therefore have a Central Health EHR. Participants also had to be self-identified women over the age of 18 years with less than a four-year college degree. We used stratified sampling to recruit both White women and women of color, including women who self-identified as Black, Latina, and Asian Pacific. We recruited women residing in counties in the Mid-Atlantic region of the United States that are ranked among the lowest in health indicators, including poor physical and mental health days, as well as teen births, preventable hospital stays, and incidences of violent crimes and injury deaths [13]. We first recruited respondents through convenience sampling, visiting venues such as laundromats, bars and gaming rooms, public libraries, grocery stores, bus stops, dollar stores, public community events, churches, and job training workshops. We then employed snowball sampling, asking for each participant to recommend one to two people who might also participate in the study.

In-Depth Interviews

We conducted one-on-one interviews in-person until we needed to shift to remote interviewing, via video conferencing, due to

the COVID-19 pandemic. The interviews probed women about their past and current medical issues; their interaction with Central Health; their mental health; their understanding of required treatment and barriers to compliance; their experiences of stress, violence, discrimination, and economic deprivation; their families, children, and social support systems; and their sense of self-efficacy, trust, and the future. The interview allowed for participants to share their perspectives in their own words, rather than through a standard survey format. The authors conducted about two-thirds of the interviews, with trained student research assistants completing the remaining ones. We compensated interviewees for their time with $50 in cash at the completion of the interview. Interviews were recorded with the permission of the participant and fully transcribed by a professional firm.

Electronic Health Record Data

Central Health is a single, large, integrated health system serving a mix of rural and urban communities, including communities designated as Medically Underserved Areas [14]. The health system facilities include multiple inpatient hospitals and more than 125 primary and specialty clinic sites. The system uses Epic EHR software modules, including ambulatory, inpatient, surgery, emergency department, e-prescribing, computerized physician order entry, registration, and scheduling.

We extracted a set of discrete and free-text clinician notes on all consenting interview participants, matching participants to their record based on first name, last name, and date of birth (Textbox 1). We included participant data for any contact with the health system between 1996 and the year of the completed interview (between 2017-2021). Using these data, the health system research team created an EHR narrative, a free-text summary of the health of the participants and their interactions with the health system. We developed these narratives using a three-phase approach implemented by trained research assistants. First, the research assistant wrote a summary of the participant's health and contact with the health system in consecutive order, going from the first health system contact to the last, using only information from the discrete data fields. Next, the team members enriched this summary of discrete data with direct quotes from clinician notes. Research assistants included notes if the text addressed an SDoH (eg, occupation, education, marital status, history of trauma, transportation challenges, etc) or if the notes clarified discrete data (eg, a patient who did not fill prescriptions because of lack of insurance coverage). Notes were enclosed in quotation marks or highlighted to distinguish them from information obtained from discrete EHR data. The team was trained to be over-inclusive of potential SDoH, including any reference to life circumstances beyond health conditions and health care.

Finally, the research assistant revised the summary so that, rather than being purely consecutive, chronic conditions (eg, chronic pain and mental health) that were managed over time were described in their own separate sections, enabling readers to obtain the breadth, depth, and longevity of these chronic issues. In a pilot study of this methodology, the principal investigator and each research assistant completed narratives for the same five patients and refined training and instructions to ensure a consistent approach to the narratives. The principal investigator at the health system provided a second review of all final narratives to ensure completeness and to confirm that no personal identifiers were included.

Textbox 1. Electronic health record data extracted for all interview participants.

Sociodemographic factors: date of birth, race, ethnicity, health insurance
Health behaviors: smoking history, alcohol history, illicit drug history, sexual history
Diagnoses: diagnoses associated with clinical encounters, dates of diagnoses
Medication: medication orders (name, pharmacy class, dosage, indications), dates of medication orders
Utilization: dates of inpatient, outpatient, emergency department encounters, and telemedicine
Free-text notes: telephone encounters, health system letters, encounter notes, after-visit summary

Analysis

Central Health and the universities developed a data-sharing protocol to facilitate the merging of the interview and EHR data with minimal risk to participant confidentiality. Under the authors' guidance, two trained members of the research team coded each interview and EHR narrative using the qualitative software NVivo. We employed an approach called "flexible coding," which consisted of reading and rereading the interview transcripts line by line and slowly building connections across the different sources of data [15]. We entered the coding process with a list of codes, or "nodes," that we anticipated based on our review of the literature, but our list quickly expanded to include nodes that the participants, either directly or indirectly, emphasized as salient features and events in their social contexts, daily routines, and health care experiences, leading to approximately 200 nodes. We then used NVivo to systematically analyze relationships between nodes as well as frequencies of occurrences in the data. During this process, the second and third authors, with a team of trained students, created two to three-page narrative summaries, with direct quotes, of each participant's life history from their interview transcript, providing quick "snapshots" of cases for our reference.

At the same time, we did a close reading and mapping of observations in the EHR, notes, and interviews across three major categories: SDoH; stories of "health" and health care utilization/experiences; and interactions between the patient and the clinician. We first documented the similarities and differences between the EHR and patient narrative within each participant's case, and then compared across cases to refine our concepts, specifying the conditions and contexts that seemed to explain differences across cases, and repeating this process until we had identified a working set of social processes [16].

Results

To date, we have interviewed 88 women, including 51 White, 19 Black, 14 Latina, 2 mixed Black and Latina, and 2 Asian Pacific Islanders. We have extracted the EHR data on all participants with available EHR data and completed the EHR narratives on 66 women. The women ranged in age from 18 to 90 years. We found corresponding EHR data on all but 4 of the interview participants. Participants had contact with a wide range of clinical departments, including family practice, obstetrics and gynecology, specialty care (eg, psychiatry, otolaryngology, infection disease, neurology, endocrinology, cardiology, nutrition services, and ophthalmology), surgery (eg, vascular, general), and emergency medicine. Clinical visits occurred at the main campus of the health system, as well as community practice sites and smaller affiliated hospitals. Free-text notes were documented by different clinician types, including physicians, nurses, and social workers.

A preliminary review of the paired interviews and EHR narratives revealed that the clinician notes were a source of data on a range of SDoH (Textbox 2). Clinicians documented topics such as education and occupational history, transportation challenges, financial burdens, social isolation, issues with insurance, and childhood trauma. The notes also include data on how clinicians responded to social risk factors, including transportation issues and financial strain (Textbox 2). However, the notes did not always comprehensively capture the social needs of patients, as reflected in the misalignment of the social needs documented in the notes and the social needs described by patients during the interviews (Multimedia Appendix 1). Some interviewees reported that they withheld information from their clinicians out of fear or shame. Other interviewees attested to sharing information with their clinicians but felt that their clinicians did not take them seriously, resulting in perceptions of misdiagnosis and breakdowns in communications.

Textbox 2. Selected quotes from electronic health record notes regarding social determinants of health.

Financial strain:

- "Patient does not have inhalers at home. Does not particularly like and also cost is an issue."
- "Does not feel it (antidepressant) would help her as a lot of her mood changes are due to financial situation."
- "Has financial burdens due to having to take care of the grandchildren and feed them for years."

Insurance:

- "…she stopped getting them (mammograms) anymore because her insurance stopped paying."
- "Due to transportation difficulties and lack of insurance, she only attended two prenatal appointments."
- "During one of the visits this year, *the patient* noted that she had been doing well on Lipitor years ago for cholesterol but had to go off of it due to a change in her insurance. She requested to go back to Lipitor now that she had better health insurance."

Family trauma:

- "Sexually abused by uncle as child and 'gang raped' at age 14."
- "C/o ongoing depression… Worsening over the past year. Feels it started seven years ago after her child of 1 month and 20 days old died from SIDS. Also grew up in foster care and suffered physical and emotional abuse."
- "Conflict in the home between mom and dad has increased significantly, with *patient* witnessing most of it. On occasion, she has gotten between mom and dad. Mom concerned about impact on *patient*."
- "During this evaluation, it was revealed that *the patient's* first husband had verbally and physically abused her prior to their divorce in 2003. As of 2012, *the patient* is married and lives with her husband and 2 children. Her husband also suffers from chronic pain, and her husband's brother, who suffers from alcohol abuse, lives with them."

Clinician response to social risk factors:

- "Pt… planning to apply for food stamps. Enrolled in WIC. Provided contact info for local pregnancy support programs and childcare subsidy program. Pt declined referral to Nurse/Family Partnership program."
- "I talked with *patient* about transferring her to a more long-term, supportive therapist in her community. It has been difficult for her to attend regular appointments d/t to the distance."
- "Has necessary baby supplies, except care seat, discuss rental from L+D unit."

Discussion

Principal Findings

This is an innovative study that brings together two data sources, EHR data and patient voices, to help inform the integration of SDoH into health care delivery. Examining how these two data sources converge and diverge, we will identify successes as well as missed opportunities for capturing and addressing social needs and gain insight into the unintended consequences of these efforts. The goal of this study is to provide guidance to health systems as to how to collect and respond to SDoH.

The participants in the study are women with less than a college education, living in rural communities with high levels of

community socioeconomic deprivation. We designed the study so that close to half of the sample was Black, Latina, or both Black and Latina. We selected a study population with these characteristics as this population suffers from a lack of health equity and is thus most likely to benefit from better integration of SDoH into medical care. Individuals living in rural communities, for example, are at an increased risk of death from five leading causes (heart disease, cancer, unintentional injury, chronic lower respiratory disease, and stroke) compared to their urban counterparts [17]. Among a large list of disparities for those marginalized racially and economically, life expectancy is shorter among Black individuals compared to White as well as residents of low-income communities compared to high-income communities [18]; infant mortality rates are higher in Black mothers compared to White mothers and among mothers with a high school education compared to those with a college degree; pregnancy-related mortality ratios are higher among Black women compared to White women [19]. This study provides a voice for individuals most likely to experience social barriers to good health. By enrolling White, Black, Latina, and Asia Pacific women, we are also primed to evaluate differences in health system interactions by race and ethnicity.

This study has multiple strengths. First, the study integrates two perspectives through the analysis of EHR data and patient experiences. Second, the study uses more than 25 years of EHR data, enabling the study team to examine social risk factors across the lifespan in a subset of women. Finally, we have developed a novel approach to analyzing EHR data through the development of an EHR narrative based on discrete and free-text data for qualitative analysis. The use of EHR data from a single health system may be a limitation of this study. However, the health system employs more than 1000 physicians and more than 900 advanced practitioners; therefore, the EHR data used in this study is from a range of clinicians in a variety of clinical settings. Our inability to describe the sociodemographic characteristics of the clinicians is a further limitation, given the growing body of literature on the importance of physicians' characteristics to patients' health experiences [20,21].

Conclusions

As health systems accelerate attempts to integrate SDoH into the EHR and health care decisions, it is essential that these efforts are informed by the experiences of the patients receiving care, particularly the most vulnerable patients. This study will provide unique insight into the needs and consequences of integrating SDoH into clinical care. This work comes at a pivotal point in time, as health systems, payors, and policymakers accelerate attempts to deliver care within the context of social needs.

Acknowledgments

This research was funded by the Russell Sage Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample electronic health record and interview narratives.
[DOCX File , 20 KB - resprot_v11i3e36201_app1.docx]

Multimedia Appendix 2
Peer-review report by the Russell Sage Foundation.
[PDF File (Adobe PDF File), 64 KB - resprot_v11i3e36201_app2.pdf]

References

1. Capturing Social and Behavioral Domains and Measures in Electronic Health Records. Washington DC: National Academies Press; Jan 08, 2015.
2. National AOS, Engineering, Medicine 2. Integrating Social Care into the Delivery of Health Care: Moving Upstream to Improve the Nation's Health. Washington DC: National Academies Press; Jan 08, 2015.
3. Pantell M, Adler-Milstein J, Wang M, Prather A, Adler N, Gottlieb L. A call for social informatics. J Am Med Inform Assoc 2020 Nov 01;27(11):1798-1801 [FREE Full text] [doi: 10.1093/jamia/ocaa175] [Medline: 33202020]
4. Arons A, DeSilvey S, Fichtenberg C, Gottlieb L. Documenting social determinants of health-related clinical activities using standardized medical vocabularies. JAMIA Open 2019 Apr;2(1):81-88 [FREE Full text] [doi: 10.1093/jamiaopen/ooy051] [Medline: 31984347]
5. Truong HP, Luke AA, Hammond G, Wadhera RK, Reidhead M, Joynt Maddox KE. Utilization of Social Determinants of Health ICD-10 Z-Codes Among Hospitalized Patients in the United States, 2016-2017. Med Care 2020 Dec;58(12):1037-1043 [FREE Full text] [doi: 10.1097/MLR.0000000000001418] [Medline: 32925453]
6. American Hospital Association. Medicare releases data on Z code use to document social determinants of health. 2021 Oct 07. URL: https://www.aha.org/news/headline/2021-10-07-medicare-releases-data-z-code-use-document-social-determinants-health [accessed 2020-09-01]

7. Hatef E, Rouhizadeh M, Tia I, Lasser E, Hill-Briggs F, Marsteller J, et al. Assessing the Availability of Data on Social and Behavioral Determinants in Structured and Unstructured Electronic Health Records: A Retrospective Analysis of a Multilevel Health Care System. JMIR Med Inform 2019 Aug 02;7(3):e13802 [FREE Full text] [doi: 10.2196/13802] [Medline: 31376277]
8. Zhu V, Lenert L, Bunnell B, Obeid J, Jefferson M, Halbert C. Automatically identifying social isolation from clinical narratives for patients with prostate cancer. BMC Medical Informatics and Decision Making 2019:43 [FREE Full text] [doi: 10.1186/s12911-019-0795-y]
9. Alderwick H, Gottlieb LM. Meanings and Misunderstandings: A Social Determinants of Health Lexicon for Health Care Systems. Milbank Q 2019 Jun;97(2):407-419 [FREE Full text] [doi: 10.1111/1468-0009.12390] [Medline: 31069864]
10. Reverby SM. Listening to narratives from the Tuskegee syphilis study. Lancet 2011 May 14;377(9778):1646-1647. [doi: 10.1016/s0140-6736(11)60663-6] [Medline: 21591285]
11. Hollister B, Bonham VL. Should Electronic Health Record-Derived Social and Behavioral Data Be Used in Precision Medicine Research? AMA J Ethics 2018 Sep 01;20(9):E873-E880 [FREE Full text] [doi: 10.1001/amajethics.2018.873] [Medline: 30242819]
12. P Goddu A, O'Conor KJ, Lanzkron S, Saheed MO, Saha S, Peek ME, et al. Do Words Matter? Stigmatizing Language and the Transmission of Bias in the Medical Record. J Gen Intern Med 2018 May;33(5):685-691 [FREE Full text] [doi: 10.1007/s11606-017-4289-2] [Medline: 29374357]
13. Robert WJF. County Health Rankings and Roadmaps. 2021. URL: https://www.countyhealthrankings.org/ [accessed 2020-09-01]
14. HRSA Data Warehouse: Find Shortage Areas. 2017. URL: https://data.hrsa.gov/tools/shortage-area [accessed 2020-09-01]
15. Deterding N, Waters M. Flexible Coding of In-depth Interviews: A Twenty-first-century Approach. Sociological Methods & Research 2021;50(2):708-739 [FREE Full text] [doi: 10.1177/0049124118799377]
16. Kolb S. Grounded theory and the constant comparative method: Valid research strategies for educators. Journal of emerging trends in education research and policy studies 2013:83-86 [FREE Full text]
17. Centers for Disease Control and Prevention. Rural Health. 2021. URL: https://www.cdc.gov/ruralhealth/about.html
18. U.S. Department of Health and Human Services. Special Feature on Housing and Health Inequities: Health Equity Report 2019-2020. 2020. URL: https://www.hrsa.gov/about/news/press-releases/health-equity-report-housing [accessed 2021-09-20]
19. Petersen EE, Davis NL, Goodman D, Cox S, Syverson C, Seed K, et al. Racial/Ethnic Disparities in Pregnancy-Related Deaths - United States, 2007-2016. MMWR Morb Mortal Wkly Rep 2019 Sep 06;68(35):762-765 [FREE Full text] [doi: 10.15585/mmwr.mm6835a3] [Medline: 31487273]
20. Takeshita J, Wang S, Loren AW, Mitra N, Shults J, Shin DB, et al. Association of Racial/Ethnic and Gender Concordance Between Patients and Physicians With Patient Experience Ratings. JAMA Netw Open 2020 Nov 02;3(11):e2024583 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.24583] [Medline: 33165609]
21. Alsan M, Garrick O, Graziani G. Does Diversity Matter for Health? Experimental Evidence from Oakland. American Economic Review 2019 Dec 01;109(12):4071-4111. [doi: 10.1257/aer.20181446]

Abbreviations

EHR: electronic health record
ICD: International Classification of Disease
IRB: institutional review board
NASEM: National Academies of Sciences, Engineering, and Medicine
SDoH: social determinants of health

Edited by T Derrick; submitted 05.01.22; this is a non–peer-reviewed article;accepted 14.01.22; published 11.03.22.

Please cite as:
Hirsch A, Durden TE, Silva J
Linking Electronic Health Records and In-Depth Interviews to Inform Efforts to Integrate Social Determinants of Health Into Health Care Delivery: Protocol for a Qualitative Research Study
JMIR Res Protoc 2022;11(3):e36201
URL: https://www.researchprotocols.org/2022/3/e36201
doi:10.2196/36201
PMID:35275090

Protocol

Survivorship of Patients After Long Intensive Care Stay With Exploration and Experience in a New Zealand Cohort (SPLIT ENZ): Protocol for a Mixed Methods Study

Lynsey Sutton[1,2*], MNclin, RNdip; Elliot Bell[2*], BA (Hons), BCA, MA, PGDipClinPsych, PhD; Susanna Every-Palmer[2*], MBChB, MSc, PhD; Mark Weatherall[3*], BA, MBChB, MApplStats; Paul Skirrow[2*], BSc (Hons), MPhil, DClinPsychol, PGDipClinNeuropsych

[1]Intensive Care Unit, Level 3, Wellington Regional Hospital, Capital and Coast District Health Board, Wellington, New Zealand

[2]Department of Psychological Medicine, University of Otago, Wellington, New Zealand

[3]Department of Medicine, University of Otago, Wellington, New Zealand

[*]all authors contributed equally

Corresponding Author:
Lynsey Sutton, MNclin, RNdip
Intensive Care Unit
Level 3, Wellington Regional Hospital
Capital and Coast District Health Board
49 Riddiford Street
Newtown
Wellington, 6021
New Zealand
Phone: 64 211211385
Email: lynsey.sutton-smith@ccdhb.org.nz

Related Article:

This is a corrected version. See correction statement: https://www.researchprotocols.org/2022/4/e38180

Abstract

Background: *Post Intensive Care Syndrome* (PICS) was defined by the Society of Critical Care Medicine in 2012 with subsequent international research highlighting poor long-term outcomes; reduced quality of life; and impairments, for survivors of critical illness. To date, there has been no published research on the long-term outcomes of survivors of critical illness in New Zealand.

Objective: The aim of this study is to explore long-term outcomes after critical illness in New Zealand. The primary objectives are to describe and quantify symptoms and disability, explore possible risk factors, and to identify unmet needs in survivors of critical illness.

Methods: This will be a mixed methods study with 2 components. First, a prospective cohort study of approximately 100 participants with critical illness will be followed up at 1, 6, and 12 months after hospital discharge. The primary outcome will be disability assessed using the World Health Organization Disability Assessment Scale 2.0. Secondary outcomes will focus on mental health using the Hospital Anxiety and Depression Scale and the Impact of Events Scale-revised, cognitive function using the Montreal Cognitive Assessment (Montreal Cognitive Assessment–BLIND), and health-related quality of life using the European Quality of Life-Five Dimension-Five Level. The second element of the study will use qualitative grounded theory methods to explore participants experiences of recovery and highlight unmet needs.

Results: This study was approved by the New Zealand Northern A Health and Disability Ethics Committee on August 16, 2021 (21/NTA/107), and has been registered with the Australian New Zealand Clinical Trials Registry on October 5, 2021. SPLIT ENZ is due to start recruitment in early 2022, aiming to enroll 125 patients over 2 years. Data collection is estimated to be completed by 2024-2025 and will be published once all data are available for reporting.

Conclusions: Although international research has identified the prevalence of PICS and the extent of disability in survivors of critical illness, there is no published research in New Zealand. Research in this field is particularly pressing in the context of COVID-19, an illness that may include PICS in its sequelae.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN1262100133588; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=382566&showOriginal=true&isReview=true

International Registered Report Identifier (IRRID): PRR1-10.2196/35936

(JMIR Res Protoc 2022;11(3):e35936) doi:10.2196/35936

KEYWORDS

COVID-19; critical illness; disability; intensive care unit; survivorship; Post Intensive Care Syndrome

Introduction

Overview

In-hospital mortality for patients treated in intensive care units (ICUs) in New Zealand and Australia is low, with a high proportion of patients surviving to hospital discharge and beyond [1,2]. In part, due to population aging, patients are presenting to the ICU who are increasingly complex [3]. Those with advanced age, medical frailty, and multiple comorbidities are more likely to have high acuity, have long ICU stays, and importunate reliance on multi-organ support and mechanical ventilation [4,5]. The term *persistently critically ill* (PerCI) is used to describe patients who, in addition to a long length of stay in ICU, are at high risk for ongoing disability and poor quality of life after discharge from the hospital [6,7]. This complex group have a variety of initial conditions and are a growing proportion of patients treated in the ICU [2]. Although different cutoff points for length of stay have been used to define them [3,4], it is generally considered that around days 6-10 of the ICU stay represents the juncture at which the patient moves from being acutely ill to being PerCI [4]. Although they account for only 5% of total patients presenting with critical illness in Australia and New Zealand, they use 33% of all ICU bed days and 15% of all hospital bed days [3,4]. Approximately 1 in 6 patients meet the criteria for being PerCI in Australasian ICUs at any time [3,5,6].

The recovery and survivorship journey of these patients is challenging and burdensome. Even before the COVID-19 pandemic, there had been a growing emphasis on the high morbidity and poor quality of life for patients who have been critically ill [8]. A decade ago, Cuthbertson et al [9] reported that quality of life was significantly degraded in patients who are critically ill up to 5 years after critical illness. Herridge et al [10] reported that this was compounded by exercise limitation and physical and psychological sequelae and was associated with increased use of health care services. Other studies have also consistently reported similar poor outcomes across multiple areas of functioning, including mental health, neurocognitive, social, and physical domains [8,11-13]. Since an inaugural stakeholder meeting of critical care experts from the Society of Critical Care Medicine [14], this constellation of poor outcomes following ICU discharge was defined as the *Post Intensive Care Syndrome* (PICS).

After a successful run with a COVID-19 elimination strategy, the spread of the Delta variant prompted the New Zealand government to move to a minimization and protection approach. The impact of this on ICUs in New Zealand is uncertain; however, COVID-19 is likely to become endemic worldwide [15]. Evidence from overseas is that patients who are critically ill with COVID-19 exhibit PICS-like symptoms leading to impairment and disability once they are home [16]. The mechanism for these long-term impairments is still unclear, and debate about what may be long COVID-19 and what is PICS continues [17]. Irrespective of the cause of presentation to the ICU, nearly one-third of patients who are critically ill will go on to experience an element of PICS in their recovery, which warrants urgent attention [12].

Background to the Study

Much of the understanding of PICS comes from follow-up clinics in the United Kingdom and the United States [18]. Published research consistently reports long-term impairments and other negative outcomes in patients with critical illness. Initial studies explored impairments related to specific symptoms such as fatigue, memory loss, cognitive dysfunction, depression, anxiety, posttraumatic stress disorder (PTSD), insomnia, prolonged ICU-acquired neuromuscular weakness, and alopecia [19-36]. However, more recently, a unifying definition has been used to better characterize domains of cognitive, mental health, and functional sequelae for PICS [14,37]. For example, in a multicenter cohort study, Marra et al [12] explored and quantified recovery outcomes in 406 survivors of critical illness. This study did not recruit patients with pre-existing impairments in baseline activities of daily living and cognitive dysfunction; hence, it was able to explore the role of critical illness in the development of new impairments. Three months after discharge, between 25% and 33% of patients had new cognitive impairment, new functional disabilities, or symptoms of depression, with many of these problems persisting after 12 months. Although most patients were assessed as having problems in 1 PICS domain (39% and 35% at 3 and 12 months, respectively), a substantial proportion of patients had problems in 2 domains (19% and 16% after 3 and 12 months, respectively), whereas a smaller group of patients had problems in all 3 domains (6% and 4% at 3 and 12 months, respectively).

Patient-centered outcomes in survivorship research are strongly emphasized [38]. Several studies have reported low quality of life, high levels of disability, and low rates of return to work in relation to PICS [8,9,22,30,39-42]. Hodgson et al [8] explored the concept of PICS and mapped patient difficulties to the World Health Organization International Classification of Functioning Disability and Health. Disability was assessed among 262 Australian survivors of critical illness after 6 months using the World Health Organization Disability Assessment Schedule (WHODAS) 2.0 [43]. The WHODAS was developed to measure disability across six major life domains: cognition, mobility, self-care, interpersonal relationships, work and household roles, and participation in society. The study reported disability was

highly prevalent in survivors after six months, with 75% of participants experiencing variable levels of disability. More specifically, 50% experienced mild disability and 25% had moderate to severe disability. Those with moderate to severe disability were more likely to have a history of depression and anxiety and a longer duration of mechanical ventilation; a worse health-related quality of life; and significant reductions in mobility, personal care, and activities. Furthermore, only 40% of the patients had returned to work or study because of their disabilities, consistent with results reported by other studies [27,39-41,44,45]. Return to work has important implications not only for the individual but also for the family, wider community, and society at large.

The Experience of Recovery

Qualitative work exploring patient accounts of recovery and PICS, compliments and improves the understanding of recovery after ICU treatment. This type of research adds depth and context to quantitative findings and allows patients to tell their own stories of recovery, coping, and transition to a new life. Several qualitative studies report that individual patient experiences can have elements that are unique and that the pace of recovery and trajectory are different from one person to the next [46]. Themes that are common to ICU survivorship include adjustment, acceptance, transition, and liminality (ie, giving up an old life) [47]. Kang and Jeong [48] described survivors' recovery as being characterized by a need to embrace vulnerability, struggling through, moving from crisis to crisis, and progressing onward to a period of acceptance of their new selves. Through coping and internal and external support, survivors gained a new perspective on normality [49]. Only 1 qualitative study has been completed in New Zealand to date, which was solely focused on the ICU experience and interviews with the patient while still in the ICU [50].

Critical illness has a profound long-term impact on patients undergoing recovery. Not only do impairments related to cognition, mental health, and physical function create new and lasting disabilities but quality of life, return to work, and social functioning are also affected. To date, no studies of survivorship or long-term outcomes beyond mortality have been reported in New Zealand [51]. Research on the health system in New Zealand is important because of the specific nature of its population and health care system context. These unique aspects include the relatively low availability of ICU beds in relation to other health systems in high-income countries; the distinctive ethnic makeup within New Zealand society; and elements of the New Zealand health system, such as subsidized primary care services, a no-fault system for support after accidental injury, and the geographic distribution of the population. It is important that New Zealand–based qualitative and quantitative research explores the challenges that ICU patients in New Zealand face once home in recovering, both from and with PICS. Equally pressing is the need to understand the journey of New Zealand's indigenous Māori population. To understand what Māori need to flourish during critical illness recovery, what support is needed and what are the unique health needs of our first nation population is urgently needed. This research protocol outlines and describes a mixed methods study to evaluate and quantify PICS in survivors of critical illness in New Zealand.

Aims and Objectives

The aims of this study are as follows:

1. To estimate the proportion of intensive care survivors with moderate or severe disability at 1, 6, and 12 months after hospital discharge.
2. To describe the data distributions of relevant clinical variables and baseline characteristics of ICU survivors.
3. To understand the survivorship and recovery journey and describe unmet health needs in ICU survivors. A conceptual model of barriers to and facilitators of coping will be developed using grounded theory (qualitative study).

It is anticipated, and will be formally tested, that at least 20% of the adult patients who have had a prolonged stay in the ICU will experience moderate to severe disability in at least one domain of functioning in the year following critical illness.

Trial Design

Quantitative Prospective Cohort Study

This study will be a mixed methods design with 2 components. The first component is a prospective cohort study using validated tools to assess and quantify the level of disability in the 12 months after critical illness. Outcomes will be assessed at three time points: 1, 6, and 12 months after hospital discharge. These times reflect clinically relevant points in the patient's recovery, as identified in past research [31]. Disability will be quantified and assessed according to the WHODAS 2.0 (WHODAS-12). This will be the primary outcome measure.

Secondary outcomes will be health-related quality of life, mental health, and cognition assessed using validated tools at 1, 6, and 12 months. These tools have been chosen based on recommendations from expert committees and international research [37,52,53].

Qualitative Study

The second component is a qualitative study that will explore the process and experience of recovery for participants using the grounded theory [54]. The main purpose of the qualitative study is to identify the ongoing needs (met and unmet) in the year following critical illness, to facilitate the participant to tell their story of recovery, and to tease out the themes that drive recovery and the barriers that prohibit it. The grounded theory has been chosen as the methodology because it is well suited to social science research and has been used in several studies on ICU survivorship using both constructivist and classic approaches [47,48]. It is a rigorous and pragmatic model that incorporates a systematic but flexible approach. Moreover, the grounded theory is particularly well suited to research involving life transitions and the psychological responses to them [55] and has been used effectively alongside quantitative studies in mixed methods research [56]. This design will use a nested convenience sample conducting semistructured interviews between 6 and 8 months after discharge. Recruitment will involve sequential sampling initially, moving to theoretical sampling thereafter. Depending on thematic saturation, the

grounded theory will dictate the number of participants required for the qualitative study.

This mixed methods study design has been chosen to highlight and quantify disability following a critical illness alongside qualitative data to bring context and provide a broad understanding of the recovery journey for survivors in New Zealand. Identifying the ongoing needs, barriers to recovery, and coping strategies will inform the development of resources or interventions to better support the long-term recovery of future patients with critical illness.

Methods

Participants, Interventions, and Outcomes

Study Setting

This will be a single-center prospective cohort study with all participants recruited at a 24-bed, tertiary ICU (Wellington ICU). This unit has a catchment population of approximately 1 million people in central New Zealand, admitting approximately 1800 people per year. Patients can be selected from within a geographic radius of 300 km, and research participants may be from across New Zealand's lower North Island and upper South Island at follow-up. The study may be extended to other New Zealand ICU's if the impact of COVID-19 affects recruitment such that a sample size cannot be reached in Wellington.

Eligibility Criteria and Sample

The study sample will include all adult patients admitted to the Wellington ICU who meet the inclusion criteria within the recruitment period. This is likely to comprise a mix of patients with various conditions and acuity. Where possible, the inclusion and exclusion criteria have been designed to capture patients who are likely to have evidence of PICS while minimizing and excluding the effects of any pre-existing conditions that may worsen because of critical illness.

Inclusion Criteria

All participants will be adult (aged >18 years) ICU patients admitted to the Wellington ICU who have been in an ICU for 7 or more consecutive days or patients who were mechanically ventilated for >72 hours. This inclusion criterion has been designed to capture patients who are most likely to experience poor outcomes once they are home, based on previous research [57]. Patients who have been in another New Zealand ICU or critical care unit before retrieval to the Wellington ICU will be included if both admissions combined is ≥7 days.

Mechanical ventilation is defined as a positive pressure ventilation (PPV) mode via an endotracheal, nasotracheal, or tracheostomy tube. Patients who have been extubated from PPV for a period and then reintubated will also be included if both periods of PPV exceed 72 hours.

Patients with known depression or anxiety will be included because these low prevalence mental health conditions are common, and it would be unreasonable to exclude them. This is a limitation acknowledged in the study by Marra et al [12]. If the participant has experienced previous mental health issues but where mental health information may not be evident from their medical records, this will be ascertained at the first follow-up and the patient will be directly asked the following questions:

- Have you ever been diagnosed with a mental health problem by a physician or psychologist?
- If so, what was the diagnosis they made?

Exclusion Criteria

People are not eligible for the study if they present with the exclusion criteria outlined in Textbox 1.

Textbox 1. Exclusion criteria.

Exclusion criteria

- <18 years
- Non-English speakers
- Not expected to survive their hospital stay as identified by an intensive care unit (ICU) senior medical officer once the inclusion criteria are met
- Have significant challenges for follow-up, for example, are prisoners or are homeless
- Have the following pre-existing conditions:
  - Neuromuscular disorders, for example, muscular dystrophy, multiple sclerosis, myasthenia gravis, or Guillain Barré syndrome
  - Neurodegenerative disease
  - Psychiatric disease or intellectual disability in which patients are already mentally, cognitively, or functionally impaired before ICU admission (identified from the patient's health history)
  - Moderate to severe cognitive impairment, as recorded in the patient electronic health records and medical notes
- Presented to ICU with stroke, neurotrauma, status epilepticus, hypoxic or ischemic brain injury, or encephalopathy
- Have any other disease or disorder which, in the opinion of the principal investigator may influence the result of the trial

Intervention Description

This study has no active intervention but is based on collecting clinical data from eHealth sources and more in-depth responses to the research instruments that are described in the following sections. After informed consent is obtained, baseline data will be collected, including admission details, demographics, ethnicity, and baseline function (Charlson Comorbidity Index and Clinical Frailty Score). Clinical data collected will include diagnosis and the duration of therapies such as intubation, sedation, renal replacement therapy, antibiotic therapy, and mechanical ventilation. Data regarding time to negative polymerase chain reaction test (COVID-19 patients only) and complications during the ICU stay (delirium, ventilator-associated pneumonia, sepsis, and others) will also be collected.

Follow-up will be scheduled at 3 time points during the study. These are at 1, 6, and 12 months after discharge from the hospital. At these follow-up periods, the following tools or outcome measures will be completed by telephone: World Health Organisation Disability Assessment Scale (WHODAS) 2.0, Montreal Cognitive Assessment–BLIND, Hospital Anxiety and Depression Scale, Impact of Events Scale-revised, and European Research Foundation-Five Dimension-Five Level.

Qualitative Component

Participants will be approached at their 6-month follow-up, and after questionnaire completion, they will be invited to be part of a nested sample of the main study. Sequential sampling will be used initially, moving to the theoretical sampling approach, as per the grounded theory [54]. Thematic saturation will dictate the number of participants required for the qualitative study. All interviews will be undertaken by the principal investigator (PI) to ensure consistency. Participants will take part in audio recorded interviews using either face-to-face, Zoom, or telephone interviews, a method that is feasible or preferred for the participant and in line with any government-mandated COVID-19 restrictions.

Criteria for Discontinuing or Modifying Allocated Interventions

Participants will be contacted 1-2 weeks before the time of follow-up. Those who do not respond after 3 phone calls will be considered lost to follow-up and no further data will be sought. Data collected until that point will continue to be used.

Participants who have been readmitted to the hospital at the time of follow-up will continue to be enrolled in the study, with assessments deferred until they are discharged home again. If the participant has died, the date and causes of death will be recorded and data collected up to that point will be used, and the person will be withdrawn from the study.

As part of the informed consent process, participants will be told that they can withdraw at any time and do not need to provide any reason for withdrawal. Participants who wish to withdraw, as communicated verbally or in writing, will be withdrawn immediately. Data collected until that point will still be used and communicated to participants via the information or consent form.

Strategies to Improve Adherence to Interventions

The study has been designed to minimize the burden on participants. First, several methods of obtaining consent have been created so that participants can engage the way they find easiest and most convenient, for example, web-based, paper mail out, mobile phone SMS text messages, or other means. Second, the outcome measures and questionnaires have been chosen based on participant ease, tolerability, and feasibility in mind, while also ensuring that they are responsive and valid. Third, reminder texts and emails will be sent before all follow-ups. Fourth, participants will be able to select their preferred data collection method. If in-person interviews occur, the PI will travel to the participant (there is funding for travel granted in this study). As an additional incentive, a small gift will be offered to the participants in the qualitative study.

Relevant Concomitant Care Permitted or Prohibited During the Study

All relevant concomitant care and interventions will occur during the study in relation to the routine and expected clinical care of the participants, and participants will be advised to continue recovery and rehabilitation as advised by their care provider.

Outcomes Measures for the Quantitative Study

There is currently no standardized, comprehensive tool to measure PICS, with >250 separate tools in existence [58]. Although research into validated tools such as PICS questionnaires is beginning to emerge, this research is limited and not generalizable to the New Zealand population [49]. The outcome measures chosen to emphasize a set of core outcome measures that are valid, responsive, and specific. This will ensure the same reproducible outcome measures, with comparability between studies and the generation of good quality meta-analyses [59].

Several international critical care expert committees have published recommendations in the core outcome sets (COSs) that should be used in research evaluating long-term outcomes in ICU survivors [37,52,53]. In addition, the development of COS for COVID-19 research is also emerging, with an Australian study evaluating 6-month outcomes after COVID-19 using WHODAS, European Research Foundation-Five Dimension-Five Level, Montreal Cognitive Assessment–BLIND, Hospital Anxiety and Depression Scale, and the Impact of Events Scale-6 with success [16]. WHODAS 2.0 and the secondary outcome measures used in both studies by Hodgson et al [8,16] will also be used for SPLIT ENZ, which have been shown to be reliable, valid, and responsive. Additional measures will include mortality and return to work (or study) (as assessed using the WHODAS 2.0). With guidance from the COS and other high-quality published research [8,12,16,37,52,53], the outcome measures chosen for this prospective cohort study are summarized in Table 1.

Table 1. Primary and secondary outcome measures.

Tool	Domain measured	Use and scoring
Primary outcome measure		
WHODAS[a] 2.0 [43]	Function	The 12-item WHODAS 2.0 covers 6 domains of functioning with scores from 0 (no difficulty) to 4 (extreme difficulty). The total score between 0 and 48, is then divided by 48 and multiplied by 100 to convert it to a percentage of maximum disability as follows: no disability (0%-4%), mild disability (5%-24%), moderate disability (25%-49%), severe disability (50%-95%), and complete disability 96% to 100%.
Secondary outcome measures		
EQ-5D-5L[b] [60]	Health-related quality of life	Measured in five domains: mobility, personal care, usual activities, pain, anxiety, and depression. Each dimension has 5 levels (*no problems*=1 to *extreme problems*=5). The EQ-5D-5L consists of two pages: the EQ-5D descriptive system and the EQ VAS[c]. The EQ VAS is used as a measure of overall self-rated health status as a numerical score.
HADS[d] [61]	Depression and anxiety	The HADS contains fourteen questions: 7 to assess anxiety and 7 for depression. For the 14 questions, a 4-point Likert scale (range 0-3) gives a possible score of 0 (none) to 21 (severe) for each of the two subscales: 0-7 indicate normal or no anxiety or depression symptoms, ≥8 to 10 indicate clinically significant anxiety or depression symptoms (borderline cases), and ≥11 indicate severe psychological distress.
IES-r[e] [62]	PTSD[f]	There are 22 questions that cover the three diagnostic clusters: intrusion; avoidance (8 questions each); and hyperarousal (6 questions). Respondents report on a 5-point Likert scale: *not at all* (item score 0) to *extremely* (4) how distressed they have been in the past 7 d in relation to a specific event. The IES-r yields a total score (ranging from 0 to 88) and subscale scores can also be calculated for the intrusion, avoidance, and hyperarousal subscales. The total mean IES-r score is the sum of the means of the 3 subscale scores. The maximum mean score on each of the 3 subscales is *4*, therefore the maximum *total mean* IES-r score is 12. A total IES-r score of 33 or more from a theoretical maximum of 88 signifies the likely presence of PTSD.
MOCA[g]-BLIND [63]	Cognitive function	The total possible score is 22 points; a score of 18 or above is considered normal. Cutoffs have not been validated in patients who are critically ill.

[a]WHODAS: World Health Organization Disability Assessment Schedule.

[b]EQ-5D-5L: European Research Foundation-Five Dimension-Five Level.

[c]EQ VAS: EuroQol visual analogue scale.

[d]HADS: Hospital Anxiety and Depression Scale.

[e]IES-r: Impact of Events Scale-revised.

[f]PTSD: posttraumatic stress disorder.

[g]MOCA: Montreal Cognitive Assessment.

Participant Timeline

The timeline of SPLIT ENZ study schedule is shown in Figure 1.

Figure 1. Timeline of SPLIT ENZ study schedule. ICU: intensive care unit.

	Enrollment		Follow-up time points (post hospital discharge)		
Time point	Day 7 ICU stay or >72 hours mechanical ventilation	Before hospital discharge	2-4 weeks	6 months	12 months
Enrollment:	√				
Eligibility screen	√				
Informed consent		√ Cohort study		√ Qualitative study	
Participant Interviews and Qualitative study				√	
Allocation	Not applicable				
Baseline demographic & clinical data collected	←	→			
Assessments:					
World Health Organization Disability Assessment Schedule 2.0 (WHODAS)			√	√	√
Hospital Anxiety & depression Scale (HADS)			√	√	√
Impact of Events Scale-revised (IES-r)			√	√	√
Montreal Cognitive Assessment-Blind (MOCA-blind)			√	√	√
European Quality of Life-Five Dimension-Five level (EQ-5D-5L)			√	√	√

Sample Size—Quantitative Study

All adult patients admitted to the Wellington ICU between February 2022 and February 2024 who are eligible will be invited to participate. Admission modeling suggests that this will give approximately 100 potential participants. The anticipated sample size of 100 gives a 95% CI for a proportion of plus or minus 10%. This sample size has 80% power to test that the proportion with PICS in less than approximately 10% or more than approximately 30% based on an assumed proportion of 20%.

Recruitment

Overview

Despite the broad inclusion criteria, it is likely that recruitment may be slow and that completion to follow-up may also be low. Slow recruitment may reflect the patient cohort, which is typically the emphasis on studies of PICS and likely so with this study. These patients, historically termed PerCI [3], have the longest stays in the ICU but reflect only 5% of all patient admissions to the ICUs in Australia and New Zealand. In addition, New Zealand has the lowest proportion of ICU beds per 100 compared with the rest of the Organisation for Economic Co-operation and Development countries [64], which affects overall admission rates and potential recruitment of participants.

It is anticipated that a significant proportion of patients may also fail to reach study completion because of in-hospital death (but after ICU discharge) and death at home. International studies report variable mortality rates in the post-ICU year, between 20% and 50% for a variety of critically ill subpopulations [13,57,65]. One subpopulation in particular, patients with COVID-19 are also reported to have mortality

rates as high as 30%, 6 months after discharge, and ICU mortality rates between 50% and 97% [66]. During the course of the study, the application of new therapies may change the survival experience of these patients with an uncertain effect on recruitment and follow-up [67]. The SPLIT ENZ team acknowledges these limitations and aims to ensure as many participants complete follow-up, by extending the recruitment period for up to 2 years. This should ensure that sufficient participants have data available at 6 months.

Qualitative Study Sample Size

Participants will be approached to participate in the qualitative study at the scheduled 6-month follow-up for the quantitative study. This sample will be generated sequentially from the main quantitative cohort (nested sample) and will be consented separately. Participants will be approached at the 6-month telephone follow-up, and verbal and written consent will be sought. Recruitment will continue, and interviews are completed until (thematic) data saturation is reached. As the grounded theory moves through the process of continuous recruitment, data collection, analysis, coding, and thematic exploration simultaneously, these processes inform the number of participants required to reach saturation of themes [54]. In this instance, the sample size is not preset, but it is likely to include an estimated sample of 15-20 participants.

Significance to Māori

Between 10% and 20% of patients admitted to the Wellington ICU each year are of Māori ethnicity. In New Zealand, Māori account for 17% of the total population [68] and experience health disparities across most major health, education, and psychosocial sectors. High rates of inequity, socioeconomic disparity, and increased barriers to access at all levels of health care exist for Māori [68]. Therefore, it is important that this study sample is reflective and inclusive to ensure that the voice of Māori participants is heard, something that is currently missing in New Zealand. The qualitative aspect of this study will be an important method by which we may understand the recovery journey, understand what helps Māori to flourish, and identify what unmet needs and barriers to recovery remain. Boosted sampling will be used to ensure that a high number of Māori participants are included in the qualitative part of the study sample. The researchers consulted with the Wellington, New Zealand Research Advisory Group for Māori, the Kaupapa Māori research network, and the Ngāi Tahu Research Consultation Committee at Otago University. This research protocol has been supported by these groups.

Data Collection, Management, and Analysis

Data Collection Completed Once Deemed Eligible and Once the Patient has Consented

The following data will be extracted from the ICU database, electronic and paper-based patient notes, and ICU observation charts (Textbox 2).

Textbox 2. Data to be collected once eligible and after recruitment.

Initial details to be collected once patient meets eligibility criteria

- Key intensive care unit (ICU) and hospital admission and discharge dates
- Contact details for patient and next of kin
- Demographics
- Diagnosis

Subsequent data collection once patient consented

- Clinical frailty score
- Apache II and Apache III scores
- Charlson Comorbidity Index and types of comorbidities
- Admission sequential organ failure score

Length of stay and funding

- Accident Compensation Corporation funding (yes or no)
- Hospital length of stay
- ICU length of stay

Measures of clinical status or acuity

- Number and type of clinical complications in the ICU (described and listed)
- PaO_2/FiO_2 ratio per day while mechanically ventilated
- Number of reintubations/failed extubation or extubations during ICU stay

Duration (hours/days) of interventions

- Renal replacement therapy (continuous or intermittent dialysis), extracorporeal membrane oxygenation, vasopressors, invasive hemodynamic monitoring, and antibiotic therapy duration
- Time to negative polymerase chain reaction test and if isolated for COVID-19
- Daily oxygen therapy, specifying high-flow nasal prongs and low-flow oxygen (ie, via nasal prongs), noninvasive ventilation and direct tracheostomy interface
- Oxygen free days
- Mechanical ventilation duration hours/days and mode
- Time to extubation and tracheostomy decannulation
- Tracheostomy weaning duration
- Sedation types and doses per day
- Mean Richmond Agitation Sedation scores per day
- Paralysis agents (neuromuscular blocking agents) doses and type per day
- Delirium duration hours/days as evidenced by the Confusion Assessment Method for ICU scores

Other

- Best daily ICU mobility score (if recorded)
- Chelsea Critical Care Physical Assessment Score (if completed at ICU discharge)
- If a patient diary was assigned to the patient in the ICU

Data to Be Collected at Follow-up

Follow-up data will consist of questionnaires administered by the PI over the telephone, using the primary and secondary outcome measures described in Table 1. The data will be recorded on data collection forms in accordance with the tool creators' specifications after any required training has occurred. In addition, the participants will also be asked if they have tested positive for COVID-19, as this may influence the results of questionnaires.

Qualitative Study Data

The following key themes are explored based on previous research highlighting themes common to the recovery experience [47,48]:

- Experience of recovery overall
- Coping strategies and things that facilitate recovery (eg, spiritual, family, and social connectivity)
- Barriers, including met or unmet needs

Grounded theory will be undertaken to guide data collection during semistructured recorded interviews. Several data collection strategies will be used during the process: simultaneous collection and analysis of data, open and axial coding with comparative analysis (within cases and across cases), refining theoretical ideas, and memo writing [54,69]. Interviews will be audio recorded transcribed verbatim, coded, and checked for accuracy with another researcher. The PI will independently read and code the transcripts, the codes will be examined, and by an iterative process, the codes will be condensed into similar themes [69]. A second researcher will check the transcripts for truth and completeness. To achieve saturation of the themes, researchers will move back and forth between data collection and analysis, reidentifying themes and subthemes [69]. Work completed early in the study will inform subsequent recruitment using theoretical sampling data collection and analysis.

Outcome Measures for Qualitative Study

The following outcomes or themes were used to guide the interviews: ultimately, as part of the qualitative study, we will present a conceptual model of barriers and facilitators to coping for ICU survivors.

Plans to Promote Participant Retention and Complete Follow-up

Strategies will be used to maximize continued recruitment throughout the study. These include ensuring that consent is easy, follow-up is not arduous, and patient or family preferences are understood. Regular contact will also be made with participants to ensure engagement throughout the study period (text or mail reminders), and clear written information will be delivered before follow-up.

Data Management

Patients will be allocated a unique study number by the PI. Study data and enrollment logs will be kept separate in a locked research office at the study center. Once the unique participant number is allocated, documentation will be deidentified and referred to only using that number. Data will be coded and entered into Microsoft Excel spreadsheets and Microsoft Word documents. Data will be stored securely in a locked research office for 10 years or in a password-protected secure computer file.

Results

Statistical Methods for Primary and Secondary Outcomes

Overview

Continuous variables will be described by mean and SD, median and IQR, and minimum to maximum values. Appropriate frequency histograms and boxplots will also be used to summarize the data distributions. Categorical variables will be described by numerators and denominators, and proportions will be expressed as percentages.

Proportions will be estimated together with CIs using standard binomial methods. It is anticipated that asymptotic methods for the CIs will be satisfactory; however, if there are many small frequency counts, exact binomial methods will also be used (aim I).

Associations between disability measured by the WHODAS and potential univariate predictors will be examined by logistic regression with disability categorized and moderate or severe versus lesser degrees of disability. As a sensitivity analysis, WHODAS will be treated on a continuous scale and ordinary regression used. For the latter, normality assumptions will be assessed by residual analyses to determine if a data transformation will be needed or if another form of regression such as ordinal regression might be more suitable. With an anticipated 25-30 participants with moderate or severe disability, this gives limited scope for multivariate analysis, but as discussed in the following sections, a more limited number of potential predictors will be used in a multivariate model to determine if associations remain after adjusting for confounding.

We selected a priori the following covariates as potential predictors of disability after intensive care admission: age, gender, ethnicity, APACHE II and III score, sequential organ failure score, duration of sepsis (days), frailty score, Charlson Comorbidity Index, length of ICU stay reported in days, duration of mechanical ventilation (in hours), duration of delirium (reported as days the patient was Confusion Assessment Method for the ICU positive), total doses of sedation, use of benzodiazepines, use of neuromuscular blocking agents, previous history of depression, anxiety, or PTSD (obtained from medical records, ICU database admission, and Medical Admissions Portal) and whether the ICU admission was for cardiothoracic surgery whereby the patient underwent cardiopulmonary bypass. Drug doses (sedative agents, benzodiazepines, and neuromuscular blocking agents) will be transformed into mean doses per day and analyzed over the number of days received.

Each of these potential predictors will be examined using univariate predictors with accompanying illustrative plots. In general, the analysis strategy will treat disability as a dichotomous variable and use logistic regression to estimate odds ratios for association and as discussed to explore linear regression and ordinal regression treating the WHODAS as a continuous response variable and possibly as an ordinal response variable. Although the primary interest is in disability after 12 months, the associations at earlier points, 1 month and 6 months,

will also be estimated. At least one study has categorized disability based on the WHODAS-12 as none, mild, and moderate to severe disability [70]; however, it is likely to be more useful to explore if the instrument can be used on its native scale or use ordinal regression based on the full range of scores rather than other cutoff values.

Although not directly related to the study aims, mortality will also be assessed using the Kaplan–Meier curves and associated estimates of median or, where relevant, other percentiles, survival.

Subgroup Analyses

Subgroups of patients of interest will be those with a history of depression, anxiety or PTSD; those who have undergone postcardiac surgery and cardiopulmonary bypass in the preceding 3 months before ICU admission (and are at high risk for postoperative cognitive dysfunction) [71], those categorized by COVID-19 and non–COVID-19 illness, and ethnicity. Whether proportions with PICS or other outcomes differ in relation to these subgroups will be explored using interaction terms in the regression analyses.

Data Monitoring

Data monitoring for quality will be maintained within the SPLIT ENZ team, and there is no formal outside data monitoring committee that will be used as part of this study. No formal interim analysis will be performed as part of this study.

Reporting Adverse Events and Patient Safety

Any adverse events as part of this study will be discussed and managed within the SPLIT ENZ team and documented in the study file notes or data collection sheet. If there is a patient safety issue, this will be managed by the coordinating investigator in conjunction with the SPLIT ENZ team, and the escalation plan will be followed and documented.

There are no provisions for posttrial care, as this is a low-risk study. However, it is anticipated there may be a proportion of patients who require ongoing support from their health care provider (general practitioner [GP]) during the follow-up period. If the patient is found to be in distress with unmanageable symptoms at any of the follow-up periods, consent will be sought to contact the patient's GP will be asked to provide help and treatment where possible. The PI will be responsible for ensuring that the GP is contacted and made aware, and a note to file will be created in the patient's notes. At all study follow-up calls, the participant will be encouraged to have a support person with them if they wish.

Ethics, Consent, and Dissemination

Research Ethics Approval and Protocol Amendments

This study has received full ethics approval from the New Zealand Northern A Health and Disability Ethics Committee on August 16, 2021 (21/NTA/107), and has been registered with the Australian New Zealand Clinical Trials Registry on October 5, 2021 (12621001335886). In light of the sudden outbreak of the Delta and Omicron strain of COVID-19 in New Zealand in August 2021, further protocol amendments were sought to ensure the overall study design, consent processes, and outcome measures were appropriate for patients admitted to the ICU with both COVID-19 and non–COVID-19 disease. Approval was granted for the protocol amendments on November 10, 2021, from the New Zealand Northern A Health and Disability Ethics Committee.

Any further changes to the protocol that may affect the conduct of the study or patient safety, including modifications to study objectives, study design, patient population, sample sizes, or study procedures, will require formal amendments to the protocol. Any such amendments will be agreed to by the SPLIT ENZ research group and must be approved by the Health and Disability Ethics Committee before implementation and notified to the health authorities and study sites in accordance with local regulations. The Australian New Zealand Clinical Trials Registry will also be updated with any relevant changes or updates.

Minor changes to the protocol (eg, corrections and clarifications that have no effect on the way the study is conducted) will be agreed to by the research group and logged on the Health and Disability Ethics Committee website.

Consent

During the recruitment period, the ICU patient database will be screened daily by the PI. Once potential participants have been identified as meeting the inclusion criteria, they will be screened to ensure that none of the exclusion criteria apply. If the person is eligible for recruitment, they and their families will be approached by the PI. It is highly likely that patients will be unable to provide consent at that time, so their family will be given information about the study and informed that once the patient is able to give consent, they will be approached by the PI. All consent processes will be managed by the PI.

An approach consistent with section 7.4 of the New Zealand Health and Disability Code [72] will be used to obtain consent. However, it is acknowledged that the consent process may require contingencies with the uncertainty around national COVID-19 restrictions and future lockdowns. There are 2 processes and contingencies to obtain consent in these circumstances. The following section outlines the different scenarios and their contingencies.

When Consent Can Be Gained in Hospital

Consent will be gained from prospective participants before discharge from the hospital (while recovering on a ward) when they are competent to give informed consent (ie, they are able to comprehend and communicate their wishes). The patient and family will be given all the necessary information about the study by the PI in a face-to-face meeting, and consent forms will be provided to the patient. The patient will be given sufficient time to think about participation, and if they are willing to take part, signed consent forms will be collected.

Where Consent Cannot Be Gained in Hospital

If the patient cannot consent while in Wellington hospital or they are transferred to another facility or domicile before the opportunity to gain written consent occurs, there are several options that will be used. First, an initial phone call to engage with the patient and the family will be made. If verbal consent

is given, the consent form can be mailed out or emailed to participants with written information about the study. Prepaid envelopes will be provided for mailing back to the coordinating center. If the patient finds it easier to take a photo of the signed page and send it this way, it can be done via email or texting back to the PI. For participants who prefer an electronic method, a web-based consent process via the REDCap (Research Electronic Data Capture; Vanderbilt University) system has also been created.

Qualitative Study Consent

At the 6-month follow-up, the participant will be invited to further partake in qualitative interviews with the PI. All information will be mailed or emailed to the potential participants if they indicated that they would like to take part.

Confidentiality

All personal, demographic, and identifying data will be managed as stated under the *Data Management* section. Consent will be sought at the time of follow-up to alert the participant's family physician if the researchers become aware of urgent or follow-up care that is needed, and this discussion will remain confidential. All reporting and publication of data will contain fully deidentified data and will undergo stringent peer review before publication in accordance with the journal's editorial policy.

Access to Data

Patients will be allocated a unique study number by the PI. Study data and enrolment logs will be kept separate in a locked research office at the study center. Once the unique participant number is allocated, documentation will be deidentified from there on in and referred to using that number. The data will be archived for 10 years in accordance with the University of Otago guidelines. Only the PI will have access to the raw data, but the wider SPLIT ENZ team will have access to the final data set that will be published once the study concludes. Outside investigators may contact the PI for any data at the conclusion of the study and the publication of the results.

Composition of the Coordinating Center

The coordinating center consists of the PI supported by the SPLIT ENZ supervisory team (PS, SEP, EB, and MW). Although there is no formal outside trial steering committee for this study, all trial design queries and quality control oversight will be undertaken with guidance from the SPLIT ENZ supervisory team. If any further advice is required outside of this team, the PI will approach the ICU research team, where the PI is located.

Dissemination Plan

The study protocol is registered with the Australia and New Zealand Clinical Trials registration and is freely available on the web.

At consent, participants will be asked if they would like a copy of the study results, and this will be shared with them after study completion.

Data will be freely shared (with appropriate privacy and confidentiality oversight) at the completion of the study with other health professionals, researchers, and study sponsors as applicable.

Discussion

Overview

This study will be the first of its kind in New Zealand and will contribute to understanding of the challenges and level of disability patients face once home, recovering from a critical illness. Despite the appropriate and commendable government response to COVID-19 and the emergence of the Delta, Omicron, and possible future variants, New Zealand traverses a likely future with endemic COVID-19. It is crucial that research is undertaken to understand recovery, survivorship, and quality of life after critical illness. Although this study was originally designed to focus on patients without COVID-19, it is important to recognize that all patients who are critically ill are vulnerable to PICS from a variety of causes and illnesses; hence, both are eligible to be included in this study. Although some authors have postulated both PICS and long COVID-19 are the same entity [73], it is important to highlight that there are some tangible differences among them. For example, there is evidence that many patients infected with SARS-CoV-2 undergo significant acute lung injury leading to chronic dyspnea, with fatigue, mental health and memory problems persisting even in young home-isolated adults who were not even admitted to the ICU [74]. This study will include both patients with COVID-19 and non–COVID-19, with outcomes analyzed separately owing to the differences with long COVID-19 and PICS. Nonetheless, the overall emphasis is on highlighting the recovery journey for New Zealand survivors of critical illness, irrespective of cause of illness.

Conclusions

SPLIT ENZ is due to begin recruitment at the start of 2022 with an intended recruitment period of no longer than 2 years. The current protocol (version 2.5; November 15, 2021) has undergone an ethics review and Māori consultation and is registered with the Australian New Zealand Clinical Trials Registry.

Acknowledgments

The authors would like to thank Wellington Medical Research and Research for Life (NZ$6490 [US $4340.84]) and the Perpetual Guardian Nursing fund (NZ$10,000 [US $6688.51]) for providing grants for this study. The grant funding body had no direct role in the design of the study nor did they have any role during its execution; collection, analysis, and interpretation of data; writing the manuscript; or the decision to submit results.

Authors' Contributions

LS is the principal investigator of the trial. LS and PS conceived the idea of the study. LS designed the trial and coordinated the main preparation of the study documents and the grant funding applications with inputs from PS, SEP, and EB. LS and PS led the development of data collection processes. MW assisted in designing a plan for statistical analysis. LS drafted the first version of the study protocol.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer Review Report from the Health and Disability Ethics Committees, Ministry of Health (New Zealand).
[PDF File (Adobe PDF File), 237 KB - resprot_v11i3e35936_app1.pdf]

References

1. Paul E, Bailey M, Kasza J, Pilcher DV. Assessing contemporary intensive care unit outcome: development and validation of the Australian and New Zealand risk of death admission model. Anaesth Intensive Care 2017 May 01;45(3):326-343. [doi: 10.1177/0310057x1704500308]
2. Corke C, Leeuw ED, Lo SK, George C. Predicting future intensive care demand in Australia. Crit Care Resusc 2009 Dec;11(4):257-260. [Medline: 20001873]
3. Darvall JN, Boonstra T, Norman J, Murphy D, Bailey M, Iwashyna TJ, et al. Persistent critical illness: baseline characteristics, intensive care course, and cause of death. Crit Care Resusc 2019 Jun;21(2):110-118. [Medline: 31142241]
4. Iwashyna TJ, Hodgson CL, Pilcher D, Bailey M, van Lint A, Chavan S, et al. Timing of onset and burden of persistent critical illness in Australia and New Zealand: a retrospective, population-based, observational study. Lancet Respiratory Med 2016 Jul;4(7):566-573. [doi: 10.1016/s2213-2600(16)30098-4]
5. Darvall JN, Bellomo R, Paul E, Subramaniam A, Santamaria JD, Bagshaw SM, et al. Frailty in very old critically ill patients in Australia and New Zealand: a population-based cohort study. Med J Aust 2019 Oct;211(7):318-323. [doi: 10.5694/mja2.50329] [Medline: 31489652]
6. Bagshaw SM, Stelfox HT, Iwashyna TJ, Bellomo R, Zuege D, Wang X. Timing of onset of persistent critical illness: a multi-centre retrospective cohort study. Intensive Care Med 2018 Dec;44(12):2134-2144. [doi: 10.1007/s00134-018-5440-1] [Medline: 30421256]
7. Jeffcote T, Foong M, Gold G, Glassford N, Robbins R, Iwashyna TJ, et al. Patient characteristics, ICU-specific supports, complications, and outcomes of persistent critical illness. J Crit Care 2019 Dec;54:250-255. [doi: 10.1016/j.jcrc.2019.08.023] [Medline: 31630075]
8. Hodgson CL, Udy AA, Bailey M, Barrett J, Bellomo R, Bucknall T, et al. The impact of disability in survivors of critical illness. Intensive Care Med 2017 Jul;43(7):992-1001. [doi: 10.1007/s00134-017-4830-0] [Medline: 28534110]
9. Cuthbertson BH, Roughton S, Jenkinson D, Maclennan G, Vale L. Quality of life in the five years after intensive care: a cohort study. Crit Care 2010;14(1):R6 [FREE Full text] [doi: 10.1186/cc8848] [Medline: 20089197]
10. Herridge MS, Tansey CM, Matté A, Tomlinson G, Diaz-Granados N, Cooper A, Canadian Critical Care Trials Group. Functional disability 5 years after acute respiratory distress syndrome. N Engl J Med 2011 Apr 07;364(14):1293-1304. [doi: 10.1056/NEJMoa1011802] [Medline: 21470008]
11. Jackson JC, Pandharipande PP, Girard TD, Brummel NE, Thompson JL, Hughes CG, et al. Depression, post-traumatic stress disorder, and functional disability in survivors of critical illness in the BRAIN-ICU study: a longitudinal cohort study. Lancet Respiratory Med 2014 May;2(5):369-379. [doi: 10.1016/s2213-2600(14)70051-7]
12. Marra A, Pandharipande PP, Girard TD, Patel MB, Hughes CG, Jackson JC, et al. Co-occurrence of post-intensive care syndrome problems among 406 survivors of critical illness. Crit Care Med 2018 Sep;46(9):1393-1401 [FREE Full text] [doi: 10.1097/CCM.0000000000003218] [Medline: 29787415]
13. Haines KJ, Berney S, Warrillow S, Denehy L. Long-term recovery following critical illness in an Australian cohort. J Intensive Care 2018;6:8 [FREE Full text] [doi: 10.1186/s40560-018-0276-x] [Medline: 29445502]
14. Needham DM, Davidson J, Cohen H, Hopkins RO, Weinert C, Wunsch H, et al. Improving long-term outcomes after discharge from intensive care unit. Critical Care Med 2012;40(2):502-509. [doi: 10.1097/ccm.0b013e318232da75]
15. Phillips N. The coronavirus is here to stay - here's what that means. Nature 2021 Feb 16;590(7846):382-384. [doi: 10.1038/d41586-021-00396-2] [Medline: 33594289]
16. Hodgson CL, Higgins AM, Bailey MJ, Mather AM, Beach L, Bellomo R, COVID-Recovery Study Investigatorsthe ANZICS Clinical Trials Group. The impact of COVID-19 critical illness on new disability, functional outcomes and return to work at 6 months: a prospective cohort study. Crit Care 2021 Nov 08;25(1):382 [FREE Full text] [doi: 10.1186/s13054-021-03794-0] [Medline: 34749756]
17. Jaffri A, Jaffri UA. Post-Intensive care syndrome and COVID-19: crisis after a crisis? Heart Lung 2020;49(6):883-884 [FREE Full text] [doi: 10.1016/j.hrtlng.2020.06.006] [Medline: 32690219]

18. Lasiter S, Oles SK, Mundell J, London S, Khan B. Critical care follow-up clinics. Clin Nurse Spec 2016;30(4):227-237. [doi: 10.1097/nur.0000000000000219]
19. Honarmand K, Lalli RS, Priestap F, Chen JL, McIntyre CW, Owen AM, et al. Natural history of cognitive impairment in critical illness survivors. A systematic review. Am J Respir Crit Care Med 2020 Jul 15;202(2):193-201. [doi: 10.1164/rccm.201904-0816ci]
20. Collet MO, Egerod I, Thomsen T, Wetterslev J, Lange T, Ebdrup BH, et al. Risk factors for long-term cognitive impairment in ICU survivors: a multicenter, prospective cohort study. Acta Anaesthesiol Scand 2021 Jan 12;65(1):92-99. [doi: 10.1111/aas.13692] [Medline: 32852053]
21. Hopkins RO, Suchyta MR, Kamdar BB, Darowski E, Jackson JC, Needham DM. Instrumental activities of daily living after critical illness: a systematic review. Annals ATS 2017 Aug;14(8):1332-1343. [doi: 10.1513/annalsats.201701-059sr]
22. Jackson JC, Hart RP, Gordon SM, Hopkins RO, Girard TD, Ely EW. Post-traumatic stress disorder and post-traumatic stress symptoms following critical illness in medical intensive care unit patients: assessing the magnitude of the problem. Crit Care 2007;11(1):R27 [FREE Full text] [doi: 10.1186/cc5707] [Medline: 17316451]
23. Parker AM, Sricharoenchai T, Raparla S, Schneck KW, Bienvenu OJ, Needham DM. Posttraumatic stress disorder in critical illness survivors. Crit Care Med 2015;43(5):1121-1129. [doi: 10.1097/ccm.0000000000000882]
24. Nikayin S, Rabiee A, Hashem MD, Huang M, Bienvenu OJ, Turnbull AE, et al. Anxiety symptoms in survivors of critical illness: a systematic review and meta-analysis. Gen Hosp Psychiatry 2016;43:23-29 [FREE Full text] [doi: 10.1016/j.genhosppsych.2016.08.005] [Medline: 27796253]
25. Ohtake PJ, Lee AC, Scott JC, Hinman RS, Ali NA, Hinkson CR, et al. Physical impairments associated with post-intensive care syndrome: systematic review based on the World Health Organization's international classification of functioning, disability and health framework. Phys Ther 2018 Aug 01;98(8):631-645. [doi: 10.1093/ptj/pzy059] [Medline: 29961847]
26. Battle C, Lynch C, Thorpe C, Biggs S, Grobbelaar K, Morgan A, et al. Incidence and risk factors for alopecia in survivors of critical illness: a multi-centre observational study. J Crit Care 2019 Apr;50:31-35. [doi: 10.1016/j.jcrc.2018.11.015] [Medline: 30471558]
27. Brandstetter S, Dodoo-Schittko F, Blecha S, Sebök P, Thomann-Hackner K, Quintel M, et al. Influence of quality of care and individual patient characteristics on quality of life and return to work in survivors of the acute respiratory distress syndrome: protocol for a prospective, observational, multi-centre patient cohort study (DACAPO). BMC Health Serv Res 2015 Dec 17;15:563 [FREE Full text] [doi: 10.1186/s12913-015-1232-2] [Medline: 26677970]
28. Cox CE, Carson SS, Lindquist JH, Olsen MK, Govert JA, Chelluri L, Quality of Life After Mechanical Ventilation in the Aged (QOL-MV) Investigators. Differences in one-year health outcomes and resource utilization by definition of prolonged mechanical ventilation: a prospective cohort study. Crit Care 2007;11(1):R9 [FREE Full text] [doi: 10.1186/cc5667] [Medline: 17244364]
29. Dettling-Ihnenfeldt DS, de Graaff AE, Beelen A, Nollet F, van der Schaaf M. Coping style and quality of life in Dutch intensive care unit survivors. Rehabil Psychol 2016 May;61(2):165-172. [doi: 10.1037/rep0000084] [Medline: 26938222]
30. Dowdy DW, Eid MP, Sedrakyan A, Mendez-Tellez PA, Pronovost PJ, Herridge MS, et al. Quality of life in adult survivors of critical illness: a systematic review of the literature. Intensive Care Med 2005 Jun 11;31(7):1007. [doi: 10.1007/s00134-005-2704-3]
31. Griffiths J, Hatch RA, Bishop J, Morgan K, Jenkinson C, Cuthbertson BH, et al. An exploration of social and economic outcome and associated health-related quality of life after critical illness in general intensive care unit survivors: a 12-month follow-up study. Crit Care 2013 May 28;17(3):R100 [FREE Full text] [doi: 10.1186/cc12745] [Medline: 23714692]
32. Hofhuis J, Spronk P, van Stel HF, Schrijvers A, Rommes J, Bakker J. The impact of severe sepsis on health-related quality of life: a long-term follow-up study. Anesth Analg 2008 Dec;107(6):1957-1964. [doi: 10.1213/ane.0b013e318187bbd8] [Medline: 19020144]
33. Honselmann KC, Buthut F, Heuwer B, Karadag S, Sayk F, Kurowski V, et al. Long-term mortality and quality of life in intensive care patients treated for pneumonia and/or sepsis: predictors of mortality and quality of life in patients with sepsis/pneumonia. J Crit Care 2015 Aug;30(4):721-726. [doi: 10.1016/j.jcrc.2015.03.009] [Medline: 25818842]
34. Oeyen SG, Vandijck DM, Benoit DD, Annemans L, Decruyenaere JM. Quality of life after intensive care: a systematic review of the literature. Crit Care Med 2010 Dec;38(12):2386-2400. [doi: 10.1097/CCM.0b013e3181f3dec5] [Medline: 20838335]
35. Parsons EC, Hough CL, Vitiello MV, Zatzick D, Davydow DS. Insomnia is associated with quality of life impairment in medical-surgical intensive care unit survivors. Heart Lung 2015 Mar;44(2):89-94. [doi: 10.1016/j.hrtlng.2014.11.002] [Medline: 25592203]
36. Wintermann G, Rosendahl J, Weidner K, Strauß B, Hinz A, Petrowski K. Fatigue in chronically critically ill patients following intensive care - reliability and validity of the multidimensional fatigue inventory (MFI-20). Health Qual Life Outcomes 2018 Feb 20;16(1):37 [FREE Full text] [doi: 10.1186/s12955-018-0862-6] [Medline: 29463245]
37. Mikkelsen M, Still M, Anderson B, Bienvenu OJ, Brodsky MB, Brummel N, et al. Society of Critical Care Medicine's International Consensus Conference on prediction and identification of long-term impairments after critical illness. Crit Care Med 2020 Nov;48(11):1670-1679. [doi: 10.1097/CCM.0000000000004586] [Medline: 32947467]

38. Dinglas VD, Chessare CM, Davis WE, Parker A, Friedman LA, Colantuoni E, et al. Perspectives of survivors, families and researchers on key outcomes for research in acute respiratory failure. Thorax 2018 Jan 29;73(1):7-12 [FREE Full text] [doi: 10.1136/thoraxjnl-2017-210234] [Medline: 28756400]
39. Heydon E, Wibrow B, Jacques A, Sonawane R, Anstey M. The needs of patients with post-intensive care syndrome: a prospective, observational study. Aust Crit Care 2020 Mar;33(2):116-122. [doi: 10.1016/j.aucc.2019.04.002] [Medline: 31160217]
40. Hodgson CL, Haines KJ, Bailey M, Barrett J, Bellomo R, Bucknall T, ICU-Recovery Investigators. Predictors of return to work in survivors of critical illness. J Crit Care 2018 Dec;48:21-25. [doi: 10.1016/j.jcrc.2018.08.005] [Medline: 30138904]
41. Norman BC, Jackson JC, Graves JA, Girard TD, Pandharipande PP, Brummel NE, et al. Employment outcomes after critical illness. Critical Care Medicine 2016;44(11):2003-2009. [doi: 10.1097/ccm.0000000000001849]
42. Deacon KS. Re-building life after ICU: a qualitative study of the patients' perspective. Intensive Crit Care Nurs 2012 Apr;28(2):114-122. [doi: 10.1016/j.iccn.2011.11.008] [Medline: 22390918]
43. Üstün TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization disability assessment schedule 2.0. Bull World Health Organ 2010 May 20;88(11):815-823. [doi: 10.2471/blt.09.067231]
44. Kamdar BB, Suri R, Suchyta MR, Digrande KF, Sherwood KD, Colantuoni E, et al. Return to work after critical illness: a systematic review and meta-analysis. Thorax 2020 Jan 08;75(1):17-27 [FREE Full text] [doi: 10.1136/thoraxjnl-2019-213803] [Medline: 31704795]
45. McPeake J, Mikkelsen ME, Quasim T, Hibbert E, Cannon P, Shaw M, et al. Return to employment after critical illness and its association with psychosocial outcomes. A systematic review and meta-analysis. Annals ATS 2019 Oct;16(10):1304-1311. [doi: 10.1513/annalsats.201903-248oc]
46. Hashem MD, Nallagangula A, Nalamalapu S, Nunna K, Nausran U, Robinson KA, et al. Patient outcomes after critical illness: a systematic review of qualitative studies following hospital discharge. Crit Care 2016 Oct 26;20(1):345 [FREE Full text] [doi: 10.1186/s13054-016-1516-x] [Medline: 27782830]
47. Kean S, Salisbury LG, Rattray J, Walsh TS, Huby G, Ramsay P. 'Intensive care unit survivorship' - a constructivist grounded theory of surviving critical illness. J Clin Nurs 2017 Oct 21;26(19-20):3111-3124. [doi: 10.1111/jocn.13659] [Medline: 27875013]
48. Kang J, Jeong YJ. Embracing the new vulnerable self: a grounded theory approach on critical care survivors' post-intensive care syndrome. Intensive Crit Care Nurs 2018 Dec;49:44-50. [doi: 10.1016/j.iccn.2018.08.004] [Medline: 30193868]
49. Jeong YJ, Kang J. Development and validation of a questionnaire to measure post-intensive care syndrome. Intensive Crit Care Nurs 2019 Dec;55:102756. [doi: 10.1016/j.iccn.2019.102756] [Medline: 31522829]
50. Minton C, Batten L, Huntington A. The impact of a prolonged stay in the ICU on patients' fundamental care needs. J Clin Nurs 2018 Jun;27(11-12):2300-2310. [doi: 10.1111/jocn.14184] [Medline: 29149460]
51. Carden GP, Graham JW, McLennan S, Celi LA. Long-term outcome of long stay ICU and HDU patients in a New Zealand hospital. Crit Care Shock 2008 Mar;11(1):26-34 [FREE Full text] [Medline: 23766668]
52. Needham DM, Sepulveda KA, Dinglas VD, Chessare CM, Friedman LA, Bingham CO, et al. Core outcome measures for clinical research in acute respiratory failure survivors. An international modified Delphi consensus study. Am J Respir Crit Care Med 2017 Nov;196(9):1122-1130. [doi: 10.1164/rccm.201702-0372oc]
53. Spies CD, Krampe H, Paul N, Denke C, Kiselev J, Piper SK, et al. Instruments to measure outcomes of post-intensive care syndrome in outpatient care settings - Results of an expert consensus and feasibility field test. J Intensive Care Soc 2021 May;22(2):159-174 [FREE Full text] [doi: 10.1177/1751143720923597] [Medline: 34025756]
54. Charmaz K. Premises, principles, and practices in qualitative research: revisiting the foundations. Qual Health Res 2004 Sep 01;14(7):976-993. [doi: 10.1177/1049732304266795] [Medline: 15296667]
55. Vilma Z. Implementing ethical principles in social research: challenges, possibilities and limitations. Vocational Training Res Realities 2018;29(1):19-43. [doi: 10.2478/vtrr-2018-0003]
56. Guetterman TC, Babchuk WA, Howell Smith MC, Stevens J. Contemporary approaches to mixed methods–grounded theory research: a field-based analysis. J Mixed Methods Res 2017 Jun 01;13(2):179-195. [doi: 10.1177/1558689817710877]
57. Detsky ME, Harhay MO, Bayard DF, Delman AM, Buehler AE, Kent SA, et al. Six-month morbidity and mortality among intensive care unit patients receiving life-sustaining therapy. A prospective cohort study. Annals ATS 2017 Oct;14(10):1562-1570. [doi: 10.1513/annalsats.201611-875oc]
58. Turnbull A, Rabiee A, Davis W, Nasser M, Venna VR, Lolitha R, et al. Outcome measurement in ICU survivorship research from 1970 to 2013: a scoping review of 425 publications. Crit Care Med 2016 Jul;44(7):1267-1277 [FREE Full text] [doi: 10.1097/CCM.0000000000001651] [Medline: 26992067]
59. Hodgson C, Watts N, Iwashyna T. Long-term outcomes after critical illness relevant to randomized clinical trials. In: Annual Update in Intensive Care and Emergency Medicine. New York: Springer; 2016.
60. The EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9]
61. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]

62. Weiss D. The impact of event scale: revised. In: Cross-Cultural Assessment of Psychological Trauma and PTSD. Boston: Springer; 2007.
63. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
64. Adhikari NK, Fowler RA, Bhagwanjee S, Rubenfeld GD. Critical care and the global burden of critical illness in adults. Lancet 2010 Oct;376(9749):1339-1346. [doi: 10.1016/s0140-6736(10)60446-1]
65. Garland A, Olafson K, Ramsey CD, Yogendran M, Fransoo R. A population-based observational study of intensive care unit–related outcomes. With emphasis on post-hospital outcomes. Annals ATS 2015 Feb;12(2):202-208. [doi: 10.1513/annalsats.201405-201cme]
66. Oliveira E, Parikh A, Lopez-Ruiz A, Carrilo M, Goldberg J, Cearras M, et al. ICU outcomes and survival in patients with severe COVID-19 in the largest health care system in central Florida. PLoS One 2021 Mar 25;16(3):e0249038 [FREE Full text] [doi: 10.1371/journal.pone.0249038] [Medline: 33765049]
67. Dennis J, McGovern A, Vollmer S, Mateen B. Improving survival of critical care patients with coronavirus disease 2019 in England: a national cohort study, March to June 2020. Crit Care Med 2021 Feb 01;49(2):209-214 [FREE Full text] [doi: 10.1097/CCM.0000000000004747] [Medline: 33105150]
68. Slim MA, Lala HM, Barnes N, Martynoga RA. Māori health outcomes in an intensive care unit in Aotearoa New Zealand. Anaesth Intensive Care 2021 Jun 22;49(4):292-300. [doi: 10.1177/0310057x21989715]
69. Chun Tie Y, Birks M, Francis K. Grounded theory research: a design framework for novice researchers. SAGE Open Med 2019;7:2050312118822927 [FREE Full text] [doi: 10.1177/2050312118822927] [Medline: 30637106]
70. Shulman M, Myles P, Chan M, McIlroy D, Wallace S, Ponsford J. Measurement of disability-free survival after surgery. Anesthesiology 2015 Mar;122(3):524-536 [FREE Full text] [doi: 10.1097/ALN.0000000000000586] [Medline: 25689757]
71. Relander K, Hietanen M, Rantanen K, Rämö J, Vento A, Saastamoinen K, et al. Postoperative cognitive change after cardiac surgery predicts long-term cognitive outcome. Brain Behav 2020 Sep 17;10(9):e01750 [FREE Full text] [doi: 10.1002/brb3.1750] [Medline: 32681544]
72. The HDC Code of Health and Disability Services Consumers' Rights Regulation 1996. Health and Disability Commissioner. URL: https://www.hdc.org.nz/your-rights/about-the-code/code-of-health-and-disability-services-consumers-rights/(full) [accessed 2022-02-21]
73. Rousseau A, Minguet P, Colson C, Kellens I, Chaabane S, Delanaye P, et al. Post-intensive care syndrome after a critical COVID-19: cohort study from a Belgian follow-up clinic. Ann Intensive Care 2021 Jul 29;11(1):118 [FREE Full text] [doi: 10.1186/s13613-021-00910-9] [Medline: 34324073]
74. Nakanishi N, Liu K, Kawakami D, Kawai Y, Morisawa T, Nishida T, et al. Post-intensive care syndrome and its new challenges in coronavirus disease 2019 (COVID-19) pandemic: a review of recent advances and perspectives. J Clin Med 2021 Aug 28;10(17):3870 [FREE Full text] [doi: 10.3390/jcm10173870] [Medline: 34501316]

Abbreviations

COS: core outcome set
GP: general practitioner
ICU: intensive care unit
PerCI: persistently critically ill
PI: principal investigator
PICS: Post Intensive Care Syndrome
PPV: positive pressure ventilation
REDCap: Research Electronic Data Capture
WHODAS: World Health Organization Disability Assessment Scale

Edited by T Derrick; submitted 22.12.21; this is a non–peer-reviewed article;accepted 06.01.22; published 17.03.22.

Please cite as:
Sutton L, Bell E, Every-Palmer S, Weatherall M, Skirrow P
Survivorship of Patients After Long Intensive Care Stay With Exploration and Experience in a New Zealand Cohort (SPLIT ENZ): Protocol for a Mixed Methods Study
JMIR Res Protoc 2022;11(3):e35936
URL: https://www.researchprotocols.org/2022/3/e35936
doi:10.2196/35936
PMID:35297773

Protocol

Delivery of a Mental Health Intervention for Chronic Pain Through an Artificial Intelligence–Enabled App (Wysa): Protocol for a Prospective Pilot Study

Megha Gupta[1], PhD; Tanya Malik[1], MA; Chaitali Sinha[1], MA

Wysa Inc, Boston, MA, United States

Corresponding Author:
Tanya Malik, MA
Wysa Inc
131 Dartmouth St, 3rd floor
Boston, MA, 02116
United States
Phone: 1 987 480 3442
Email: tanya@wysa.io

Abstract

Background: Patients with chronic pain often experience coexisting, long-term and debilitating mental health comorbidities such as depression and anxiety. Artificial intelligence–supported cognitive behavioral therapy (AI-CBT) interventions could offer cost-effective, accessible, and potentially effective resources to address this problem. However, there is not enough research conducted about the efficacy of AI-CBT interventions for chronic pain.

Objective: This prospective cohort study aims to examine the efficacy and use of an AI-CBT intervention for chronic pain (Wysa for Chronic Pain app, Wysa Inc) using a conversational agent (with no human intervention). To the best of our knowledge, this is the first such study for chronic pain using a fully-automated, free-text–based conversational agent.

Methods: Participants with self-reported chronic pain (n=500) will be recruited online on a rolling basis from April 2022 through posts on US-based internet communities within this prospective cohort. Informed consent is received from participants within the app, and the Wysa for Chronic Pain intervention is delivered remotely for 8 weeks. Outcome measures including a numeric pain rating scale and Patient-Reported Outcomes Measurement Information System–Pain Interference, Generalized Anxiety Disorder–7, and Patient Health Questionnaire–9 questionnaires administered to test the effectiveness of the intervention on reducing levels of pain interference, depression, and anxiety. The therapeutic alliance created with the conversational agent will be assessed through the Working Alliance Inventory–Short Revised instrument. Retention and use statistics will be observed for adherence and engagement.

Results: The study will open for recruitment in April 2022, and data collection is expected to be completed by August 2022. The results for the primary outcomes are expected to be published by late 2022.

Conclusions: Mental health conversational agents driven by artificial intelligence could be effective in helping patients with chronic pain learn to self-manage their pain and common comorbidities like depression and anxiety. The Wysa for Chronic Pain app is one such digital intervention that can potentially serve as a solution to the problems of affordability and scalability associated with interventions that include a human therapist. This prospective study examines the efficacy of the app as a treatment solution for chronic pain. It aims to inform future practices and digital mental health interventions for individuals with chronic pain.

International Registered Report Identifier (IRRID): PRR1-10.2196/36910

(JMIR Res Protoc 2022;11(3):e36910) doi:10.2196/36910

KEYWORDS

chronic pain; AI-enabled mental health assistant; digital health intervention; mental health conversational agent; artificial intelligence; depression; mental health; anxiety; health care cost; conversational agent; chatbot; digital health

Introduction

Background

Chronic pain is a long-term debilitating health concern that affects physical, psychological, cognitive, and social functioning [1-3] resulting in significant health care costs [4]. According to the American Academy of Pain Medicine, more than 1.5 billion people around the world have chronic pain [5]. At least 116 million US adults—more than the number affected by heart disease, diabetes, and cancer combined—live with common chronic pain conditions [6]. Research also suggests that 13% to 50% of adults in the United Kingdom are affected by chronic pain, with 10% to 14% experiencing moderate-to-severe disabling chronic pain [1,7]. The total health care cost of chronic pain ranges from $560 billion to $635 billion annually in the United States [8]. Chronic pain conditions are comorbid with depression, anxiety, sleep disturbances, fatigue, neurocognitive changes, and other symptoms [9].

Chronic Pain and Mental Health

Mental health difficulties such as depression and anxiety are among the most common comorbidities present with chronic pain [10,11]. Between 20% to 50% of patients with chronic pain have comorbid depression [1,12], and patients with severe pain are more likely to be depressed [13]. Chronic pain often leads to patients being unable to focus on normal day-to-day activities and being constantly distracted by their pain, which is associated with poor mental health [14]. Moreover, patients experiencing from pain-related depression and anxiety are also likely to have worse outcomes from chronic pain [15,16]. Numerous studies have also suggested that individuals living with chronic pain are 2 to 3 times more susceptible to suicidal ideation and suicidal behaviors [17-19].

Similarly, research suggests that anxiety and fear arising from chronic pain are likely to interfere with recovery from the pain [15]. This has been observed with the development of maladaptive behaviors including catastrophizing [20], fear avoidance, and the reduction in helpful activities [21]. Chronic pain shares a bidirectional relationship with mental health [22], and the resolution of depression and anxiety are important components in the effective management of chronic pain.

Pain Management

Chronic pain management through psychotherapeutic techniques is important since a commonly observed impact of chronic pain is the reduction in coherent behavior due to the consistent threat of pain [23]. Individuals report experiencing dissembling, with the pain moving beyond an acute experience and evolving into the prolonged state of a disabling disease [24].

Psychological treatments for chronic pain are usually cognitive or behavioral strategies aimed at reducing mental suffering and promoting active engagement with life [23]. They serve to overcome the pain perseverance paradox, which is the occurrence of self-defeating behaviors by those living with chronic pain [25]. Studies have demonstrated that treating mental health concerns such as depression can often result in better outcomes for treatment of pain [26]. Treatment strategies include behavioral techniques of relaxation, biofeedback, contingency management, exposure, and cognitive behavioral techniques that typically include programs with components of education, coping strategies training, and cognitive therapy [23].

Since depression in patients with chronic pain often goes unrecognized, it therefore remains untreated [27]. There is also a significant gap in the needs of patients with chronic pain and the care provided due to lack of therapeutic resources, acute stress on outpatient facilities [28], and worsening mental health experienced during the ongoing COVID-19 pandemic [29,30]. Several barriers to self-management of pain have been identified that include lack of support and resources, lack of time or fear of worsening the pain by engaging in physical activities, and difficulty in patient-physician interactions [31,32].

Research suggests that factors like support and encouragement from caregivers, access to a variety of self-management techniques, and treatment of depression can serve as facilitators in improving self-pain management [32]. Research also indicates the positive impact of a therapeutic alliance in improving the impact of treatment for individuals with chronic pain [33,34].

Digital Mental Health Interventions for Chronic Pain

Digital mental health interventions can potentially address the issues of accessibility, rising health care costs, low availability of therapeutic care, and other barriers [35,36] associated with standard in-person treatments for chronic pain [37] and mental health [38,39]. Studies have reported that digitally delivered mental health interventions have a positive role to play in the self-management of depression and anxiety [40,41] and have been found effective in reducing the impact of chronic pain on the quality of life [42].

Currently, one way in which these digital mental health interventions address the barriers of accessibility is by augmenting digitally delivered tools with human coaches [43-46]. However, the success of this approach is dependent on the availability and degree of involvement of therapists, which influences the cost and scalability of such interventions. Another approach that can effectively tackle the problems of scalability and cost is an artificial intelligence (AI)-enabled conversational agent that mimics human dialog with users [47]. Health care interventions involving artificial intelligence–supported cognitive behavioral therapy (AI-CBT) conversational agents, aimed at reducing symptoms or improving self-management of mental health conditions and increasing mental well-being, are increasingly being used with positive outcomes [48-50]. Many of these conversational agent–based interventions have shown significant improvements on measures of common mental health problems like depression, psychological distress, anxiety, fear of heights, and loneliness [43,51-56]. Also, these studies suggest a positive correlation between engagement level with the conversational agent and reduction in psychological distress. However, there is a need for more evidence-based studies that focus on the efficacy of the interventions driven by AI-enabled conversational agents [57,58].

Wysa for Chronic Pain

The intervention proposed in this study employs one such AI-CBT intervention, the Wysa for Chronic Pain app (Wysa Inc), that uses an AI-enabled conversational agent with a

free-text conversational interface. It listens and responds to the user's distress by recommending techniques and self-care tools based on CBT, behavioral reinforcement, and mindfulness, among other evidence-based therapies. Wysa supports the user in dealing with multiple challenges such as anxiety, sleep, low energy, motivation, loss, and pain. The app has been shown to be effective through mean improvements in symptoms of major depression (Patient Health Questionnaire–9 item [PHQ-9]) among users who were highly engaged with the app when compared to a low engagement group [59].

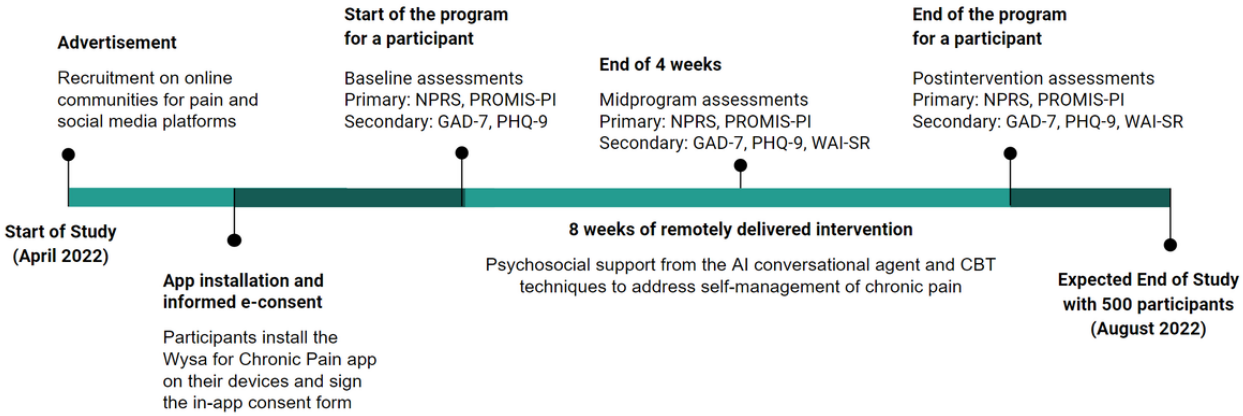
This is an exploratory study following a quantitative research design that aims to study the efficacy of a digital mental health intervention program for chronic pain when delivered solely by a conversational agent. The duration of the study is 8 weeks, and the efficacy will be calculated through the use of standardized statistical measures.

Methods

Overview

In this study, 500 individuals living with chronic pain will be recruited and administered an 8-week program on the Wysa for Chronic Pain app (a specific version of the publicly available Wysa app). This is an anonymous study where the participants enroll themselves by installing the app on their mobile phones and agree to participate in the study. We will restrict this version of the app to the Apple App Store and Google Playstore in the United States. Once enrolled in the program, participants will complete an assessment at the start of the study (comprising 4 measures), complete an 8-week program with the AI-led conversational agent during which another assessment (5 measures) will be taken halfway through the program, and complete the final set of assessments (5 measures) at the end of 8 weeks (Figure 1). Postintervention changes in pain interference will serve as the primary outcome for this study. Our secondary outcomes will include postintervention changes in depression and anxiety, as well as the therapeutic alliance between the user and conversational agent. The app will be free and available 24×7 to the participants.

Figure 1. Timeline of the study. NPRS: numerical pain rating scale; PROMIS-PI: Patient-Reported Outcomes Measurement Information System Pain Inference; GAD-7: Generalized Anxiety Disorder, 7-item; PHQ-9: Patient Health Questionnaire, 9-item; WAI-SR: Working Alliance Inventory–Short Revised; CBT: cognitive behavioral therapy.



Screening

Eligibility criteria include being aged 18 years or older, experiencing chronic pain, and not receiving any form of professional mental health support. Individuals who do not self-report presenting with chronic pain will be excluded as a part of the recruitment criteria. The data of participants who score below 4 points on a 1-item numeric pain rating scale (NPRS) at the first assessment will be excluded from the analysis [60,61].

Recruitment

Participants are currently being recruited on a rolling basis from US-based internet communities centered around the experiences of living with chronic pain. We will share information about the research study by posting on social media platforms. The purpose of these posts is to introduce potential participants to the Wysa for Chronic Pain app, inform them of the aims of this research study, and subsequently invite them to download the app on their devices. All participants will need to have an Android or iOS phone and have access to the internet while using the app. Interested participants will be screened on the basis of self-reported scores on the NPRS and their responses to a set of questions that will collect baseline demographics (age range), type of chronic pain, and the time since onset.

Ethics

Participants will remain anonymous for the duration of the study and are being recruited from a US-only population. After participants download the Wysa for Chronic Pain app and agree to the app's Terms of Service and Privacy Policy, they will opt in to the study using a consent screen following Wysa's security and compliance guidelines. The consent screen states the purpose of the trial, potential risks and benefits associated with participation in the trial study, and mechanism for opting out of the study at any time. They will also be informed of the ways in which the usage data generated from their participation in the study could potentially be used.

Intervention

A separate version of the publicly available Wysa app, Wysa for Chronic Pain, is being used for this study. This is a conversational agent–only version (ie, it does not include one-on-one human coach support, which is an option with the publicly available app). The AI-enabled free-text conversational agent offers participants a space to talk about their pain, depression, anxiety, and other issues arising from disturbances due to chronic pain. The conversational agent acts as a companion in their journey (Figure 2) of learning to live with pain and helps them build resilience by guiding them through self-care tools based on CBT and other techniques. The goal of the intervention is to improve the quality of life for the user as their ability to manage the pain improves.

The conversational agent builds an 8-week road map for the user and does daily morning and evening check-ins to encourage adherence to the program (Figure 3). Users are encouraged to set a daily goal for themselves based on what gives them joy, and ideas are suggested to help them engage in these activities, even if in a small way. Completing these activities on a regular basis gives the users a sense of joy and achievement and fosters the belief that they are capable of simultaneously managing their pain and living a normal life [62,63].

Steps included in the user's journey includes participating in activities that comprise a mix of positive psychology, acceptance and commitment therapy, and CBT techniques that will be conducted in an organized, weekly fashion to encourage the development of skills aimed at coping with pain and the reduction of depression and anxiety symptoms.

In contrast to traditional worksheets or psychoeducation materials common in digital mental health [64,65], all therapeutic interventions in this study are conducted in a conversational format via the Wysa AI-enabled agent, which guides the user through the program in order to create a sense of support similar to in-person therapy [66]. The conversational nature of the interaction is meant to engage the user more, thereby increasing adherence to the effective intervention for chronic pain. The AI-enabled conversational agent provides empathetic support to the user, answers their questions, and guides them through uncertainty and a potential lack of trust in the bot as they proceed with their journey [67]. A rapport is built between the user and Wysa by the conversational agent's ability to add personalized elements to the conversation, including welcoming messages, addressing the user by the nickname chosen by them during onboarding, and providing clarifying examples. All these are important elements of a therapeutic conversation [68].

The app provides the user an open space for conversation by giving the user the ability to converse with the conversational agent through free text (Figure 4). The free-text input enables flexibility and genuineness in the conversation, and allows the user to speak their mind. The app uses natural language processing, natural language understanding, and machine learning to interpret the user's messages, and provides the user with supportive listening and appropriate responses to unique situations using interventive techniques, thus creating an efficient interactive environment [55,69]. The conversational agent provides users with empathetic support in a nonjudgmental environment, an important aspect of health care [70]. The AI uses custom models trained by data scientists, clinicians, and conversation designers that detect different aspects of what the user is conveying [71].

Figure 2. Onboarding conversation with the conversational agent that sets the expectation for the program and administers the first assessment.

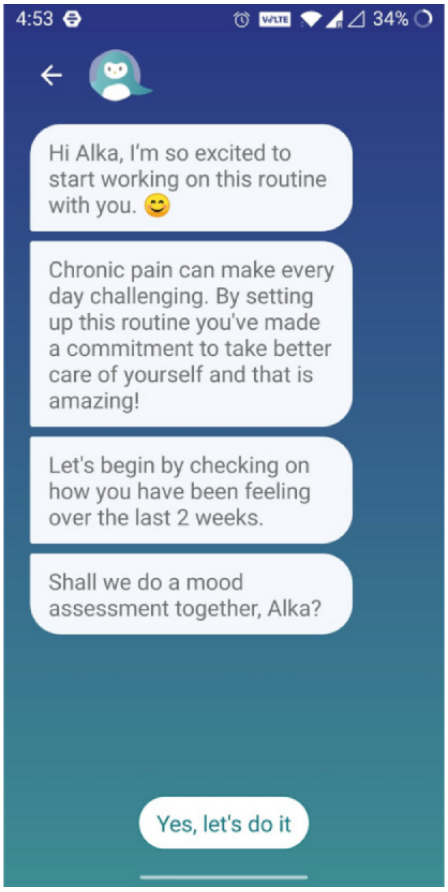


Figure 3. First mood check-in with the conversational agent.

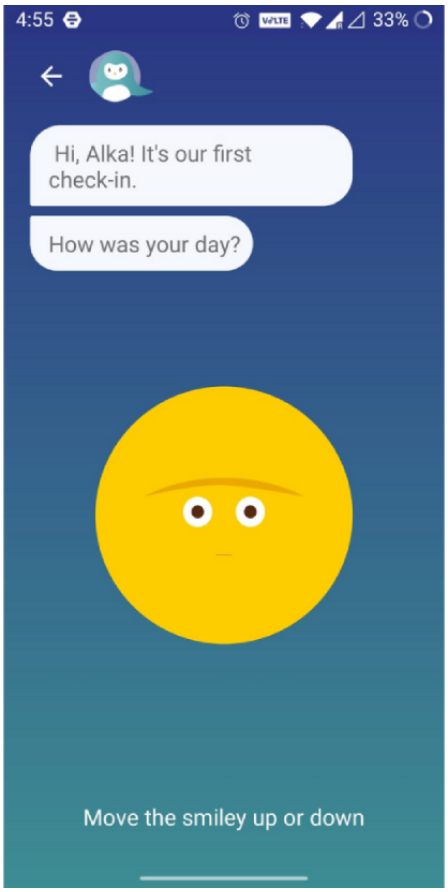
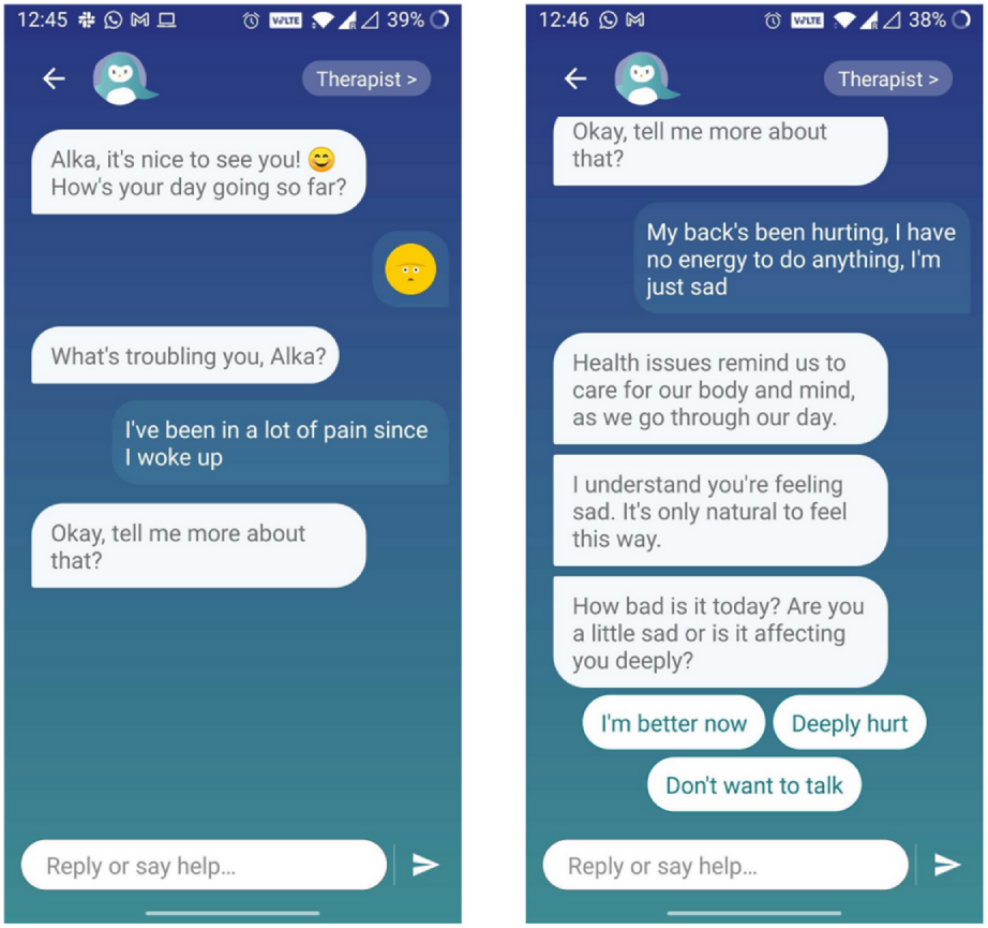


Figure 4. Example screenshots showing the conversational style of the app and the flexibility given to the user through free-text inputs.



Primary Outcome Measure

Study participants will complete the self-administered measures, NPRS and the Patient-Reported Outcomes Measurement Information System–Pain Inference (PROMIS-PI) short form 6b, at the baseline, midpoint, and after the completion of the study. Both measures will be administered through the Wysa app and will be completed on the participants' own devices.

Numeric Pain Rating Scale

The NPRS is a unidimensional measure of pain intensity in adults with chronic pain [72]. It is a segmented 11-point numeric scale in which the respondent rates the intensity of their pain on an integer scale of 0 (no pain) to 10 (worst pain imaginable). The minimal clinically important change score is taken to be a 2 point reduction [73,74].

Patient-Reported Outcomes Measurement Information System–Pain Interference

The 6-item PROMIS-PI was developed by the National Institutes of Health to measure the negative effects of pain on physical, mental, cognitive, emotional, social, and recreational functioning. This study uses a 6-item pain interference scale that aims to measure the extent to which pain causes a hindrance in various aspects of an individual's life, including those related to activities done for fun as well as day-to-day activities. A 5-point ordinal rating scale is used with options being 1 (not at all), 2 (a little bit), 3 (somewhat), 4 (quite a bit), and 5 (very much). The raw scores are converted to T-scores based on a scoring manual. Although the PROMIS-PI forms are relatively new, their responsiveness has been found to be comparable to legacy pain measures [75]. A reduction in PROMIS-PI scores at the end of the study will indicate a positive impact of the intervention. Minimal clinically important change score is defined as at least 2.0 points for pain interference [76].

Secondary Outcome Measures

In order to observe ancillary impacts of the intervention, study participants will complete 2 additional self-administered measures (PHQ-9 and Generalized Anxiety Disorder, 7-item [GAD-7]) at the baseline, midpoint, and after the completion of the study. The Working Alliance Inventory–Short Revised (WAI-SR) will be administered at the midpoint and after completion of the study only. All measures will be administered through the Wysa app itself and will be completed on the participants' own devices.

Generalized Anxiety Disorder

GAD-7 is a widely adopted, standardized, and validated 7-item self-report questionnaire developed to assess symptoms of generalized anxiety disorder. This 7-item anxiety scale has been found to have good reliability as well as construct, procedural, and factorial validity [77]. This assessment asks the participants to assess themselves over the last 2 weeks with questions about whether they felt nervous or anxious, had trouble relaxing, became easily annoyed or irritated, etc. The response categories include 0 (not at all), 1 (several days), 2 (more than half the days), and 3 (nearly every day). The final score is calculated by

summing over the 7 questions, and scores of 5, 10, and 15 are taken as the cutoff points for mild, moderate, and severe anxiety, respectively. A reduction in the GAD-7 score at the end of the study will indicate a positive impact of the intervention. Clinically meaningful effect size is defined as at least 4.0 points for the GAD-7 [78].

Patient Health Questionnaire

PHQ-9 is a widely used, standardized, validated 9-item survey assessment tool for depression that has been shown to be valid and reliable as a longitudinal clinical tool [79]. It has demonstrated high criterion validity among populations with high rates of physical distress and has been found to be valid for diverse modalities of administration, including those via touchscreens [80]. This assessment asks the participants to assess themselves over the last 2 weeks with questions about whether they felt pleasure in doing things, felt tired or having little energy, had thoughts of hurting themselves, etc. The response categories and associated numerical scores are identical to that of GAD-7. The overall score is calculated by summing over the 9 questions and is used to monitor the severity of depression and response to treatment. A reduction in the PHQ-9 scores at the end of the study will indicate a positive impact of the intervention. Clinically meaningful effect size is defined as at least 5.0 points for the PHQ-9 [79].

Working Alliance Inventory–Short Revised

The WAI-SR is a well-established measure of therapeutic alliance, consisting of a total score and 3 subscales: bond, goal, and task [81]. This questionnaire asks the participants 12 questions related to whether the therapy gave them new ways of looking at their problem, whether their therapist cares for and respects them, and whether they agree with their therapist on the goals set for them. The response categories lie on a 5-point Likert scale and include 1 (seldom), 2 (sometime), 3 (often), 4 (very often), and 5 (always). Ratings are summed at the end, with a higher total indicating better therapeutic alliance. The WAI-SR will be administered at the end of the study via the app's conversational interface, in which the word therapist will be changed to Wysa. This measure demonstrates high internal consistency (Cronbach α>.90) [82]. Mean total scores of 3.59 (out of 5) are considered high [83].

Statistical Analysis

We will use the Wilcoxon signed-rank test to measure the efficacy of the intervention by comparing baseline and postintervention assessment scores on the NPRS, GAD-7, and PHQ-9 scales. The Wilcoxon signed-rank test is a nonparametric statistical hypothesis test used to compare the differences between 2 populations using a set of matched samples. We will also use the Wilcoxon signed-rank test for measuring the median of therapeutic alliance using the WAI-SR scale. We chose this test because we believe that our data will not be normally distributed, in which case a nonparametric test is more suitable. Finally, we will do a paired *t* test for measuring if the intervention resulted in any significant changes on the PROMIS-PI scale since its scores are mapped to T-scores. Rolling and ongoing recruitment will mediate any potential concerns regarding the power of the study and the multiple testing problem where Type I errors may get inflated.

Results

Recruitment for this study will begin in April 2022 and continue on a rolling basis until 500 participants have been recruited. The baseline, midpoint, and postintervention assessment data for participants who have completed their 8-week intervention will be evaluated to determine if the intervention had any significant effect on their pain interference, depression, and anxiety scales. Although participants are encouraged to check in twice a day, there is no mandate for continuation in the program. Final outcomes will be based on end of program assessments, and engagement is evaluated across the cohort to evaluate efficacy in a real-world use setting. Another round of recruitment may be needed if the retention rate is low. App use, engagement, and retention will also be assessed to determine the acceptability of an AI-only conversational agent to support chronic pain management. From this pilot, we hope to learn more about what frequency and how many check-ins can be considered as sufficient use of the app so as to have desirable outcomes for the user. Data collection for all participants is expected to be completed by August 2022 and results to be published by late 2022.

Discussion

Summary

To the best of our knowledge, this is the first study to assess the efficacy of a fully automated, free-text–based conversational agent as an assistant for chronic pain patients. Prior to this, Hauser-Ulrich et al [84] evaluated the efficacy and acceptance of the Smartphone-Based Health Care Chatbot to Promote Self-Management of Chronic Pain (SELMA) in pain self-management for chronic pain patients. Some of the shortcomings of SELMA, as pointed out by its users, were the conversation being too structured, there not being enough answer options, and no ability to write free text. In addition, the notifications were sent at a fixed time and there was no way to personalize the app.

Wysa for Chronic Pain overcomes these shortcomings. Moreover, apart from using a conversational flow tailored for chronic pain, it also provides participants with a wide array of self-care tools they can use to deal with other issues like insomnia, depression, anxiety, and negative thoughts anytime they want [59,85,86]. Wysa for Chronic Pain has proven high engagement and efficacy when the intervention uses a conversational agent enhanced by a human coach [87]. This study aims to assess the efficacy of this intervention using the conversational agent alone. The quality of the therapeutic alliance built between the user and the conversational agent will also be assessed, which will further inform the utility of digital mental health interventions.

Treatment of chronic pain often requires a multidisciplinary and multidimensional approach, targeting both the physical and psychological aspects of the disease. Many techniques like behavioral activation, mindfulness exercises, CBT, interpersonal

psychotherapy, and psychoeducational tools have been tried for different kinds of pain, with each showing promise for different aspects of pain management [88]. This is one of the reasons why digital health interventions are becoming increasingly popular as they remove the barriers of experienced disrespect, distrust, and dismissal encountered while seeking care for chronic pain [89]. There have been several efforts at building digital health interventions for chronic pain and studying their efficacy for various kinds of patient groups. While many have shown promising results for different aspects of pain management, most of them involved human therapists [45,46], focused on passive consumption of psychoeducational content [45,90], or required an additional device [91,92]. Any requirement of a therapist or an additional device greatly limits the scalability and accessibility of a digital mental health intervention, and passive consumption of psychoeducational content isn't interactive, and hence, may not be very engaging [57].

To address these challenges, this study proposes to investigate the efficacy of a conversational agent–led intervention for chronic pain without the need for human coaches or special devices. The Wysa for Chronic Pain app involves no human coach in the loop, but the free-text–based AI-enabled conversational agent is capable of understanding and responding to a user's messages like a human therapist would. The advantages of using a digital mental health conversational agent like Wysa are manifold. First, therapy and self-care tools grounded in CBT, mindfulness, and other evidence-based therapies become immediately accessible to the patient via their mobile phone. Second, the strain on human resources is reduced. Third, anonymity allows patients to share their thoughts and feelings more freely, thereby increasing alliance. Fourth, especially in the context of chronic pain, when even routine activities seem overwhelming, a mobile assistant puts the resources literally in the user's hands and takes away the effort, stress, and additional challenges required to go meet a therapist. Finally, the conversational agent also acts like a companion who is always there to talk to [93].

Another facet of the treatment for chronic pain includes the therapeutic alliance between the individual and the support system. Emerging evidence for chronic musculoskeletal pain indicates that a strong therapeutic alliance may improve pain outcomes [94]. Therapeutic alliance was consistently a predictor of outcome in a study with 182 patients with lower back pain that assessed function, global perceived effect of treatment, pain, and disability [95]. Another study showed that enhanced therapeutic alliance combined with treatment led to clinically meaningful improvements in pain intensity and muscle pain sensitivity in patients with chronic low back pain [96]. This study looks at the levels of therapeutic alliance formed with the AI-enabled conversational agent and the impact on treatment outcomes for the participants.

Strengths and Limitations

A challenge associated with most digital mental health interventions is ensuring adherence and engagement [90]. Adherence is key to benefiting from a self-management program. To address this potential problem and promote adherence, the app has been built using the principles of behavioral activation, which are supported conversationally and have specific customizations for chronic pain [87].

To encourage participation, the study has intentionally been structured to be anonymous. The app does not require the creation of an account nor does it ask for personal identification data (name, age, location, etc). Any sensitive information, if provided accidentally by a participant in a conversation, is identified and redacted by an algorithm to prevent retention in the system. This kind of anonymity may promote trust in the app among the users and encourage them to feel more open and engage with the conversational agent. It is only available in English, which is a limitation of the service.

Although recruitment for this study is being done from online communities of people living with chronic pain, one challenge that remains is the inability to verify chronic pain in the cohort recruited. All participants will be self-referred, thus making it difficult to assess whether they satisfy the inclusion criteria. In addition, their presence in online communities oriented to the support of pain may indicate sampling bias toward individuals with a greater willingness to learn about pain self-management [97]. Another potential limitation of this study may be the self-reported nature of the assessments, which may indicate subjective efficaciousness due to the absence of physician-interventive metrics for change. Self-reporting also involves the risk of random answers offered by some participants. We will examine the data for outliers (any data points that lie 2 to 3 standard deviations away from the mean) and use an appropriate statistical methodology to improve data quality.

Conclusions

Chronic pain is one of the leading causes of long-term disability in the world. It is a complex problem with roots in social, economic, physical, and psychological aspects and affects more than 1.5 billion people worldwide. Treatment is difficult, and it often requires continuous pain management. Depression and anxiety are the most common comorbidities occurring with chronic pain, and psychological interventions have been shown to result in better outcomes for treatment of chronic pain. Digital mental health solutions make these interventions accessible and affordable. This study describes a novel AI-CBT intervention, the Wysa for Chronic Pain app, which is completely led by a free-text–based, AI-enabled conversational agent. The intervention takes the user through a variety of techniques and self-care tools grounded in CBT and mindfulness, which enable the user to learn how to self-manage their pain on a regular basis. The results from this study will be important in understanding the efficacy of such an intervention that can potentially serve as a scalable and cost-effective resource for chronic pain patients around the globe.

Conflicts of Interest

CS and TM report being employed by and owning equity in Wysa Inc. MG reports being employed by Wysa Inc.

References

1. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006 May;10(4):287-333 [FREE Full text] [doi: 10.1016/j.ejpain.2005.06.009] [Medline: 16095934]
2. Nicholson B, Verma S. Comorbidities in chronic neuropathic pain. Pain Med 2004 Mar 01;5 Suppl 1(suppl 1):S9-S27 [FREE Full text] [doi: 10.1111/j.1526-4637.2004.04019.x] [Medline: 14996227]
3. Demyttenaere K, Bruffaerts R, Lee S, Posada-Villa J, Kovess V, Angermeyer M, et al. Mental disorders among persons with chronic back or neck pain: results from the World Mental Health Surveys. Pain 2007 Jun;129(3):332-342 [FREE Full text] [doi: 10.1016/j.pain.2007.01.022] [Medline: 17350169]
4. Gaskin DJ, Richard P. The economic costs of pain in the United States. J Pain 2012 Aug;13(8):715-724. [doi: 10.1016/j.jpain.2012.03.009] [Medline: 22607834]
5. Gardner J, Sachdeva H. Causes of pain worldwide. In: Abd-Elsayed A, editor. Pain: A Review Guide. Cham: Springer International Publishing; 2019:1191-1192.
6. Tsang A, Von Korff M, Lee S, Alonso J, Karam E, Angermeyer MC, et al. Common chronic pain conditions in developed and developing countries: gender and age differences and comorbidity with depression-anxiety disorders. J Pain 2008 Oct;9(10):883-891. [doi: 10.1016/j.jpain.2008.05.005] [Medline: 18602869]
7. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open 2016 Jun 20;6(6):e010364 [FREE Full text] [doi: 10.1136/bmjopen-2015-010364] [Medline: 27324708]
8. Simon LS. Relieving pain in America: a blueprint for transforming prevention, care, education, and research. J Pain Pall Care Pharmacother 2012 Jul 05;26(2):197-198 [FREE Full text] [doi: 10.3109/15360288.2012.678473]
9. Dahan A, van Velzen M, Niesters M. Comorbidities and the complexities of chronic pain. Anesthesiology 2014 Oct;121(4):675-677 [FREE Full text] [doi: 10.1097/ALN.0000000000000402] [Medline: 25099749]
10. Fishbain DA. Approaches to treatment decisions for psychiatric comorbidity in the management of the chronic pain patient. Med Clin North Am 1999 May;83(3):737-760 [FREE Full text] [doi: 10.1016/s0025-7125(05)70132-2]
11. Romano JM, Turner JA. Chronic pain and depression: does the evidence support a relationship? Psychol Bull 1985;97(1):18-34 [FREE Full text] [doi: 10.1037/0033-2909.97.1.18]
12. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012 Jul;380(9836):37-43. [doi: 10.1016/S0140-6736(12)60240-2]
13. de Heer EW, Ten Have M, van Marwijk HW, Dekker J, de Graaf R, Beekman AT, et al. Pain as a risk factor for common mental disorders. Results from the Netherlands Mental Health Survey and Incidence Study-2: a longitudinal, population-based study. Pain 2018 Apr 15;159(4):712-718 [FREE Full text] [doi: 10.1097/j.pain.0000000000001133] [Medline: 29252911]
14. Eccleston C. Chronic pain and distraction: an experimental investigation into the role of sustained and shifting attention in the processing of chronic persistent pain. Behav Res Ther 1995 May;33(4):391-405. [doi: 10.1016/0005-7967(94)00057-q]
15. Boersma K, Linton S. Expectancy, fear and pain in the prediction of chronic pain and disability: a prospective analysis. Eur J Pain 2006 Aug;10(6):551-557 [FREE Full text] [doi: 10.1016/j.ejpain.2005.08.004] [Medline: 16199189]
16. Nijrolder I, van der Windt D, van der Horst H. Prediction of outcome in patients presenting with fatigue in primary care. Br J Gen Pract 2009 Apr 01;59(561):e101-e109 [FREE Full text] [doi: 10.3399/bjgp09X420329] [Medline: 19341545]
17. Tang NKY, Crane C. Suicidality in chronic pain: a review of the prevalence, risk factors and psychological links. Psychol Med 2006 May;36(5):575-586. [doi: 10.1017/S0033291705006859] [Medline: 16420727]
18. Racine M. Chronic pain and suicide risk: a comprehensive review. Prog Neuropsychopharmacol Biol Psych 2018 Dec 20;87(Pt B):269-280 [FREE Full text] [doi: 10.1016/j.pnpbp.2017.08.020] [Medline: 28847525]
19. Androulakis XM, Guo S, Zhang J, Sico J, Warren P, Giakas A, et al. Suicide attempts in US veterans with chronic headache disorders: a 10-year retrospective cohort study. J Pain Res 2021 Aug;14:2629-2639 [FREE Full text] [doi: 10.2147/JPR.S322432] [Medline: 34466030]
20. Peres MFP, Lucchetti G. Coping strategies in chronic pain. Curr Pain Headache Rep 2010 Oct 3;14(5):331-338 [FREE Full text] [doi: 10.1007/s11916-010-0137-3] [Medline: 20680705]
21. Wakaizumi K, Yamada K, Oka H, Kosugi S, Morisaki H, Shibata M, et al. Fear-avoidance beliefs are independently associated with the prevalence of chronic pain in Japanese workers. J Anesth 2017 Apr 3;31(2):255-262 [FREE Full text] [doi: 10.1007/s00540-016-2303-1] [Medline: 28050703]
22. Ang DC, Bair MJ, Damush TM, Wu J, Tu W, Kroenke K. Predictors of pain outcomes in patients with chronic musculoskeletal pain co-morbid with depression: results from a randomized controlled trial. Pain Med 2010 Apr;11(4):482-491. [doi: 10.1111/j.1526-4637.2009.00759.x] [Medline: 20002592]

23. Eccleston C, Morley S, Williams A. Psychological approaches to chronic pain management: evidence and challenges. Br J Anaesth 2013 Jul;111(1):59-63 [FREE Full text] [doi: 10.1093/bja/aet207] [Medline: 23794646]
24. Chapman CR, Gavrin J. Suffering: the contributions of persistent pain. Lancet 1999 Jun;353(9171):2233-2237 [FREE Full text] [doi: 10.1016/S0140-6736(99)01308-2]
25. Eccleston C. A normal psychology of everyday pain. Int J Clin Pract Suppl 2013 Jan(178):47-50 [FREE Full text] [doi: 10.1111/ijcp.12051] [Medline: 23163549]
26. Mossey JM, Gallagher RM. The longitudinal occurrence and impact of comorbid chronic pain and chronic depression over two years in continuing care retirement community residents. Pain Med 2004 Dec 01;5(4):335-348 [FREE Full text] [doi: 10.1111/j.1526-4637.2004.04041.x] [Medline: 15563319]
27. Lee H, Choi EJ, Nahm FS, Yoon IY, Lee PB. Prevalence of unrecognized depression in patients with chronic pain without a history of psychiatric diseases. Korean J Pain 2018 Apr 30;31(2):116-124 [FREE Full text] [doi: 10.3344/kjp.2018.31.2.116] [Medline: 29686810]
28. Eccleston C, Blyth FM, Dear BF, Fisher EA, Keefe FJ, Lynch ME, et al. Managing patients with chronic pain during the COVID-19 outbreak: considerations for the rapid introduction of remotely supported (eHealth) pain management services. Pain 2020 May;161(5):889-893 [FREE Full text] [doi: 10.1097/j.pain.0000000000001885] [Medline: 32251203]
29. Xiong J, Lipsitz O, Nasri F, Lui LM, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. J Affect Disord 2020 Dec 01;277:55-64 [FREE Full text] [doi: 10.1016/j.jad.2020.08.001] [Medline: 32799105]
30. Vindegaard N, Benros ME. COVID-19 pandemic and mental health consequences: systematic review of the current evidence. Brain Behav Immun 2020 Oct;89:531-542 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.048] [Medline: 32485289]
31. Al-Mahrezi A. Towards effective pain management: breaking the barriers. Oman Med J 2017 Sep;32(5):357-358 [FREE Full text] [doi: 10.5001/omj.2017.69] [Medline: 29026465]
32. Bair MJ, Matthias MS, Nyland KA, Huffman MA, Stubbs DL, Kroenke K, et al. Barriers and facilitators to chronic pain self-management: a qualitative study of primary care patients with comorbid musculoskeletal pain and depression. Pain Med 2009 Oct 01;10(7):1280-1290 [FREE Full text] [doi: 10.1111/j.1526-4637.2009.00707.x] [Medline: 19818038]
33. Hall A, Ferreira P, Maher C, Latimer J, Ferreira M. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. Phys Ther 2010 Aug;90(8):1099-1110 [FREE Full text] [doi: 10.2522/ptj.20090245] [Medline: 20576715]
34. Schönberger M, Humle F, Zeeman P, Teasdale TW. Working alliance and patient compliance in brain injury rehabilitation and their relation to psychosocial outcome. Neuropsychol Rehabil 2006 Jun;16(3):298-314 [FREE Full text] [doi: 10.1080/09602010500176476] [Medline: 16835153]
35. Rosser BA, Vowles KE, Keogh E, Eccleston C, Mountain GA. Technologically-assisted behaviour change: a systematic review of studies of novel technologies for the management of chronic illness. J Telemed Telecare 2009 Oct 08;15(7):327-338 [FREE Full text] [doi: 10.1258/jtt.2009.090116] [Medline: 19815901]
36. Heapy AA, Higgins DM, Cervone D, Wandner L, Fenton BT, Kerns RD. A systematic review of technology-assisted self-management interventions for chronic pain. Clin J Pain 2015;31(6):470-492. [doi: 10.1097/ajp.0000000000000185]
37. Gardiner P, D'Amico S, Luo M, Haas N. An innovative electronic health toolkit (our whole lives for chronic pain) to reduce chronic pain in patients with health disparities: open clinical trial. JMIR Mhealth Uhealth 2020 Mar 30;8(3):e14768 [FREE Full text] [doi: 10.2196/14768] [Medline: 32224487]
38. Kretzschmar K, Tyroll H, Pavarini G, Manzini A, Singh I, NeurOx Young People's Advisory Group. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. Biomed Inform Insights 2019 Mar 05;11:1178222619829083 [FREE Full text] [doi: 10.1177/1178222619829083] [Medline: 30858710]
39. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Dec 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
40. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
41. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct 09;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
42. Palermo TM, Law EF, Fales J, Bromberg MH, Jessen-Fiddick T, Tai G. Internet-delivered cognitive-behavioral treatment for adolescents with chronic pain and their parents: a randomized controlled multicenter trial. Pain 2016 Jan;157(1):174-185 [FREE Full text] [doi: 10.1097/j.pain.0000000000000348] [Medline: 26335910]
43. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 06;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]

44. Ly KH, Ly A, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. Internet Interv 2017 Dec;10:39-46 [FREE Full text] [doi: 10.1016/j.invent.2017.10.002] [Medline: 30135751]
45. Gentili C, Zetterqvist V, Rickardsson J, Holmström L, Simons LE, Wicksell RK. ACTsmart: development and feasibility of digital Acceptance and Commitment Therapy for adults with chronic pain. NPJ Digit Med 2020 Feb 13;3(1):20 [FREE Full text] [doi: 10.1038/s41746-020-0228-4] [Medline: 32128450]
46. Bailey JF, Agarwal V, Zheng P, Smuck M, Fredericson M, Kennedy DJ, et al. Digital care for chronic musculoskeletal pain: 10,000 participant longitudinal cohort study. J Med Internet Res 2020 May 11;22(5):e18250 [FREE Full text] [doi: 10.2196/18250] [Medline: 32208358]
47. Reddy S, Fox J, Purohit MP. Artificial intelligence-enabled healthcare delivery. J R Soc Med 2019 Jan;112(1):22-28. [doi: 10.1177/0141076818815510] [Medline: 30507284]
48. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. J Med Internet Res 2019 May 09;21(5):e13216 [FREE Full text] [doi: 10.2196/13216] [Medline: 31094356]
49. Boucher EM, Harake NR, Ward HE, Stoeckl SE, Vargas J, Minkel J, et al. Artificially intelligent chatbots in digital mental health interventions: a review. Expert Rev Med Devices 2021 Dec 31;18(sup1):37-49 [FREE Full text] [doi: 10.1080/17434440.2021.2013200] [Medline: 34872429]
50. Ramesh A, Kambhampati C, Monson J, Drew P. Artificial intelligence in medicine. Ann R Coll Surg Engl 2004 Sep 01;86(5):334-338 [FREE Full text] [doi: 10.1308/147870804290] [Medline: 15333167]
51. Pinto MD, Greenblatt AM, Hickman RL, Rice HM, Thomas TL, Clochesy JM. Assessing the critical parameters of eSMART-MH: a promising avatar-based digital therapeutic intervention to reduce depressive symptoms. Perspect Psychiatr Care 2016 Jul 19;52(3):157-168. [doi: 10.1111/ppc.12112] [Medline: 25800698]
52. Burton C, Szentagotai Tatar A, McKinstry B, Matheson C, Matu S, Moldovan R, et al. Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression. J Telemed Telecare 2016 Jul 09;22(6):348-355. [doi: 10.1177/1357633x15609793]
53. Ring L, Shi L, Totzke K, Bickmore T. Social support agents for older adults: longitudinal affective computing in the home. J Multimodal User Interfaces 2014 Jun 18;9(1):79-88. [doi: 10.1007/s12193-014-0157-0]
54. Pinto MD, Hickman RL, Clochesy J, Buchner M. Avatar-based depression self-management technology: promising approach to improve depressive symptoms among young adults. Appl Nurs Res 2013 Feb;26(1):45-48 [FREE Full text] [doi: 10.1016/j.apnr.2012.08.003] [Medline: 23265918]
55. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence (TESS) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR Ment Health 2018 Dec 13;5(4):e64 [FREE Full text] [doi: 10.2196/mental.9782] [Medline: 30545815]
56. Freeman D, Haselton P, Freeman J, Spanlang B, Kishore S, Albery E, et al. Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. Lancet Psychiatry 2018 Aug;5(8):625-632. [doi: 10.1016/s2215-0366(18)30226-8]
57. Gaffney H, Mansell W, Tai S. Conversational agents in the treatment of mental health problems: mixed-method systematic review. JMIR Ment Health 2019 Oct 18;6(10):e14166 [FREE Full text] [doi: 10.2196/14166] [Medline: 31628789]
58. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. J Med Internet Res 2017 Jul 21;19(8):e267 [FREE Full text] [doi: 10.2196/jmir.7023] [Medline: 28784594]
59. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: 10.2196/12106] [Medline: 30470676]
60. Krebs EE, Carey TS, Weinberger M. Accuracy of the pain numeric rating scale as a screening test in primary care. J Gen Intern Med 2007 Oct 1;22(10):1453-1458 [FREE Full text] [doi: 10.1007/s11606-007-0321-2] [Medline: 17668269]
61. Fejer R, Jordan A, Hartvigsen J. Categorising the severity of neck pain: establishment of cut-points for use in clinical and epidemiological research. Pain 2005 Dec 15;119(1-3):176-182 [FREE Full text] [doi: 10.1016/j.pain.2005.09.033] [Medline: 16298059]
62. Müller R, Gertz KJ, Molton IR, Terrill AL, Bombardier CH, Ehde DM, et al. Effects of a tailored positive psychology intervention on well-being and pain in individuals with chronic pain and a physical disability. Clin J Pain 2016;32(1):32-44. [doi: 10.1097/Ajp.0000000000000225]
63. Peters M, Smeets E, Feijge M, van Breukelen G, Andersson G, Buhrman M, et al. Happy despite pain: a randomized controlled trial of an 8-week internet-delivered positive psychology intervention for enhancing well-being in patients with chronic pain. Clin J Pain 2017 Nov;33(11):962-975 [FREE Full text] [doi: 10.1097/AJP.0000000000000494] [Medline: 28379873]
64. Baumel A, Faber K. Evaluating triple P online: a digital parent training program for child behavior problems. Cogn Behav Pract 2018 Nov;25(4):538-543 [FREE Full text] [doi: 10.1016/j.cbpra.2017.10.001]

65. Wadon ME, Winter M, Peall KJ. Internet-based cognitive behavioural therapy programme as an intervention for people diagnosed with adult-onset, focal, isolated, idiopathic cervical dystonia: a feasibility study protocol. Pilot Feasibility Stud 2020 Jul 15;6(1):100 [FREE Full text] [doi: 10.1186/s40814-020-00641-x] [Medline: 32685184]
66. Ryan RM, Deci EL. A self-determination theory approach to psychotherapy: the motivational basis for effective change. Can Psychol 2008 Aug;49(3):186-193. [doi: 10.1037/a0012753]
67. Feller C, Cottone R. The importance of empathy in the therapeutic alliance. J Humanistic Couns Educ Devel 2012;42(1):53-61 [FREE Full text] [doi: 10.1002/j.2164-490X.2003.tb00168.x]
68. Althoff T, Clark K, Leskovec J. Large-scale analysis of counseling conversations: an application of natural language processing to mental health. Transact Assoc Computational Ling 2016 Dec;4:463-476. [doi: 10.1162/tacl_a_00111]
69. Xu B, Zhuang Z. Survey on psychotherapy chatbots. Concurrency Comput 2020 Dec 28;34(7):1 [FREE Full text] [doi: 10.1002/cpe.6170]
70. Shields CG, Finley MA, Elias CM, Coker CJ, Griggs JJ, Fiscella K, et al. Pain assessment: the roles of physician certainty and curiosity. Health Commun 2013 Oct;28(7):740-746 [FREE Full text] [doi: 10.1080/10410236.2012.715380] [Medline: 23356451]
71. Kuchlous S, Kadaba M. Short text intent classification for conversational agents. 2020 IEEE 17th India Council Int Conf 2020:1-4 [FREE Full text] [doi: 10.1109/INDICON49873.2020.9342516]
72. Jensen M, McFarland C. Increasing the reliability and validity of pain intensity measurement in chronic pain patients. Pain 1993;55:195-203 [FREE Full text] [doi: 10.1016/0304-3959(93)90148-I]
73. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94(2):149-158. [doi: 10.1016/S0304-3959(01)00349-9]
74. Michener L, Snyder A, Leggin B. Responsiveness of the numeric pain rating scale in patients with shoulder pain and the effect of surgical status. J Sport Rehabil 2011 Feb;20(1):115-128 [FREE Full text] [doi: 10.1123/jsr.20.1.115] [Medline: 21411827]
75. Chen CX, Kroenke K, Stump T, Kean J, Krebs EE, Bair MJ, et al. Comparative responsiveness of the PROMIS pain interference short forms with legacy pain measures: results from three randomized clinical trials. J Pain 2019 Jun;20(6):664-675 [FREE Full text] [doi: 10.1016/j.jpain.2018.11.010] [Medline: 30529442]
76. Chen CX, Kroenke K, Stump TE, Kean J, Carpenter JS, Krebs EE, et al. Estimating minimally important differences for the PROMIS pain interference scales: results from 3 randomized clinical trials. Pain 2018 Apr 1;159(4):775-782 [FREE Full text] [doi: 10.1097/j.pain.0000000000001121] [Medline: 29200181]
77. Spitzer RL, Kroenke K, Williams JBW, 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]
78. Toussaint A, Hüsing P, Gumz A, Wingenfeld K, Härter M, Schramm E, et al. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). J Affect Disord 2020 Mar 15;265:395-401. [doi: 10.1016/j.jad.2020.01.032] [Medline: 32090765]
79. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. Med Care 2004 Dec;42(12):1194-1201. [doi: 10.1097/00005650-200412000-00006] [Medline: 15550799]
80. Fann JR, Berry DL, Wolpin S, Austin-Seymour M, Bush N, Halpenny B, et al. Depression screening using the Patient Health Questionnaire-9 administered on a touch screen computer. Psychooncology 2009 Jan;18(1):14-22 [FREE Full text] [doi: 10.1002/pon.1368] [Medline: 18457335]
81. Hatcher RL, Gillaspy JA. Development and validation of a revised short version of the working alliance inventory. Psychother Res 2006 Jan;16(1):12-25. [doi: 10.1080/10503300500352500]
82. Munder T, Wilmers F, Leonhart R, Linster HW, Barth J. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. Clin Psychol Psychother 2010;17(3):231-239. [doi: 10.1002/cpp.658] [Medline: 20013760]
83. Jasper K, Weise C, Conrad I, Andersson G, Hiller W, Kleinstäuber M. The working alliance in a randomized controlled trial comparing Internet-based self-help and face-to-face cognitive behavior therapy for chronic tinnitus. Internet Interv 2014 Apr;1(2):49-57. [doi: 10.1016/j.invent.2014.04.002]
84. Hauser-Ulrich S, Künzli H, Meier-Peterhans D, Kowatsch T. A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 3;8(4):e15806. [doi: 10.2196/15806]
85. Woolfolk RL, Carr-Kaffashan L, McNulty TF, Lehrer PM. Meditation training as a treatment for insomnia. Behav Ther 1976 May;7(3):359-365 [FREE Full text] [doi: 10.1016/S0005-7894(76)80064-0]
86. Floyd M, Scogin F. Cognitive-behavior therapy for older adults: how does it work? Psychother Theory Res Pract Train 1998;35(4):459-463 [FREE Full text] [doi: 10.1037/h0087770]
87. Leo A, Schuelke M, Hunt D, Metzler J, Miller J, Areán P, et al. A digital mental health intervention in an orthopedic setting for patients with symptoms of depression and/or anxiety: feasibility prospective cohort study. JMIR Form Res 2022 Feb 21;6(2):e34889 [FREE Full text] [doi: 10.2196/34889] [Medline: 35039278]

88. Kohrt BA, Griffith JL, Patel V. Chronic pain and mental health: integrated solutions for global problems. Pain 2018 Sep;159 Suppl 1(1):S85-S90 [FREE Full text] [doi: 10.1097/j.pain.0000000000001296] [Medline: 30113952]
89. Upshur CC, Bacigalupe G, Luckmann R. "They don't want anything to do with you": patient views of primary care management of chronic pain. Pain Med 2010 Dec 01;11(12):1791-1798 [FREE Full text] [doi: 10.1111/j.1526-4637.2010.00960.x] [Medline: 21029353]
90. Palermo TM, de la Vega R, Murray C, Law E, Zhou C. A digital health psychological intervention (WebMAP Mobile) for children and adolescents with chronic pain: results of a hybrid effectiveness-implementation stepped-wedge cluster randomized trial. Pain 2020 Dec 10;161(12):2763-2774 [FREE Full text] [doi: 10.1097/j.pain.0000000000001994] [Medline: 32658147]
91. Greenberg J, Popok PJ, Lin A, Kulich RJ, James P, Macklin EA, et al. A mind-body physical activity program for chronic pain with or without a digital monitoring device: proof-of-concept feasibility randomized controlled trial. JMIR Form Res 2020 Jun 08;4(6):e18703 [FREE Full text] [doi: 10.2196/18703] [Medline: 32348281]
92. Rockett M. Diagnosis, mechanisms and treatment of complex regional pain syndrome. Curr Opin Anaesthesiol 2014 Oct;27(5):494-500 [FREE Full text] [doi: 10.1097/ACO.0000000000000114] [Medline: 25111604]
93. Malik T, Ambrose A, Sinha C. User feedback analysis of an AI-enabled CBT mental health application (Wysa). JMIR Hum Factors 2022 Mar 06:1 [FREE Full text] [doi: 10.2196/35668] [Medline: 35249886]
94. Kinney M, Seider J, Beaty AF, Coughlin K, Dyal M, Clewley D. The impact of therapeutic alliance in physical therapy for chronic musculoskeletal pain: a systematic review of the literature. Physiother Theory Pract 2020 Aug 28;36(8):886-898. [doi: 10.1080/09593985.2018.1516015] [Medline: 30265840]
95. Ferreira P, Ferreira M, Maher C, Refshauge K, Latimer J, Adams R. The therapeutic alliance between clinicians and patients predicts outcome in chronic low back pain. Phys Ther 2013 Apr;93(4):470-478 [FREE Full text] [doi: 10.2522/ptj.20120137] [Medline: 23139428]
96. Fuentes J, Armijo-Olivo S, Funabashi M, Miciak M, Dick B, Warren S, et al. Enhanced therapeutic alliance modulates pain intensity and muscle pain sensitivity in patients with chronic low back pain: an experimental controlled study. Phys Ther 2014 Apr;94(4):477-489 [FREE Full text] [doi: 10.2522/ptj.20130118] [Medline: 24309616]
97. Rubinelli S, Schulz PJ, Vago F. Designing and evaluating online communities for promoting self-management of chronic low back pain. Int J Web Based Commun 2008;4(1):80 [FREE Full text] [doi: 10.1504/IJWBC.2008.016492]

Abbreviations

AI: artificial intelligence
AI-CBT: artificial intelligence–supported cognitive behavioral therapy
CBT: cognitive behavioral therapy
GAD-7: Generalized Anxiety Disorder, 7-item
NPRS: numerical pain rating scale
PHQ-9: Patient Health Questionnaire, 9-item
PROMIS PI: Patient-Reported Outcomes Measurement Information System–Pain Inference short form 6b
SELMA: Smartphone-Based Health Care Chatbot to Promote Self-Management of Chronic Pain
WAI-SR: Working Alliance Inventory–Short Revised

Edited by T Leung; submitted 29.01.22; peer-reviewed by S Hoermann, H Tanaka; comments to author 21.02.22; revised version received 07.03.22; accepted 21.03.22; published 31.03.22.

Please cite as:
Gupta M, Malik T, Sinha C
Delivery of a Mental Health Intervention for Chronic Pain Through an Artificial Intelligence–Enabled App (Wysa): Protocol for a Prospective Pilot Study
JMIR Res Protoc 2022;11(3):e36910
URL: https://www.researchprotocols.org/2022/3/e36910
doi:10.2196/36910
PMID:35314423

Protocol

Classification of Patients for Whom Benefit of Long-term Opioid Therapy No Longer Outweighs Harm: Protocol for a Delphi Study

Raymond Van Cleve[1,2], MS, PhD; Sara Edmond[3,4], PhD; Jennifer Snow[4], MSc; Anne C Black[5], PhD; Jamie L Pomeranz[6], PhD; William Becker[4,5], MD

[1]Center for Innovation to Implementation, Menlo Park, CA, United States

[2]Stanford University, Palo Alto, CA, United States

[3]Department of Psychiatry, Yale School of Medicine, New Haven, CT, United States

[4]Pain Research, Informatics, Multimorbidities, and Education Center, VA Connecticut Healthcare System, New Haven, CT, United States

[5]Department of Internal Medicine, Yale School of Medicine, New Haven, CT, United States

[6]Department of Occupational Therapy, University of Florida, Gainsville, FL, United States

Corresponding Author:
Raymond Van Cleve, MS, PhD
Center for Innovation to Implementation
795 Willow Road (152-MPD)
Menlo Park, CA, 94025
United States
Phone: 1 (650) 617 2746
Email: vancleve.raymond@gmail.com

Abstract

Background: Patients with chronic pain prescribed long-term opioid therapy may come to a point where the benefits of the therapy are outweighed by the risks and tapering is indicated. At the 2019 Veterans Health Administration State of the Art Conference, there was an acknowledgment of a lack of clinical guidance with regard to treating this subset of patients. Some of the participants believed clinicians and patients would both benefit from a new diagnostic entity describing this situation.

Objective: The aim of this study was to determine if a new diagnostic entity was needed and what the criteria of the diagnostic entity would be. Given the ability of the Delphi method to synthesize input from a broad range of experts, we felt this technique was the most appropriate for this study.

Methods: We designed a modified Delphi technique involving 3 rounds. The first round is a series of open-ended questions asking about the necessity of this diagnostic entity, how this condition is different from opioid use disorder, and what its possible diagnostic criteria would be. After synthesizing the responses collected, a second round will be conducted to ask participants to rate the different responses offered by their peers. These ratings will be collected and analyzed, and will generate a preliminary definition for this clinical phenomena. In the third round, we will circulate this definition with the aim of achieving consensus.

Results: The modified Delphi study was initiated in July of 2020 and analysis is currently underway.

Conclusions: This protocol has been approved by the Internal Review Board at the Connecticut Veterans Affairs and the study is in process. This protocol may assist other researchers conducting similar studies.

International Registered Report Identifier (IRRID): DERR1-10.2196/33310

(JMIR Res Protoc 2022;11(3):e33310) doi:10.2196/33310

KEYWORDS

modified Delphi technique; long-term opioid treatment; chronic pain; opioid therapy; opioids; pain management; Delphi study

Introduction

Background

Although the number of opioid prescriptions has decreased since 2012 [1], long-term opioid therapy (LTOT) remains a common treatment for chronic pain [2], with the duration of therapy potentially lasting for years [3,4]. Patients prescribed LTOT for pain are at risk for adverse outcomes, including worsening pain and function and developing opioid use disorder (OUD) [5]. A challenge in clinical, research, and policy spheres is determining whether and how to apply the Diagnostic and

Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)'s OUD criteria to patients receiving LTOT for whom benefit is no longer outweighing harm and tapering is thus indicated per consensus guidelines [6-8]. The DSM-5 criteria for OUD are designed to identify a condition in which patients have compulsive use of opioids leading to adverse consequences and do not necessarily have comorbid chronic pain [9].

Thus, some experts have argued that a new diagnostic entity specifically intended for patients on LTOT for whom harm is outweighing benefit would help advance research and clinical care. These experts believe concurrent opioid dependence and chronic pain are more than the sum of their parts, and a new diagnostic entity could address this complexity [10], something clinicians and researchers have struggled with in using the OUD diagnosis on its own. During the 2019 Veterans Health Administration (VHA) Health Services Research and Development (HSR&D) State of the Art (SOTA) Conference, consensus emerged on the need for a Delphi study to understand if a new diagnostic entity is needed and, if so, to develop consensus on its criteria and characteristics [11]. The modified Delphi method is designed to gain consensus on a discrete topic that does not yet have a clear definition [12] and has been used in previous studies concerning OUD. Given the problem articulated during the SOTA Conference, this method lends itself well to investigating a problem of this complexity [13,14]. The aim of this paper is to describe the protocol for a Delphi study to explore the need for and criteria of a new diagnostic entity characterizing the clinical scenario of benefit no longer outweighing harm of LTOT for chronic pain.

Study Objectives

The objectives of this Delphi study are to (1) explore perspectives on the merits of creating a new diagnostic entity, separate from but not replacing OUD, that better characterizes the scenario of benefits no longer outweighing harm of LTOT for chronic pain and (2) develop consensus on its definition and diagnostic criteria. We will present questions in an open-ended format to avoid bias and allow the responses of the participants to guide the analysis.

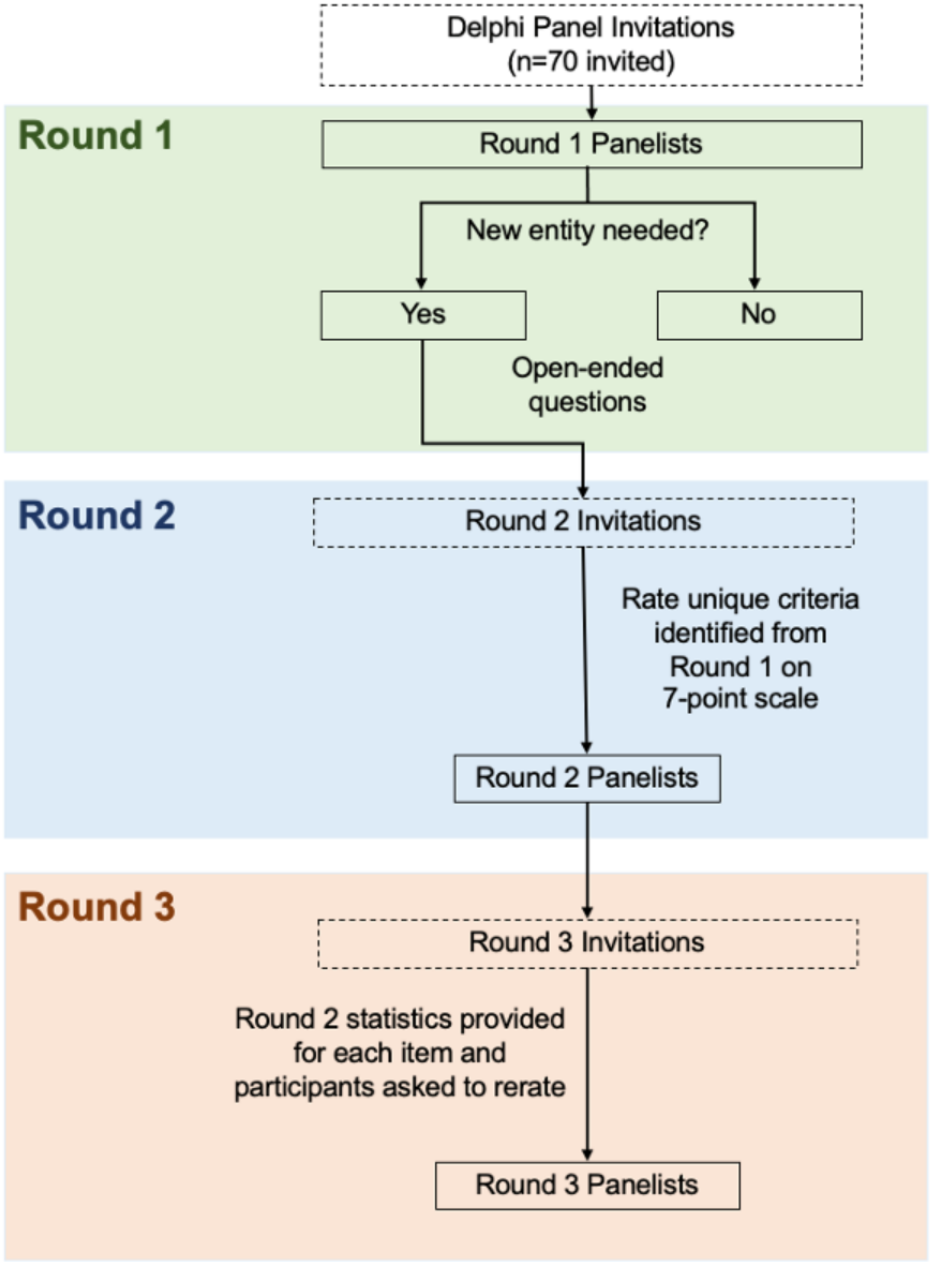
Methods

Overview

The Delphi method is well suited for surveying expert panels to gain consensus on a clinical problem that is not well defined and determine the defining elements of that problem. Consistent with other related studies, we will conduct the Delphi study remotely to promote independent contribution and avoid "bandwagon" or "halo" effects, and generate as many ideas as possible from each individual contributor [15]. The halo effect refers to an instance where a positive response given by a member of a panel influences subsequent responses [16] and the bandwagon effect is when one theme identified by a member of a panel alters subsequent responses [17]. Given the structure of the Delphi method, we were not concerned about these biases occurring.

The Delphi study flow is shown in Figure 1. The first round of this Delphi process will have a screening question, described below, and will then elicit responses to open-ended questions on the potential new diagnostic entity. Once these responses are collected, they will be analyzed using rapid qualitative matrix and content analysis [18-20]. Results from the qualitative analysis will summarize expert-proposed diagnostic criteria that will be evaluated using numeric rating scales in round 2. Round 2 group statistics summarizing ratings of each proposed criterion (ie, means, standard deviation, the medians, and interquartile ranges) will be presented to the expert panel in round 3, when experts will be asked to review their previous ratings relative to the measures of central tendency, and to re-evaluate each item. Results from the third round will be summarized in accordance with the consensus criteria originally established prior to the initiation of the Delphi study.

Figure 1. CONSORT flow diagram for a 3-round Delphi study.



Selection of Delphi Participants and Inclusion Criteria

As above, VHA HSR&D convened a 2-day SOTA Conference on effective pain management and opioid safety in September 2019. The purpose of the SOTA Conference was to convene a multidisciplinary group of experts to help HSR&D develop specific research priorities related to increasing access to medications for OUD, managing LTOT for pain including tapering, and treating co-occurring pain and substance use disorders [21]. To ensure as robust a process as possible, HSR&D invited a diverse array of subject matter experts from within and outside the VHA with expertise in pain, OUD, opioid safety, and clinical research representing the following disciplines: general internal medicine, psychology, addiction medicine, addiction psychiatry, nursing, pharmacy, pain medicine, neurology, clinical epidemiology, health services research, and health policy. We will recruit participants for the Delphi study from the list of invitees to the SOTA Conference (N=70). We explicitly chose to recruit invitees rather than participants in the SOTA Conference as we did not want to exclude the perspectives of experts who arbitrarily were unavailable to attend the SOTA Conference.

We set a date for the completion of each round of the Delphi study and will send out regular reminders to participants who have not completed that round to try to minimize attrition. We do not anticipate significant attrition given that many of the participants are already collaborating on similar projects and have voiced an interest in participating in this type of study. However, if we observe significant attrition (eg, >20%), we will consider alternative data collection methods such as offering to conduct data collection via phone.

Delphi Process

Overview

All 3 rounds of the survey will be administered by Qualtrics XM (Qualtrics International Inc), which has been licensed to our research team through Yale University. We will send individual emails to participants explaining the Delphi process and why they were selected to participate. In the email, we will acknowledge the discussion at the 2019 VHA HSR&D SOTA Conference that prompted this study. The survey will contain a link to the DSM-5 criteria for OUD for participant reference. We will allow approximately 1 month to complete each round, with some reminder emails if necessary.

Round 1

Aim

The aims of round 1 are as follows: (1) to understand the various expert perspectives on the merits of creating a separate diagnostic entity and (2) to gather information about the defining characteristics of the clinical phenomenon requiring operationalization, as well as information about how the phenomenon relates to OUD via several open-ended questions. We designed a survey that will take approximately 25-30 minutes to complete. Questions such as "Please describe a person who would be diagnosed with Condition X" are meant to prime participants to envision how these patients might present in a clinical setting. We also wanted to encourage participants to consider the bounds of this condition, defining both what it is and what it is not.

Data Collection

The round 1 survey can be found in Textbox 1. To meet aim 1, we will use the question, "Do you think a new diagnostic entity is needed for patients who have been taking opioids and for whom the potential harms of the therapy outweigh the benefits of the therapy?" as a screening question; only participants who answer "yes" to this question will be invited to complete the rest of the survey. As shown in Textbox 1, the remainder of round 1 items are a mix of yes/no and free-text questions.

Textbox 1. Questions included in round 1 of this Delphi study.

Note: Questions 2-9 were only asked of those who said "yes" to question 1.

1. Do you think a new diagnostic entity is needed for patients on long-term opioid therapy for pain that is related to but distinct from the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition's opioid use disorder definition? (yes/no)
   - If yes, please explain why. If no, please explain why not.
2. Please describe a person who would be diagnosed with Condition X (How did he or she present? What were they prescribed? How did the treatment course go? What behaviors manifest themselves over time? How is this person different from a person with opioid use disorder)?
3. How would you differentiate Condition X from opioid use disorder?
4. Please complete the sentence "Condition X is defined as ____"
5. Please list the diagnostic criteria for Condition X:
6. Should Condition X have different gradations (eg, mild, medium, severe)? (yes/no)
   - If yes, what gradations would you recommend? How would you distinguish between the different gradations?
   - If no, why not?
7. Is Condition X related to opioid use disorder? (yes/no)
   - If yes, how so?
8. What are the differences in how Condition X should be treated compared to opioid use disorder?
9. We have been using the term Condition X as a placeholder for ease of discussion. What do you think this condition should be called?

Data Analysis

We will use rapid qualitative analysis [20] to summarize free-text, open-ended answers. Summaries of individual responses will be inputted into a data matrix that will be used to identify distinct concepts within each section of the survey. We will then perform a content analysis to assess the concentration of each concept and generate a list of potential diagnostic criteria. The list of potential diagnostic criteria will be the basis for the second round. We will only include criteria mentioned by at least two Delphi participants in the round 2 survey.

Round 2

Aim

The objective of the second round will be to begin building consensus on the potential criteria defining the proposed new diagnostic entity. We will ask the participants to evaluate and rate the relevance of each criterion to the diagnostic entity. The experts will have the opportunity to critique the initial responses (that have been analyzed and reformatted) and suggest the criteria that should be retained.

Data Collection

The round 2 survey will list each potential criterion and ask participants to answer, "to what extent do you agree that each

of the following features/criteria should be included as a feature/criterion of Condition X?" Participants will answer on a 7-point Likert scale from "strongly disagree" to "strongly agree." We will also ask participants to comment on their preferred wording among alternatives, and provide suggestions for alternative wording for diagnostic criteria. Finally, we will ask participants to indicate their favored proposed names for the new diagnostic entity.

Data Analysis

For each item, we will calculate the mean, standard deviation, median, and interquartile range of ratings. We agreed a priori that diagnostic criteria with a median of 5 or greater will indicate consensus for inclusion, median scores between 3 and 5 will indicate the need for further exploration, and a median score below 3 will indicate that an item should not be included. We will use rapid qualitative analysis to explore the qualitative feedback. Qualitative feedback will be used to identify any additional potential criteria for inclusion in round 3.

Round 3

Aim

This final round will ask panelists to rerate potential criteria in the context of central tendency statistics in an effort to generate further consensus.

Data Collection

We will present to the group the mean, standard deviation, median, and interquartile range of ratings for each proposed criterion from round 2, along with that panelist's initial response on the 7-point Likert scale. Participants will be asked to rerate each item from round 2 and provide initial ratings for any additional items identified based on round 2 qualitative feedback.

Data Analysis

Consistent with prior Delphi studies, items with a median of 5 or greater after round 3 will be recommended for inclusion in the criteria for the new diagnostic entity.

Ethics Approval

This protocol was approved by the Institutional Review Board of VA Connecticut Healthcare System (approval number: 1624733-3).

Results

Data collection for round 1 began in July 2020; we have completed the first two rounds and are currently collecting round 3 data. We plan to publish qualitative findings from round 1 and overall findings from the entire study.

Discussion

Herein, we describe a protocol for a Delphi study designed to develop criteria to characterize the phenomenon experienced by patients with chronic pain prescribed LTOT when benefits of the therapy no longer outweigh harms. The protocol uses rigorous methods to iteratively generate consensus among a diverse group of subject matter experts; as such, it holds promise in providing momentum forward on a topic that has become increasingly relevant during the United States' efforts to improve the safety of opioid prescribing, including tapering when harm outweighs benefit.

Criteria for the novel diagnostic entity may be used to standardize the definition of the new entity, promote targeted research, inform clinical practice guidelines for its treatment, and ultimately improve quality of care. The Delphi method provides the process to accomplish this goal by iteratively developing expert consensus regarding the definition of the entity and a preliminary list of its diagnostic criteria.

If consensus is achieved around a definition and list of criteria for a new diagnostic entity, the next steps may include convening a work group to refine the criteria and evaluating the criteria in practice with regard to sensitivity and specificity. Identification of an entity distinct from OUD could inform practice specific to that entity. Future work will include exploring best pathways for disseminating our findings to frontline clinicians who work with patients prescribed LTOT, and determining how to use our findings to inform future research, policy work, and clinical care, such as developing better treatments for patients with chronic pain prescribed LTOT and helping frontline clinicians assess and provide tailored treatment options.

Acknowledgments

This work was partially supported by the Center for Innovation to Implementation at the Palo Alto VA Medical Center via the Office of Academic Affiliation's Big Data Scientist Training Enhancement Program and the Pain Research, Informatics, Medical Comorbidities, and Education Center of Innovation of the VA Connecticut Healthcare System.

This material is based on work supported by the Pain/Opioid Consortium of Research (CORE), funded by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Health Services Research and Development.

We thank Brent Moore, PhD, Alicia Heapy, PhD, and Erin Krebs, MD, for their assistance with this work.

Disclaimer

The views expressed in this publication do not represent the views of the Veterans Health Administration or the United States Government.

Conflicts of Interest

None declared.

References

1. Zhu W, Chernew ME, Sherry TB, Maestas N. Initial Opioid Prescriptions among U.S. Commercially Insured Patients, 2012-2017. N Engl J Med 2019 Mar 14;380(11):1043-1052 [FREE Full text] [doi: 10.1056/NEJMsa1807069] [Medline: 30865798]
2. Ballantyne JC. Opioids for the Treatment of Chronic Pain: Mistakes Made, Lessons Learned, and Future Directions. Anesth Analg 2017 Nov;125(5):1769-1778. [doi: 10.1213/ANE.0000000000002500] [Medline: 29049121]
3. Mojtabai R. National trends in long-term use of prescription opioids. Pharmacoepidemiol Drug Saf 2018 May 06;27(5):526-534. [doi: 10.1002/pds.4278] [Medline: 28879660]
4. Barry DT, Marshall BD, Becker WC, Gordon AJ, Crystal S, Kerns RD, et al. Duration of opioid prescriptions predicts incident nonmedical use of prescription opioids among U.S. veterans receiving medical care. Drug Alcohol Depend 2018 Oct 01;191:348-354 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.07.008] [Medline: 30176548]
5. Frank JW, Lovejoy TI, Becker WC, Morasco BJ, Koenig CJ, Hoffecker L, et al. Patient Outcomes in Dose Reduction or Discontinuation of Long-Term Opioid Therapy: A Systematic Review. Ann Intern Med 2017 Aug 01;167(3):181-191 [FREE Full text] [doi: 10.7326/M17-0598] [Medline: 28715848]
6. Manhapra A, Sullivan MD, Ballantyne JC, MacLean RR, Becker WC. Complex Persistent Opioid Dependence with Long-term Opioids: a Gray Area That Needs Definition, Better Understanding, Treatment Guidance, and Policy Changes. J Gen Intern Med 2020 Dec 06;35(Suppl 3):964-971 [FREE Full text] [doi: 10.1007/s11606-020-06251-w] [Medline: 33159241]
7. Card P. VA/DoD Clinical Practice Guideline for Opioid Therapy for Chronic Pain. Department Veterans Aff Department Defense. 2017. URL: https://www.va.gov/HOMELESS/nchav/resources/docs/mental-health/substance-abuse/VA_DoD-CLINICAL-PRACTICE-GUIDELINE-FOR-OPIOID-THERAPY-FOR-CHRONIC-PAIN-508.pdf [accessed 2022-02-16]
8. Dowell D, Haegerich TM, Chou R. CDC Guideline for Prescribing Opioids for Chronic Pain--United States, 2016. JAMA 2016 Apr 19;315(15):1624-1645 [FREE Full text] [doi: 10.1001/jama.2016.1464] [Medline: 26977696]
9. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Vol. 5th. Washington, DC: American Psychiatric Association; 2013:540-546.
10. Manhapra A, Becker WC. Pain and Addiction: An Integrative Therapeutic Approach. Med Clin North Am 2018 Jul;102(4):745-763. [doi: 10.1016/j.mcna.2018.02.013] [Medline: 29933827]
11. Becker WC, Krebs EE, Edmond SN, Lin LA, Sullivan MD, Weiss RD, et al. A Research Agenda for Advancing Strategies to Improve Opioid Safety: Findings from a VHA State of the Art Conference. J Gen Intern Med 2020 Dec 03;35(Suppl 3):978-982 [FREE Full text] [doi: 10.1007/s11606-020-06260-9] [Medline: 33145681]
12. Murray TJ. Delphi methodologies: A review and critique. Urban Systems 1979 Jan;4(2):153-158. [doi: 10.1016/0147-8001(79)90013-5]
13. Merlin JS, Young SR, Azari S, Becker WC, Liebschutz JM, Pomeranz J, et al. Management of problematic behaviours among individuals on long-term opioid therapy: protocol for a Delphi study. BMJ Open 2016 May 06;6(5):e011619 [FREE Full text] [doi: 10.1136/bmjopen-2016-011619] [Medline: 27154486]
14. Merlin JS, Young SR, Starrels JL, Azari S, Edelman EJ, Pomeranz J, et al. Managing Concerning Behaviors in Patients Prescribed Opioids for Chronic Pain: A Delphi Study. J Gen Intern Med 2018 Feb;33(2):166-176 [FREE Full text] [doi: 10.1007/s11606-017-4211-y] [Medline: 29204977]
15. Steinert M. A dissensus based online Delphi approach: An explorative research tool. Technological Forecasting and Social Change 2009 Mar;76(3):291-300. [doi: 10.1016/j.techfore.2008.10.006]
16. Nisbett RE, Wilson TD. The halo effect: Evidence for unconscious alteration of judgments. Journal of Personality and Social Psychology 1977;35(4):250-256. [doi: 10.1037/0022-3514.35.4.250]
17. Nadeau R, Cloutier E, Guay J. New Evidence About the Existence of a Bandwagon Effect in the Opinion Formation Process. International Political Science Review 2016 Jun 29;14(2):203-213. [doi: 10.1177/019251219301400204]
18. Palinkas LA, Mendon SJ, Hamilton AB. Innovations in Mixed Methods Evaluations. Annu Rev Public Health 2019 Apr 01;40(1):423-442 [FREE Full text] [doi: 10.1146/annurev-publhealth-040218-044215] [Medline: 30633710]
19. Nevedal AL, Reardon CM, Widerquist MA, Jackson GL, Cutrona SL, White BS, et al. Rapid Versus Traditional Qualitative Analysis Using the Consolidated Framework for Implementation Research (CFIR). Implementation Science 2021 Dec 06:1-15. [doi: 10.21203/rs.3.rs-137424/v1]

20. Hamilton AB, Finley EP. Qualitative methods in implementation research: An introduction. Psychiatry Res 2019 Oct;280:112516 [FREE Full text] [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
21. Becker WC, Humphreys K, Atkins D, Clancy CM. Mounting a Scientifically Informed Response to the Opioid Crisis in the Veterans Health Administration. J Gen Intern Med 2020 Dec;35(Suppl 3):883-885 [FREE Full text] [doi: 10.1007/s11606-020-06349-1] [Medline: 33277686]

Abbreviations

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
HSR&D: Health Services Research and Development
LTOT: long-term opioid therapy
OUD: opioid use disorder
SOTA: State of the Art
VHA: Veterans Health Administration

Edited by G Eysenbach; submitted 01.09.21; peer-reviewed by R Lee, SW Bae, M Lall; comments to author 18.12.21; revised version received 30.12.21; accepted 31.12.21; published 04.03.22.

Please cite as:
Van Cleve R, Edmond S, Snow J, Black AC, Pomeranz JL, Becker W
Classification of Patients for Whom Benefit of Long-term Opioid Therapy No Longer Outweighs Harm: Protocol for a Delphi Study
JMIR Res Protoc 2022;11(3):e33310
URL: https://www.researchprotocols.org/2022/3/e33310
doi:10.2196/33310
PMID:35254277

Protocol

Goal Attainment Scaling in Outpatient Physical Therapy for Chronic Low Back Pain: Protocol for a Mixed Methods Study

Douglas Haladay[1], MHS, DPT, PhD; Rebecca Edgeworth Ditwiler[1], DPT; Aimee B Klein[1], DPT, DSc; Rebecca Miro[1], PhD; Matthew Lazinski[1], DPT; Laura Lee Swisher[1], MDiv, PhD; Jason Beckstead[2], PhD; Jay Wolfson[2], DrPH, JD; Dustin Hardwick[1], DPT, PhD

[1]School of Physical Therapy & Rehabilitation Sciences, Morsani College of Medicine, University of South Florida, Tampa, FL, United States
[2]College of Public Health, University of South Florida, Tampa, FL, United States

Corresponding Author:
Douglas Haladay, MHS, DPT, PhD
School of Physical Therapy & Rehabilitation Sciences
Morsani College of Medicine
University of South Florida
12901 Bruce B Downs Blvd MDC 77
Tampa, FL, 33612
United States
Phone: 1 8139741971
Email: dhaladay@usf.edu

Abstract

Background: Patient engagement in decisions regarding their health care may lead to improved outcomes and improved adherence to treatment plans. While there are several options for involving patients in their health care, goal setting is a readily accessible method for physical therapists to increase the involvement of patients in health care decisions. Physical therapy goals are often generated by health care providers based on subjective information or standardized, fixed-item, patient-reported outcome measures. However, these outcome measures may not fully reveal the activity and participation limitations of individual patients. Goal attainment scaling (GAS) is a patient-centered approach that allows patients to set meaningful goals. While GAS has been shown to be reliable, valid, and sensitive to change in various populations, there is limited evidence in the United States on utilizing GAS in physical therapy for patients with chronic low back pain (LBP).

Objective: The purpose of this paper is to describe the protocol for a study to (1) develop a way to apply GAS procedures for physical therapists treating patients with chronic LBP in the United States and (2) test the feasibility of applying GAS procedures for chronic LBP in an outpatient physical therapy setting.

Methods: This study used a mixed methods design with 2 phases: qualitative and quantitative. The qualitative phase of the study employed focus groups of patients with chronic LBP to identify an inventory of goals that were important and measurable. A series of prompts was developed from this inventory to assist physical therapists in collaboratively establishing goals with patients in a clinical setting. The quantitative phase of the study pilot-tested the inventory developed in the qualitative phase in patients with chronic LBP to determine feasibility, reliability, validity, and responsiveness. We also plan to compare how well GAS reveals change over time relative to traditional, fixed-item, patient-reported measures.

Results: Phase 1 data collection was completed in June 2020, while data collection for phase 2 was performed between March 2021 and December 2021. We anticipate that this study will demonstrate that GAS can be implemented successfully by outpatient physical therapists, and that it will demonstrate clinically important changes in patients with chronic LBP.

Conclusions: GAS represents an opportunity for patient-centered care in the physical therapy management of chronic LBP. While GAS is not new, it has never been studied in real-world physical therapy for chronic LBP in a clinical setting. Due to unique time and productivity constraints, for GAS to be successfully implemented in this environment, we must demonstrate that clinicians can be trained efficiently and reliably, that GAS can be implemented in a clinical setting in under 15 minutes, and that GAS is able to detect clinically meaningful changes in patient outcomes.

International Registered Report Identifier (IRRID): DERR1-10.2196/32457

(JMIR Res Protoc 2022;11(3):e32457) doi:10.2196/32457

KEYWORDS

goal attainment scaling; goal setting; low back pain; chronic pain; physical therapy; patient engagement; adherence; rehabilitation; physical therapist

Introduction

The patient experience and patient-centered care are the core of the Institute for Healthcare Improvement's "Triple Aim" (eg, population health, experience of care, and per capita cost) for optimizing the performance of health systems in the United States [1]. Patient engagement in decisions regarding their health care may lead to improved outcomes and improved adherence to treatment plans [2,3]. The interaction between the patient and provider is an essential element of patient-centered care [4]. The qualities of these interactions may best be judged by patients themselves. Patients value providers who listen to them, share information via dialogue, and consider their individual preferences in management of their health conditions [5]. In physical therapy, patients are satisfied if their physical therapist communicates effectively and spends adequate time explaining treatment options throughout the course of care [6]. Furthermore, outcomes and perceived quality of care may improve when patients are actively engaged in their own care [7].

While there are several options for involving patients in their health care, goal setting is a readily accessible method for physical therapists to increase the involvement of patients in health care decisions. Goal setting is an important part of physical therapy in episodes of care and in direct interventions, but the practice and implementation of goal setting is varied across the profession [8]. When physical therapists set goals, it is an opportunity to involve patients and to design interventions that consider individual patient needs [7].

Physical therapy goals are often provider generated and based on subjective information or standardized, fixed-item, patient-reported outcome measures [9]. However, these outcome measures may not fully reveal the activity and participation limitations of individual patients. As individuals have varied needs, their goals may not be identified by standardized measures, and therefore a patient's particular goals may not be reflected in provider-directed goals. We have investigated patients' views on whether pain, disability, and recovery measures are meaningful and have found that people with chronic low back pain (LBP) feel that standard measures used to classify patient goals do not reveal what is meaningful to them. Participants in focus groups often state that the standard outcome measures do not capture the fluctuating nature of symptoms or assess improvements in more active pursuits and often do not reveal the complex nature of social roles. While standardized outcome measures are useful for comparing populations, they may be of limited value when assessing individual patient-centered goals [9].

There are several patient-centered approaches used to involve patients in setting meaningful, individualized goals, including the Canadian Occupational Performance Measure, goal attainment scaling (GAS), and self-identified goals assessment [10-12]. GAS has been identified as one of the most time-efficient and reliable ways to involve patients in goal generation during clinical care in real-world settings [11]. GAS procedures are highly variable, with little consensus on the time needed to complete them in clinical practice. The reported time to complete the GAS process ranges from 5 minutes to 60 minutes [13-15]. This range in time to complete may be due to variations in GAS methods, such as the extent of patient involvement, family involvement, and whether GAS was completed by a team or an individual provider. In addition, setting specific goals may be more time-consuming in certain patient populations.

During GAS, patients are engaged in a dialogue to set specific, measurable, achievable, realistic, relevant, and time-based (SMART) goals [16]. Physical therapists often write goals using the SMART format and are well versed in writing SMART patient goals. The GAS procedure involves a discussion between the patient and provider about patient-directed goals and expected outcomes of treatment. This provides an opportunity for physical therapists to capture fluctuating pain levels, specific activities, and complex social responsibilities that are important to patients. The GAS process is readily accessible, free, and follows defined stages (ie, identifying goals, weighting goals, identifying expected outcomes, establishing a baseline score, and judging actual outcome versus expected outcome at follow-up) [11]. In contrast to standardized, fixed-item, patient-reported outcome measures, GAS generates a T score, which provides a numerical outcome of achievement that can be used for goals across the domains of the International Classification of Functioning, Disability, and Health (ICF) with varying difficulty, importance, and expected achievement [17]. The individualized outcome generated from GAS may prove helpful in enhancing traditional methods of collecting outcome measures and setting goals.

GAS is helpful when comparing heterogenous patient populations, who may have complex presentations and backgrounds [17]. Therefore, most of the literature related to applying GAS describes findings in a rehabilitation setting with pediatric patients or patients with neurological deficits [17,18]. Recently, GAS has been applied in patients with chronic LBP [2,9,19-22], as patients with chronic LBP have varying clinical presentations, severity levels, and treatment options [23]. Considering that LBP is the most common reason patients seek physical therapy in an outpatient setting [24] and that it is one of the most common causes of disability in the United States [25-27], patients with chronic LBP may be an ideal population for assessing the feasibility of GAS in a physical therapy outpatient setting. While there is evidence for the reliability, validity, and feasibility of GAS procedures [17], there is limited evidence in the United States about physical therapist use of GAS in the management of patients with chronic LBP. Furthermore, a recent systematic review found significant variability in GAS procedures used for patients with LBP and recommended development of a standardized approach and training for clinicians applying GAS [28]. GAS is a promising method of focusing on patient-centered outcomes and goals,

but it is not clear how this method may complement standard, pre-existing outcome measures used by physical therapists and how feasible it is in an outpatient setting for patients with chronic LBP. GAS is novel because it provides a standardized means to set patient-provided goals that are quantifiable and can be used to track progress of the patient and compare outcomes across patients. Therefore, the purpose of this paper is to describe the protocol for a study to (1) develop a new application of GAS procedures to be used by physical therapists treating patients with chronic LBP in the United States and (2) test the feasibility of applying GAS procedures in the treatment of patients with chronic LBP in an outpatient physical therapy setting.

Methods

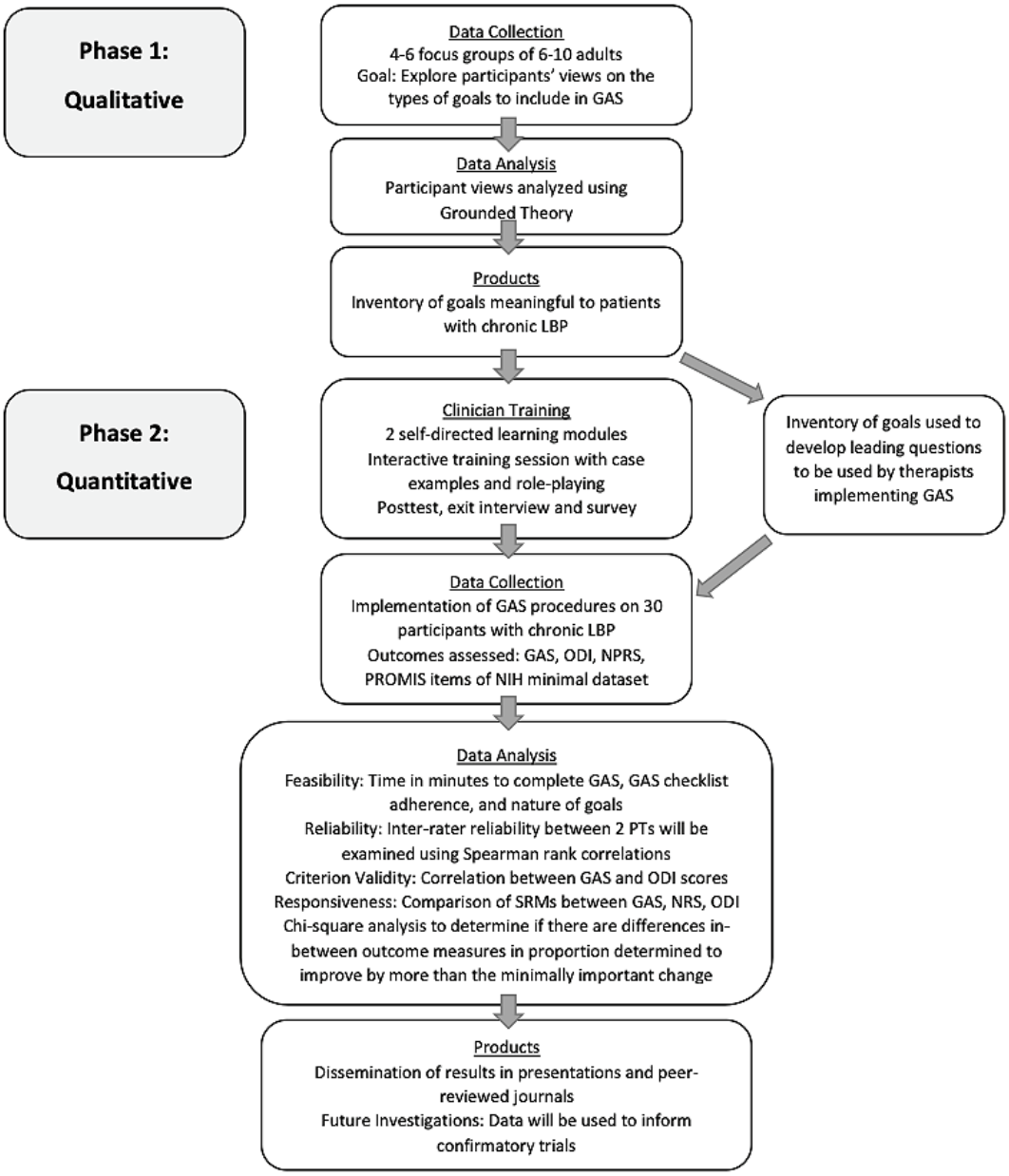
Ethics Approval

The University of South Florida (USF) Institutional Review Board approved this study on April 10, 2019 (Pro00035236), and the approval has been maintained in good standing.

Study Design

This study used a mixed methods design with 2 phases: qualitative and quantitative. Figure 1 shows a study overview. The qualitative phase of the study employed focus groups of patients with chronic LBP to identify an inventory of goals that are important and measurable. This inventory was used to develop a series of prompts that will allow physical therapists to assist patients in establishing goals in a clinical setting [22]. The quantitative phase of the study pilot-tested the inventory developed in the qualitative phase in patients with chronic LBP to determine feasibility, reliability, validity, and responsiveness. We will also compare how well GAS identifies changes over time compared to traditional, fixed-item, patient-reported measures.

Figure 1. Study overview. GAS: goal attainment scaling; LBP: low back pain; NIH: National Institutes of Health; NPRS: numerical pain rating scale; ODI: Oswestry disability index; PROMIS: Patient-Reported Outcomes Measurement Information System; PT: physical therapist; SRM: standardized response mean.



Phase 1

Participants

We assembled 4 to 6 focus groups comprising 6 to 10 adults with chronic LBP from the local community using research alerts sent via the university email listservs. We expected that approximately 30 adults would be needed for this study, as a minimum of 4 focus group discussions are needed to reach code saturation [29-31]. Participants were included if they were adults (aged 21-64 years) with a history of nonspecific chronic LBP lasting >12 weeks with or without radicular symptoms. Participants were excluded if they had a structural spinal deformity, spinal fracture, osteoporosis, or systemic disease; had undergone previous spinal surgery; were pregnant or had given birth within the last 6 months; had pending litigation related to worker's compensation; or were undergoing treatment covered under worker's compensation. Participants were recruited from the Tampa Bay area.

Procedures

Focus groups were conducted face-to-face via internet-based meetings to explore participants' views about the types of goals that should be included in GAS for patients with chronic LBP. There were 6 meetings with 6 to 10 participants that lasted approximately 2 hours [29]. An experienced facilitator led discussions. Field notes were taken during the interviews and

audio recordings of each focus group were transcribed verbatim for further analysis.

Data Analysis

Participant views were examined with grounded theory principles [32,33]. A grounded theory approach was chosen because the intent of the qualitative portion of the study was to understand what was important about patients' self-identified goals and why. This understanding was used to develop a standardized language and inventory of goals to facilitate the clinical implementation of GAS. We initially became familiar with the data by verifying the transcripts against the audio files and the field notes from each focus group to ensure accuracy and validation of speech allocation to individual participants. Coding began after the transcripts were read and became familiar to the researchers. Data were collected and coded until no new information was found (ie, saturation of the data was achieved). Qualitative data management software (MAXQDA, VERBI Software) was used to facilitate this process.

Once collected and transcribed, the data were independently coded, compared, and organized by 2 or more researchers using a constant comparison method. This qualitative procedure allows meaningful statements from the transcripts to be conceptualized in new ways [34]. The data coding process included open coding to determine code categories. Open coding was guided by sensitizing concepts found in the literature, including the ICF domains. Coding proceeded to the axial and selective phases to identify patterns in the data and identify the central themes that emerged. Several strategies were used to enhance rigor, including analytical triangulation using multiple coders, setting the goal for intercoder reliability to κ=0.80 (Cohen κ) using a coder-by-coder agreement matrix, peer debriefing group meetings to minimize researcher bias, and member checking to verify findings after focus group participation.

Phase 2

Participants

We recruited approximately 30 patients with chronic LBP who sought physical therapy from the USF Physical Therapy Center at the USF Morsani Center for Advanced Healthcare. A sample of 30 participants is considered appropriate for pilot studies of feasibility, and sufficient to estimate effect sizes for confirmatory trials [35,36]. In order to estimate effect size, we used the findings from this study and retrospective change scores from patients with chronic LBP who did not participate in GAS (from our own clinic and from published data) as a comparison group. The USF Physical Therapy Center is the faculty practice of the USF School of Physical Therapy and Rehabilitation Science. This center services the Tampa Bay area and admits approximately 30 patients with chronic LBP each month. The demographics of Tampa Bay closely match national demographics and are very diverse socioeconomically, racially, and ethnically, making Tampa a strategically desirable location for clinical trials.

Participants who met the following criteria were sequentially recruited: aged 21-64, LBP located between the lower rib cage and gluteal fold [37], pain lasting >12 weeks [37-39], pain that was not attributable to a specific pathology [40], pain on at least 50% of days in the past 6 months [37], and average pain intensity >2 out of 10 on the Numerical Pain Rating Scale. Participants with a spinal deformity, surgery or fracture, rheumatoid arthritis, extremity pain, physical therapy treatment within the past 6 months, or automobile- or work-related injury were excluded.

Procedures

Physical therapists from the USF Physical Therapy Center were trained in GAS procedures based upon the principles of Williams and Stieg [2]; these procedures have been described for use in rehabilitation [41]. Therapist training consisted of 2 self-directed learning modules that covered background information, goal setting and negotiation, the benefits of GAS in diverse patient populations, the use of GAS in patients with chronic LBP, and the implementation of the stages of GAS. Case examples, including videos, were interwoven throughout the modules and an assessment was completed upon conclusion of training. The final step in the training was an interactive session with a study investigator that included role-playing the GAS procedures and providing feedback on performance. Once the therapists were trained (but before they saw patients as part of this study), they completed a posttest and a short survey and were briefly interviewed to examine their views regarding the feasibility of the GAS process. The training design will be streamlined for future studies to facilitate deployment to clinicians. Therapist interviews were transcribed for thematic analysis.

Once the therapists completed training and met all assessment criteria, therapist and patient encounters using GAS commenced. To recruit patients from the USF Morsani Physical Therapy Center, a clinician (included as a research staff member in this study) reviewed electronic medical records of incoming patients. Patients with chronic LBP were contacted via email or traditional mail (if a patient did not have or did not provide an email address) ahead of their visit with information regarding the study. This allowed the potential participant sufficient time prior to their first visit to consider whether they wished to take part in the study. Interested patients were instructed to contact members of the study team. Potential participants were screened and provided consent before participation. During the clinician encounter, which was part of their normal therapy visit, the patient and therapist jointly completed the first part of the GAS using the GAS-Back form. In addition, the participants self-reported the following measures: the Oswestry Disability Index, the Numerical Pain Rating Scale, and the National Institutes of Health Minimal Dataset. These patient-reported outcomes are commonly recommended for use in clinical and research settings and measure a patient's perceptions of impairments in body structure and function, activity limitations, and participation restrictions (see Multimedia Appendix 1 for details) [37,42-44]. At the completion of the first visit, the patient completed a form for patient satisfaction with the goal-setting process and the clinician completed a form for patient level of engagement in goal setting. Audio recordings of patient visits were made. This was necessary to determine the reliability and feasibility of GAS in a clinical setting. Additionally, recording the encounter instead of having a research team member observing in the patient room allowed for a more natural interaction between clinician and patient.

Following the initial visit, therapy sessions proceeded as determined in each participant's physical therapy plan of care.

At the final physical therapy session (ie, at discharge), the Oswestry Disability Index and Numerical Pain Rating Scale were completed along with an abbreviated version of the National Institutes of Health Minimal Dataset, containing only the Patient-Reported Outcomes Measurement Information System items [37]. Additionally, the second part of the GAS-Back form, which regards goal achievement, was finalized, global perceived effect and patient satisfaction were measured, and the "collaboRATE" shared decision-making questionnaire was completed. We estimated these could be completed within 20 minutes. If a patient did not return for the last visit, the therapist called to follow up and the final questionnaire was completed over the phone or by using an online platform (Qualtrics).

Data Analysis

The success of GAS will be assessed based on feasibility, reliability, and validity. To consider the implementation of GAS in routine clinical practice, it must be administered in a timely (<15 minutes average examination time) and consistent (checklist adherence >80%) manner. The GAS process, feasibility, and fidelity will be evaluated by measuring the time (in minutes) to perform the GAS process, checklist adherence (as a percentage), and the nature of the goals identified (using the ICF domains) [45,46]. Interrater reliability of GAS scores will be assessed by examining the association between 2 independent physical therapist examiners using the Spearman rank correlation [33]. Criterion validity of GAS scores will be assessed by examining the association between GAS and Oswestry Disability Index scores [33,41]. The standardized response means (SRMs) for the Numerical Pain Rating Scale, Oswestry Disability Index, and GAS will be determined and compared to evaluate responsiveness [20]. A larger SRM indicates increased responsiveness of the measure to change. The SRM will be calculated as the ratio of change from pre- to posttest divided by the standard deviation of the change score [33]. A chi-square analysis will determine if different outcome measures show different proportions of patients determined to improve on that test by more than the minimally important change [33]. While a generally accepted and standardized definition of success for management of chronic LBP has not been established [42], our operational definition of success is as follows: if a patient shows changes that exceed the minimally important change or cutoff point for each individual measure, that patient's outcome will be defined as successful. Minimally important changes for this analysis will be set at 2 points for the Numerical Pain Rating Scale [44], 10 points for the Oswestry Disability Index [47], and 2 points for the global perceived effect rating [48]. Cutoff points will be set at ≥50 points for GAS [20]. A GAS score of 50 indicates that the expected outcome was achieved, while a score greater than 50 indicates performance exceeding the expected outcome [20]. As analyses using change scores have weaknesses, we will also apply alternative approaches, such as analysis of covariance and residual change score [49]. The measures proposed in this study, including GAS, have acceptable reliability, validity, and responsiveness (Multimedia Appendix 1).

Results

Anticipated Results

Overall, we anticipate that this study will demonstrate that GAS can be implemented in a consistent and timely manner by outpatient physical therapists, and that patients with chronic LBP will demonstrate clinically important changes that are also important to them. In phase 1 we anticipate that the inventory of goals will accurately represent the domains most important to patients with chronic LBP. This inventory of goals should allow for a series of prompts that can be used by physical therapists to expedite the GAS process in an outpatient setting. In phase 2 we anticipate finding that GAS will be feasible to implement in an outpatient setting. To demonstrate this feasibility, we anticipate that we will find that physical therapist training results in a reliable and timely use of GAS. Furthermore, we anticipate that GAS will demonstrate validity and responsiveness to change when compared to outcome measures commonly used in chronic LBP.

Study Timeline

Phase 1

Data collection was completed with the last focus group being held in June 2020. Preliminary data analysis was completed, and the information gleaned from this analysis was used to develop a series of prompts that were used to support the GAS process in phase 2.

Phase 2

Therapist training was completed in March 2021. Subject recruitment commenced following therapist training, and data collection began in March 2021. Data collection was completed in December 2021, and we expect data analysis to be completed by March 2022. This study is expected to conclude in late 2022.

Discussion

Principal Aims

This study aims to develop and test the feasibility of a novel application of GAS by physical therapists treating chronic LBP. It represents an important innovation because it facilitates patient-provider interaction and produces goals that encompass the activities and participation that are important to the patient. While goal setting is already part of the routine practice of physical therapists, the process is highly variable, with goals that are traditionally provider generated [8,50]. Knowledge of patient-initiated and patient-centered goals will enable health care providers to offer interventions that are more individualized and focused toward specific goals, leading to improved outcomes [18]. Furthermore, traditional, standardized, fixed-item patient-reported measures used for patients with chronic LBP (ie, the Numerical Pain Rating Scale and Oswestry Disability Index) may fail to measure constructs that are important to all patients, and therefore GAS may be better able to detect clinical changes that are meaningful to the patient. It is important to note that we are not recommending that we abandon current traditional, fixed-item, patient-reported outcome measures (eg, the Oswestry Disability Index and Numerical Pain Rating Scale).

Rather, we believe that GAS can provide complementary information that augments these more established measures.

Conclusion

GAS represents an opportunity for patient-centered care in the physical therapy management of chronic LBP. While GAS is not new, it has never been studied in real-world physical therapy for chronic LBP in a clinical setting, a type of practice that has unique time and productivity constraints. For GAS to be successfully implemented in this environment, we must demonstrate that clinicians can be trained efficiently and reliably, that GAS can be implemented in a clinical setting in under 15 minutes, and that GAS is able to detect clinically meaningful changes in patient outcomes.

Acknowledgments

The research reported in this publication was supported by the Agency for Healthcare Research and Quality of the National Institutes of Health under award HS025288-01A1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Reliability and validity.
[DOCX File , 41 KB - resprot_v11i3e32457_app1.docx]

References

1. Institute for Healthcare Improvement. The IHI Triple Aim. URL: http://www.ihi.org/Engage/Initiatives/TripleAim/Pages/default.aspx [accessed 2021-07-27]
2. Williams R, Stieg R. Validity and Therapeutic Efficacy of Individual Patient Goal Attainment Procedures in a Chronic Pain Treatment Center. Clin J Pain 1986 Jul 06;2(4):219-228 [FREE Full text] [doi: 10.1097/00002508-198612000-00003]
3. Levack WMM, Dean SG, Siegert RJ, McPherson KM. Purposes and mechanisms of goal planning in rehabilitation: the need for a critical distinction. Disabil Rehabil 2006 Jun 30;28(12):741-749. [doi: 10.1080/09638280500265961] [Medline: 16754571]
4. Cooper K, Smith BH, Hancock E. Patient-centredness in physiotherapy from the perspective of the chronic low back pain patient. Physiotherapy 2008 Sep;94(3):244-252. [doi: 10.1016/j.physio.2007.10.006]
5. Jaensch D, Baker N, Gordon S. Contemporaneous patient and health professional views of patient-centred care: a systematic review. Int J Qual Health Care 2019 Dec 31;31(10):G165-G173. [doi: 10.1093/intqhc/mzz118] [Medline: 31788686]
6. Hush J, Cameron K, Mackey M. Patient satisfaction with musculoskeletal physical therapy care: a systematic review. Phys Ther 2011 Jan;91(1):25-36. [doi: 10.2522/ptj.20100061] [Medline: 21071504]
7. Arnetz JE, Almin I, Bergström K, Franzén Y, Nilsson H. Active patient involvement in the establishment of physical therapy goals: Effects on treatment outcome and quality of care. Adv Physiother 2009 Jul 11;6(2):50-69. [doi: 10.1080/14038190310017147]
8. Baker S, Marshak H, Rice G, Zimmerman G. Patient participation in physical therapy goal setting. Phys Ther 2001;81(5):1118-1126. [doi: 10.1093/ptj/81.5.1118]
9. Hazard RG, Spratt KF, McDonough CM, Carayannopoulos AG, Olson CM, Reeves V, et al. The Impact of Personal Functional Goal Achievement on Patient Satisfaction With Progress One Year Following Completion of a Functional Restoration Program for Chronic Disabling Spinal Disorders. Spine 2009;34(25):2797-2802. [doi: 10.1097/brs.0b013e3181a9e640]
10. Toomey M, Nicholson D, Carswell A. The clinical utility of the Canadian Occupational Performance Measure. Can J Occup Ther 1995 Dec 22;62(5):242-249. [doi: 10.1177/000841749506200503] [Medline: 10152880]
11. Kiresuk TJ, Sherman RE. Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. Community Ment Health J 1968 Dec;4(6):443-453. [doi: 10.1007/bf01530764]
12. Melville LL, Baltic TA, Bettcher TW, Nelson DL. Patients' perspectives on the self-identified goals assessment. Am J Occup Ther 2002 Nov 01;56(6):650-659. [doi: 10.5014/ajot.56.6.650] [Medline: 12458857]
13. Stolee P, Rockwood K, Fox RA, Streiner DL. The use of goal attainment scaling in a geriatric care setting. J Am Geriatr Soc 1992 Jun 27;40(6):574-578. [doi: 10.1111/j.1532-5415.1992.tb02105.x] [Medline: 1587973]
14. Rockwood K, Stolee P, FoxP RA. Use of goal attainment scaling in measuring clinically important change in the frail elderly. J Clin Epidemiol 1993 Oct;46(10):1113-1118. [doi: 10.1016/0895-4356(93)90110-m]
15. Yip AM, Gorman MC, Stadnyk K, Mills WG, MacPherson KM, Rockwood K. A standardized menu for Goal Attainment Scaling in the care of frail elders. Gerontologist 1998 Dec 01;38(6):735-742. [doi: 10.1093/geront/38.6.735] [Medline: 9868853]

16. Bovend'Eerdt TJH, Botell RE, Wade DT. Writing SMART rehabilitation goals and achieving goal attainment scaling: a practical guide. Clin Rehabil 2009 Apr;23(4):352-361. [doi: 10.1177/0269215508101741] [Medline: 19237435]
17. Hurn J, Kneebone I, Cropley M. Goal setting as an outcome measure: A systematic review. Clin Rehabil 2006 Sep 27;20(9):756-772. [doi: 10.1177/0269215506070793] [Medline: 17005500]
18. Levack WM, Taylor K, Siegert RJ, Dean SG, McPherson KM, Weatherall M. Is goal planning in rehabilitation effective? A systematic review. Clin Rehabil 2006 Sep 27;20(9):739-755. [doi: 10.1177/0269215506070791] [Medline: 17005499]
19. Hazard RG, Spratt KF, McDonough CM, Olson CM, Ossen ES, Hartmann EM, et al. Patient-centered evaluation of outcomes from rehabilitation for chronic disabling spinal disorders: the impact of personal goal achievement on patient satisfaction. Spine J 2012 Dec;12(12):1132-1137 [FREE Full text] [doi: 10.1016/j.spinee.2012.09.003] [Medline: 23067862]
20. Mannion AF, Caporaso F, Pulkovski N, Sprott H. Goal attainment scaling as a measure of treatment success after physiotherapy for chronic low back pain. Rheumatology (Oxford) 2010 Sep 28;49(9):1734-1738. [doi: 10.1093/rheumatology/keq160] [Medline: 20511346]
21. Fisher K, Hardie RJ. Goal attainment scaling in evaluating a multidisciplinary pain management programme. Clin Rehabil 2002 Dec 01;16(8):871-877. [doi: 10.1191/0269215502cr554oa] [Medline: 12501949]
22. Mullis R, Hay E. Goal scaling for low back pain in primary care: development of a semi-structured interview incorporating minimal important change. J Eval Clin Pract 2010 Dec;16(6):1209-1214. [doi: 10.1111/j.1365-2753.2009.01296.x] [Medline: 20695953]
23. Chou R, Loeser JD, Owens DK, Rosenquist RW, Atlas SJ, Baisden J, American Pain Society Low Back Pain Guideline Panel. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain: an evidence-based clinical practice guideline from the American Pain Society. Spine (Phila Pa 1976) 2009 May 01;34(10):1066-1077. [doi: 10.1097/BRS.0b013e3181a1390d] [Medline: 19363457]
24. Wu A, March L, Zheng X, Huang J, Wang X, Zhao J, et al. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. Ann Transl Med 2020 Mar;8(6):299 [FREE Full text] [doi: 10.21037/atm.2020.02.175] [Medline: 32355743]
25. Marras WS. Occupational low back disorder causation and control. Ergonomics 2000 Jul 10;43(7):880-902. [doi: 10.1080/001401300409080] [Medline: 10929824]
26. Guo HR, Tanaka S, Halperin WE, Cameron LL. Back pain prevalence in US industry and estimates of lost workdays. Am J Public Health 1999 Jul;89(7):1029-1035. [doi: 10.2105/ajph.89.7.1029] [Medline: 10394311]
27. Brault M, Hootman J, Helmick C, Theis K, Armour B. Prevalence and Most Common Causes of Disability Among Adults -- United States, 2005. (Cover story). Article. MMWR Morb Mortal Wkly Rep 2009;58(16):421-426 [FREE Full text]
28. Haladay D, Swisher L, Hardwick D. Goal attainment scaling for patients with low back pain in rehabilitation: A systematic review. Health Sci Rep 2021 Sep 22;4(3):e378-e312 [FREE Full text] [doi: 10.1002/hsr2.378] [Medline: 34589616]
29. Carpenter C. Using qualitative focus groups to evaluate health programmes and service delivery. In: Hammell KW, Carpenter C, editors. Qualitative Research in Evidence-Based Rehabilitation. London, UK: Churchill Livingstone; 2004:51-64.
30. Hennink MM, Kaiser BN, Marconi VC. Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough? Qual Health Res 2017 Mar 26;27(4):591-608. [doi: 10.1177/1049732316665344] [Medline: 27670770]
31. Thomson S. Sample Size and Grounded Theory. JOAAG 2011;5(1):45-52 [FREE Full text]
32. Walker D, Myrick F. Grounded theory: an exploration of process and procedure. Qual Health Res 2006 Apr;16(4):547-559. [doi: 10.1177/1049732305285972] [Medline: 16513996]
33. Portney P, Watkins M. Foundations of Clinical Research: Applications to Practice 3rd ed. Philadelphia: FA Davis Company; 2015:569-658.
34. Oktay J. Grounded theory. In: Pocket guides to social work research methods. Oxford: Oxford University Press; 2012:52-103.
35. Browne RH. On the use of a pilot sample for sample size determination. Stat Med 1995 Sep 15;14(17):1933-1940. [doi: 10.1002/sim.4780141709] [Medline: 8532986]
36. Billingham SAM, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. BMC Med Res Methodol 2013;13:104 [FREE Full text] [doi: 10.1186/1471-2288-13-104] [Medline: 23961782]
37. Deyo RA, Dworkin SF, Amtmann D, Andersson G, Borenstein D, Carragee E, et al. Report of the NIH Task Force on Research Standards for Chronic Low Back Pain. Spine J 2014 Aug 01;14(8):1375-1391. [doi: 10.1016/j.spinee.2014.05.002] [Medline: 24950669]
38. Delitto A, George SZ, Van Dillen L, Whitman JM, Sowa G, Shekelle P, et al. Low Back Pain. J Orthop Sports Phys Ther 2012 Apr;42(4):A1-A57. [doi: 10.2519/jospt.2012.42.4.a1]
39. Chou R, Qaseem A, Snow V, Casey D, Cross JT, Shekelle P, Clinical Efficacy Assessment Subcommittee of the American College of Physicians, American College of Physicians, American Pain Society Low Back Pain Guidelines Panel. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med 2007 Oct 02;147(7):478-491 [FREE Full text] [doi: 10.7326/0003-4819-147-7-200710020-00006] [Medline: 17909209]
40. Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. The Lancet 2012 Feb;379(9814):482-491. [doi: 10.1016/s0140-6736(11)60610-7]

41. Turner-Stokes L. Goal attainment scaling (GAS) in rehabilitation: a practical guide. Clin Rehabil 2009 Apr 29;23(4):362-370. [doi: 10.1177/0269215508101742] [Medline: 19179355]
42. Chapman JR, Norvell DC, Hermsmeyer JT, Bransford RJ, DeVine J, McGirt MJ, et al. Evaluating Common Outcomes for Measuring Treatment Success for Chronic Low Back Pain. Spine 2011;36:S54-S68. [doi: 10.1097/brs.0b013e31822ef74d]
43. Hush JM, Kamper SJ, Stanton TR, Ostelo R, Refshauge KM. Standardized measurement of recovery from nonspecific back pain. Arch Phys Med Rehabil 2012 May;93(5):849-855. [doi: 10.1016/j.apmr.2011.11.035] [Medline: 22444028]
44. Ostelo RWJG, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting Change Scores for Pain and Functional Status in Low Back Pain. Spine 2008;33(1):90-94. [doi: 10.1097/brs.0b013e31815e3a10]
45. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol 2004 Sep;23(5):443-451. [doi: 10.1037/0278-6133.23.5.443] [Medline: 15367063]
46. Lewis V, Dell L, Matthews L. Evaluating the feasibility of Goal Attainment Scaling as a rehabilitation outcome measure for veterans. J Rehabil Med 2013 Apr;45(4):403-409 [FREE Full text] [doi: 10.2340/16501977-1131] [Medline: 23546308]
47. Downie A, Williams CM, Henschke N, Hancock MJ, Ostelo RWJG, de Vet HCW, et al. Red flags to screen for malignancy and fracture in patients with low back pain: systematic review. BMJ 2013 Dec 11;347:f7095 [FREE Full text] [doi: 10.1136/bmj.f7095] [Medline: 24335669]
48. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther 2009 Jul 18;17(3):163-170 [FREE Full text] [doi: 10.1179/jmt.2009.17.3.163] [Medline: 20046623]
49. Kisbu-Sakarya Y, MacKinnon DP, Aiken LS. A Monte Carlo Comparison Study of the Power of the Analysis of Covariance, Simple Difference, and Residual Change Scores in Testing Two-Wave Data. Educ Psychol Meas 2013 Feb 17;73(1):47-62 [FREE Full text] [doi: 10.1177/0013164412450574] [Medline: 26412869]
50. Scobbie L, Wyke S, Dixon D. Identifying and applying psychological theory to setting and achieving rehabilitation goals. Clin Rehabil 2009 Apr 17;23(4):321-333. [doi: 10.1177/0269215509102981] [Medline: 19293291]

Abbreviations

GAS: goal attainment scaling
ICF: International Classification of Functioning, Disability, and Health
LBP: low back pain
SMART: specific, measurable, achievable, realistic, relevant, and time-based
SRM: standardized response mean
USF: University of South Florida

Edited by T Leung; submitted 29.07.21; peer-reviewed by L Ortiz-Comino, S Rousset; comments to author 20.10.21; revised version received 22.11.21; accepted 12.01.22; published 07.03.22.

Please cite as:
Haladay D, Ditwiler RE, Klein AB, Miro R, Lazinski M, Swisher LL, Beckstead J, Wolfson J, Hardwick D
Goal Attainment Scaling in Outpatient Physical Therapy for Chronic Low Back Pain: Protocol for a Mixed Methods Study
JMIR Res Protoc 2022;11(3):e32457
URL: https://www.researchprotocols.org/2022/3/e32457
doi:10.2196/32457
PMID:35254282

Protocol

Effects of Integrating Family Planning With Maternal, Newborn, and Child Health Services on Uptake of Voluntary Modern Contraceptive Methods in Rural Pakistan: Protocol for a Quasi-experimental Study

Zahid Ali Memon[1,2], MPH; Sophie Reale[3], PhD; Wardah Ahmed[1], MSPH; Rachael Spencer[4], PhD; Talib Hussain Lashari[5], MHPM; Zulfiqar Bhutta[1], PhD; Hora Soltani[2], PhD

[1]Centre of Excellence in Women and Child Health, Aga Khan University, Karachi, Pakistan

[2]Health Research Institute, Sheffield Hallam University, Sheffield, United Kingdom

[3]College of Health, Wellbeing and Life Sciences, Sheffield Hallam University, Sheffield, United Kingdom

[4]Department of Nursing and Midwifery, Sheffield Hallam University, Sheffield, United Kingdom

[5]Population Welfare Department, Karachi, Pakistan

Corresponding Author:
Zahid Ali Memon, MPH
Centre of Excellence in Women and Child Health
Aga Khan University
Stadium Road
Karachi, 74800
Pakistan
Phone: 92 3085550859
Email: zahid.memon@aku.edu

Abstract

Background: The uptake of modern contraceptive methods (MCMs) remains low, with 25% of women reporting their use in Pakistan. The overarching interventions covering service delivery platforms at facility and community levels necessitate the integration of family planning (FP) with maternal, newborn, and child health (MNCH) services.

Objective: The main aim of this study is to evaluate the impact of an integrated FP-MNCH service delivery model to increase coverage of MCMs in rural Pakistan. Moreover, we aim to measure the level of effectiveness of interventions regarding the uptake of MCMs.

Methods: A quasi-experimental, sequential, mixed methods study design with pre- and postevaluation will be adopted to evaluate the impact of integration of FP with MNCH services. The interventions include the following: (1) capacity strengthening of health care providers, including technical trainings; training in counseling of women who attend immunization centers, antenatal care (ANC) clinics, and postnatal care (PNC) clinics; and provision of job aids; (2) counseling of women and girls attending ANC, PNC, and pediatric clinics; (3) ensuring sustained provision of supplies and commodities; (4) community engagement, including establishing adolescent-friendly spaces; and (5) use of District Health Information System data in decision-making. Descriptive statistics will be used to estimate prevalence (ie, proportions) and frequencies of outcome indicators. A univariate difference-in-difference analytical approach will be used to estimate the effect of the interventions. In addition, a Blinder-Oaxaca decomposition analysis will be conducted to identify and quantify determinants of the modern contraceptive prevalence rate.

Results: The intervention phase began in July 2021 and will run until June 2022. The impact assessment will be conducted from July to September 2022.

Conclusions: This project will evaluate the impact of integrating FP with MNCH services. Furthermore, this study will identify the drivers and barriers in uptake of MCMs and will simultaneously help in modifying the interventional strategies that can be scaled up through existing service delivery platforms within the public and private sectors, according to the local sociocultural and health system context.

Trial Registration: ClinicalTrials.gov NCT05045599; https://clinicaltrials.gov/ct2/show/NCT05045599

International Registered Report Identifier (IRRID): DERR1-10.2196/35291

(JMIR Res Protoc 2022;11(3):e35291) doi:10.2196/35291

KEYWORDS

family planning; integrated health services; contraceptive prevalence rate; modern contraceptive prevalence rate; modern contraceptive method; rural Pakistan

Introduction

Background

The scaling up of family planning (FP) programs has yielded a multitude of benefits in terms of improving health outcomes and paving opportunities for poverty reduction and women's empowerment. FP programs have effectively contributed to a significant reduction of 32% of all maternal deaths and 10% of all child deaths. Moreover, these programs enhance women's empowerment and gender equity, and, within that, they are aimed at achieving a higher level of universal female primary education. Over the past four decades, FP programs have led to a globally increased contraceptive prevalence rate (CPR), from 10% to 60%, and a substantial reduction in fertility rate from 6 to 3 children per woman in low- and middle-income countries [1,2].

Modern contraceptive methods (MCMs), such as condoms, contraceptive pills, injectables, and intrauterine devices, serve as significant measures of fertility control [1]. There is a strong body of evidence suggesting that providing a woman with opportunities to opt for the MCM of her choice through a continuum of care is viable and cost-effective, not only at the individual level but also at the broader health system level [2,3].

In Pakistan, FP programs started in the mid-1960s [1]. Later, a Lady Health Worker (LHW) outreach program was initiated in 1994, which primarily focused on FP services at the national level. Pakistan has been a signatory on various international commitments to improve access to reproductive health services, including the following: (1) the Millennium Development Goals (2000), aimed at increasing the CPR from 12% in 1990 to 55% in 2015 [4]; (2) the London Summit on Family Planning 2012, aimed at raising the CPR to 55% by 2020; and (3) the Sustainable Development Goals (2015), aimed at achieving universal access to reproductive health services by 2030. Pakistan has made considerable progress over time, yet these targets have not been fully reached [5].

Generally, the provincial government's Population Welfare Department and the Department of Health are both responsible for translating the government's vision of health and FP. However, the Department of Health has lagged behind in taking ownership of the FP agenda to fully optimize the resources to achieve the objectives [6]. The government's FP2020 agenda invigorated efforts to accelerate progress in achieving higher use of MCMs. Political commitment at the national level and at the Sindh province level culminated in appointing the same minister to the Department of Health and the Population Welfare Department to ensure functional integration in both departments. Moreover, the FP2030 Secretariat has been created to bring all stakeholders onto a single platform and provide facilitation for nongovernmental organizations (NGOs) and private sector entities that are working in the FP arena.

Several collaborative initiatives have contributed to increased MCM uptake in Pakistan, as demonstrated by a nearly three-fold increase in the CPR from 13% to 34% between 1990 and 2018 (0.5% annual increment) [7-9]. However, there still remains a large unmet need for FP, toward which progress has been slow and in some cases stagnant.

A review of the existing data report low use of MCMs at 25% among women aged 15 to 49 years and 7% among adolescents aged 15 to 19 years in Pakistan [7]. This indicates a high unmet need of FP, specifically among currently married women (17%). For example, out of all pregnancies in Pakistan, 46% are unintended and unsafe, and the abortion rate remains high at 50 per 1000 women [7]. It is noteworthy that 78% of nonusers of MCMs are women aged 15 to 49 years who have never had discussions regarding FP methods with any health care provider at the community or facility levels [7]. Further, of those who give birth, many do not receive (1) standard antenatal care (ANC), postnatal care (PNC), or quality services at the time of delivery or (2) appropriate management for mother and newborn complications [10].

This is the result of various barriers that have evolved within the broader political, social, and cultural context of Pakistan. A deeper understanding of these barriers and determinants is required to identify gaps within the health system and expand the coverage of FP and maternal, newborn, and child health (MNCH) services [1,11]. From the demand side, literature suggests that women's empowerment, gender equality, and religious obstacles are contributing factors that heavily influence the uptake of contraceptive methods [12]. However, a further exploration is required at the local level in terms of community perspectives of, resistance to, and attitudes toward FP and contraceptive uptake. Within the country, people living in the lower wealth quintiles (below 25%) and residing in rural areas have a lower literacy level, with young women experiencing a great unmet need for FP services and products [13]. This has translated into poor acknowledgement and adoption of practices relating to healthy timing and spacing of pregnancies and, consequently, higher fertility rates [14,15]. Preference for sons and a lack of husbands' participation also count as prominent factors in the context of Pakistan that present key barriers to the uptake of FP methods [16,17]. Studies further report that the involvement of supportive male family members has effectively increased the rate of voluntary uptake of contraceptive methods [17]. Given that the bulk of Pakistan's population is comprised of youth, the need for adolescent-focused interventions to evaluate and improve reproductive health outcomes at scale is urgently needed [8,18].

On the supply side, the literature reports lack of access to information; lack of access to health care service delivery points, especially in rural communities; and fear of side effects as key barriers to MCM uptake. For those who do adopt an MCM, method discontinuation due to poor quality of services provided,

attitudes of health care providers, and cost-bearing during treatment of side effects [9,13,18,19] are key factors in low contraceptive prevalence. Considering FP as the major factor in improving health indicators throughout the course of women's and children's lives [20], thereby building trust in government services, making needs-based alignment of FP programs in the country would substantially affect the resistance to opting for contraception [21].

Given these barriers, designing and evaluating integrated FP programs that are based on needs and delivered through existing service delivery platforms have the potential to improve the health and well-being of women and children in Pakistan. The integration of FP with immunization in Malawi resulted in a 14% increase in the uptake of FP methods. Parents experienced greater feasibility marked by reductions in time lost and transport costs as well as access to greater awareness and knowledge while accessing services on the same day. Similarly, health care providers expressed that this integration enhanced their skills, knowledge, and competence [20]. Similar results were found in Liberia and Rwanda, with an emphasis on maintaining privacy for couples in public health settings. The integration of the immunization schedule, with extended postpartum FP, resulted in an increase in uptake [22,23]. In order for the successful and equity-based implementation of integrated services, it was reported that following standard protocols for provision of care and training of health care providers at facility and community levels was necessary [24].

Moreover, a recent study conducted during the COVID-19 pandemic in Ethiopia examined the integration of FP with maternal health care services, including ANC, delivery, PNC, and immunization. The results showed a 6% increase in the uptake of integrated FP services, which is significant in a low-resource setting, given the widespread effect of the global pandemic [25].

The literature has also shown that interventions focusing on community outreach programs and interpersonal communications increase social acceptance of FP methods [26]. However, home-based counseling alone is not sufficient for the uptake and continuation of FP methods [27]; developing linkages with health facilities, maintaining privacy at a health facility, and delivering services in a socioculturally sensitive manner is also important.

In line with this, efforts have been made that involve facility- and community-level health care providers for the provision of MNCH services as the primary mandate of the National MNCH Program. However, there are still deficiencies at the interfacility and intrafacility levels, for example, (1) a lack of coordination among departments, such as pediatrics and gynecology and obstetrics; (2) a lack of management-level coordination with frontline health workers; (3) a lack of equipment and logistics management manifested as an imbalance in demand and supply; and (4) a lack of overall governing bodies [6,28]. Thus, overarching interventions covering service delivery platforms at facility and community levels necessitate the integration and scaling up of FP and MNCH services.

Implementation Context

Pakistan is comprised of four provinces (ie, Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan), two autonomous territories (ie, Gilgit-Baltistan and Azad Jammu Kashmir), and a federal territory (ie, Islamabad); these collectively encompass 150 districts [29]. This study will be conducted in two districts of Sindh province (Figure 1). Sindh has the highest rural-urban difference in fertility rates: 4.7 and 2.9 per woman, respectively. Median age at first birth is 23 years, with only 18% of this age group receiving any contraceptive method. Overall, modern CPR (mCPR) in Sindh is lower (24% vs 26%) and unmet needs are higher (22% vs 17%) as compared to national-level statistics [30].

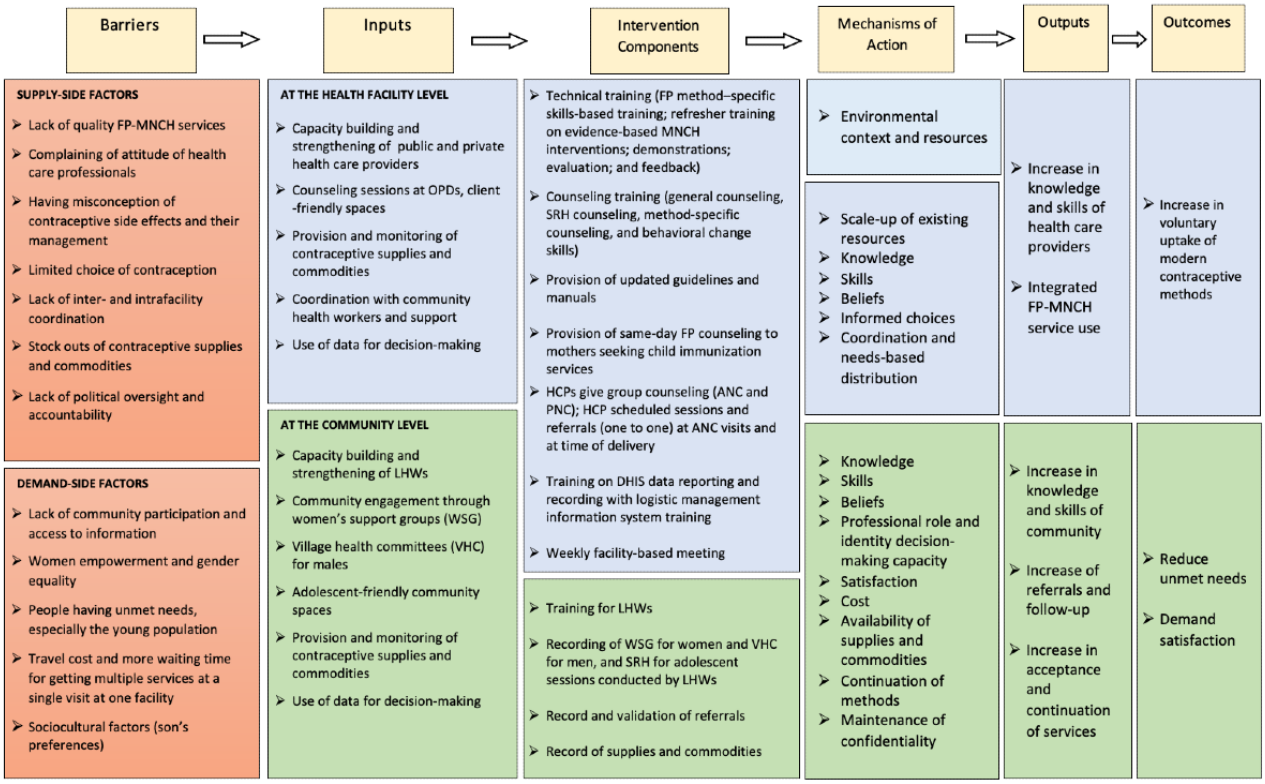
The public health system consists of a three-tiered health delivery system comprised of domiciliary and outreach services rolled out through LHWs and primary- and secondary-level health care facilities in each district [2]. Each LHW has a catchment population of approximately 1000 to 1500 people. They provide educational, preventive, and promotive services as well as some aspects of curative services in the community. They provide MNCH and FP services to eligible families.

Figure 1. Map of Pakistan showing the districts of Badin (control) and Matiari (intervention) in Sindh province.

Theoretical Framework

The theoretical underpinning of behavior change will be based on the Theoretical Domains Framework (TDF), version 2.0. The TDF will be applied to provide an in-depth exploration and understanding of factors regarding the demand and supply side and their interaction with and influences on FP uptake [31,32]. This project aims to implement a complex intervention (Figure 2) within health facilities and their catchment communities. This complex intervention includes a series of strategies involving community engagement by extensive community mobilization, availability of trained staff, and sustainable supply of commodities with the required recording and reporting system. Continuous process monitoring and quality assurance will help to replicate the success and address possible barriers during implementation of the intervention. The mechanism of action built on the TDF adopts domains and constructs, including knowledge, skills, beliefs, and intentions. Furthermore, the TDF provides a detailed understanding of complex behavior; thus, it will be used to evaluate the impact of complex interventions and strategies.

Figure 2. Theoretical Framework: integrated FP-MNCH services. ANC: antenatal care; DHIS: District Health Information System; FP: family planning; HCP: health care provider; LHW: Lady Health Worker; MNCH: maternal, newborn, and child health; OPD: outpatient department; PNC: postnatal care; SRH: sexual and reproductive health; VHC: village health committee; WSG: women support group.



Research Question

The research question that will be answered by this study is as follows: What is the impact of integrating FP with MNCH services on the uptake of voluntary MCMs in a rural district of Sindh province, Pakistan?

Aim and Objectives

The aim of this study will be to evaluate the impact of an integrated FP-MNCH service delivery model to increase coverage of MCMs in rural Pakistan.

The objectives of this study are as follows:

- To gain an understanding of the cultural and health service delivery contexts to inform a socioculturally appropriate and acceptable intervention package scalable in rural Pakistan
- To implement the intervention package at health facilities and outreach communities through existing public and private sector resources
- To measure the impact and level of effectiveness of interventions on the uptake MCMs
- To identify and quantify the drivers of improved uptake of voluntary methods of FP, especially MCMs.

Methods

Study Design

A quasi-experimental, sequential, mixed methods study design with pre- and postevaluation is proposed to assess the impact of implementing an integrated delivery model on the uptake of MCMs in a rural district of Sindh province, Pakistan.

Ethics Approval

Ethical clearance for this project was sought from the Ethical Review Committee of the Aga Khan University on June 26, 2020. The study protocol was approved by the National Bioethics Committee, Pakistan (ERC number 2021-3606-19065). The baseline assessment was completed in January 2021. The inception period included hiring and deployment of study staff and their training in the standard operating procedures of the project.

Qualitative Component

Overview

The overall objective of the qualitative component is to understand the barriers that married women and girls who are at risk of an unwanted pregnancy face in accessing and using voluntary FP. The information will be used to inform the integrated intervention design and delivery platform. The qualitative component will include focus group discussions (FGDs) with married women of reproductive age (MWRA; 18-49 years) and adolescent girls and boys (15-19 years) to understand behaviors and community-level hindrances to acceptance and use of FP methods. Due to the sensitivity of the topics, separate FGDs will be conducted with male and female participants. In-depth interviews will also be conducted with service providers and health facility managers to understand facility-level and supply-side barriers to uptake of modern FP

methods and services (see Multimedia Appendix 1 for sample sizes of FGDs and in-depth interviews).

A purposive sampling technique will be used. We aim to conduct 10 in-depth interviews and two FGDs, with 5 to 6 participants in each group. However, this will depend on reaching the theoretical point of saturation. A thematic approach will be used to analyze the data. The FGD guide for data collection will be based on the TDF.

Analysis of the Qualitative Component

Analysis will be done using NVivo software (version 11; QSR International). The TDF will be the guiding framework for coding and thematic analysis.

Systematic Review

A systematic review of the existing research will be conducted to identify effective FP interventions and strategies that led to an increase in the uptake of MCMs. The research question will look at the impact of effective interventions and strategies on improving the uptake of MCMs in South Asian countries. Various databases (PubMed, Cochrane Database of Systematic Reviews, EBSCO CINAHL, Web of Science, ProQuest Dissertations & Theses, etc) will be used to collect articles published from January 2000 to June 2021. Covidence, a systematic review management program, will be employed to facilitate collaboration across the team. A meta-analysis will be used to synthesize the data from relevant studies into a single quantitative pooled estimate. The pooled estimate will be the outcome of the meta-analysis and will be explained by a forest plot. The meta-analysis will be done using RevMan (Review Manager) software (version 5.4.1; The Cochrane Collaboration). For categorical variables, odds ratios with 95% CIs will be reported; for continuous outcomes, mean differences or standard mean differences with 95% CIs will be reported.

Quantitative Component

Overview

A nonrandomized, quasi-experimental study design with intervention and control arms is proposed to assess the impact of implementing a complex delivery model of interventions on the voluntary uptake of MCMs in two rural districts of Sindh. Two sequential face-to-face household surveys will be conducted before and after the 12-month implementation of the interventions [33]. The surveys will document the changes over time in mCPR alongside other outcomes of interest, such as unmet needs and demands satisfied.

Selection of Control District

We used the propensity score matching technique to select the control district (Table 1). This procedure estimates the probability and propensity that a study unit that has not received the intervention is similar at baseline to another unit from the intervention group, based on a set of key characteristics. As such, it reduces the problem of comparison across large numbers of key variables [34]. In the context of our study, this matching was done to select Badin as the control district to compare with the intervention district of Matiari (Figure 1). The variables used in generating propensity score matching included CPR, mCPR, unmet needs, ANC by skilled providers, deliveries by skilled birth attendants, the Human Development Index, and percentage of full immunization coverage. The district-level data of these indicators were taken from the Sindh Multiple Indicator Cluster Survey 2018 for all 23 rural districts (Multimedia Appendix 2).

Table 1. Selection of control district using propensity score matching.

District	Unmet needs, %	CPR[a], %	Modern CPR, %	Fully immunized, %	Deliveries by skilled birth attendants, %	Antenatal care, %	Propensity score
Badin (control district)	17.9	28.1	28.1	45.6	61.9	82.0	0.36
Matiari (intervention district)	25.8	32.4	30.6	68.0	65.5	85.7	0.31

[a]CPR: contraceptive prevalence rate.

Sample Size Estimation

With 95% CI and 80% power, a sample size of 880 MWRA, 15 to 49 years, per arm is required to estimate an increase from 28.9% to 36.9% (ie, 8% increase) in the proportion of MCM uptake. The assumed design effect of 1.5 and a 7% nonresponse rate is accounted for in the sample size calculations [35].

Sampling and Recruitment Strategy

A two-stage random-sampling strategy will be employed to select the clusters and households with respondents (ie, MWRA, 15-49 years). A higher response rate has been projected, as the survey will be conducted at the field level on a face-to-face basis. The face-to-face method has been accredited in the literature as an effective way of eliciting a higher rate of response from target audiences [33]. Sampling will consider each cluster level as a stratum; each cluster will be comprised of 100 to 150 households in the catchment area of the health facility. Each cluster will be a primary sampling unit. At the first stage, clusters will be randomly selected from the sampling frame of each district. Each household will be a secondary sampling unit. At the second stage, 20 eligible households will be selected using systematic random sampling to reach the calculated sample size (ie, 880 MWRA, 15-49 years, per arm).

Outcomes

The primary outcome is the mCPR. The operational definition of the primary outcome can be found in Multimedia Appendix 3.

The secondary outcomes are as follows (see Multimedia Appendix 3 for their operational definitions):

- Unmet needs
- Demands satisfied
- Proportion of women showing positive attitudes toward FP visits where clients received integrated FP-MNCH services

- Proportion of referral completers who accepted an FP method
- Proportion of follow-up visits where clients received integrated FP-MNCH services
- Percentage of women who were satisfied with the services provided
- Percentage of staff scoring 70% or above on a knowledge assessment test using validated tools, at the facility and community levels
- Number of unwanted pregnancies and births
- Proportion of married women who used ANC and PNC services
- Proportion of institutional deliveries by skilled birth attendants.

Statistical Analysis of the Quantitative Component

The data will be collected using an Android app and will be uploaded in real time to the Aga Khan University server. Data cleaning will be performed by identifying errors and missing entries under the supervision of the senior data manager. Descriptive statistics will be used to estimate prevalence (ie, proportions) and frequencies of outcome indicators. A univariate difference-in-difference analytical approach will be used to estimate the effect of the interventions. In addition, a Blinder-Oaxaca decomposition analysis will be conducted to identify and quantify determinants of mCPR. Analyses will be done using Stata Statistical Software (version 16; StataCorp LLC). Weighted analysis will be performed by target age group for each cluster.

Intervention Sites

The intervention will be implemented within existing health systems, including public and private health facilities, and among community-level health workers. The intervention duration will range from 12 to 18 months. Six public health facilities working under the Department of Health, 10 private health facilities, and all LHWs working in the catchment areas and along the referral pathway will be included in the Matiari district in Sindh (Table 2). A similar number of health facilities and outreach LHWs will be selected as control facilities and LHWs from the Badin district.

Table 2. Health system characteristics of the study districts.

District	Type of health facility, n			Lady Health Workers, n	Target population, n
	Secondary	Primary	Private		
Matiari (intervention district)	3	3	10	144	269,304
Badin (control district)	3	3	10	170	604,888

Key Strategies to Implement Integrated FP-MNCH Interventions

An integrated FP-MNCH service delivery model will be implemented. FP will be integrated into the MNCH service delivery platforms of the intervention district's public sector health care facilities, as follows:

1. ANC clinics: counseling on first ANC visit, counseling on follow-up ANC, and recruitment for postpartum FP uptake.
2. Labor and delivery rooms: counseling during admission to the health care facility. The focus of the counseling sessions will be on provision of information about the benefits of spacing of pregnancies for the mother and newborn, and information on different MCMs available and when to use them.
3. Pediatric outpatient departments: counseling of women of reproductive age and caregivers on the importance of FP and referral to the FP clinics.

In addition, LHWs will be encouraged to integrate FP into their routine MNCH activities at the household and community levels.

The proposed interventions and strategies based on the existing evidence are listed below; the strategies will be refined and finalized based on the findings of the qualitative research:

1. Capacity strengthening of health care providers serving at government secondary- and primary-level hospitals and LHWs serving in the facility catchments, as follows: (a) technical trainings; (b) training for counseling of women attending ANC, PNC, and pediatric clinics; and (c) provision of job aids.
2. Counseling of women and girls attending ANC, PNC, and pediatric clinics regarding the importance of FP and referral to the FP clinics.
3. Ensuring the sustained provision of supplies and commodities.
4. Community engagement, including adolescent and youth populations.
5. Use of District Health Information System data in decision-making.

Results

The study was registered at ClinicalTrials.gov (NCT05045599). The baseline assessment was completed in January 2021. The inception period included hiring and deployment of study staff and their training in the standard operating procedures of the project. The intervention phase began in July 2021 and will run until June 2022. The impact assessment will be conducted from July to September 2022. Details of the study timeline are outlined in Multimedia Appendix 4.

Discussion

Despite the knowledge and scalability of evidence-based interventions and commitment from the government and other NGOs to improve reproductive health, Pakistan still lags behind its regional counterparts in achieving accelerated progress in FP indicators. Addressing this lag requires exploring

context-specific interventions and strategies to ensure affordable access to, and availability of, quality FP products, information, and services within existing service delivery platforms. This provides a prime opportunity to develop scalable and sustainable models that can reach and benefit the most vulnerable and marginalized populations across the country.

This project, through its qualitative components, seeks to deconstruct the sociocultural sensitivities and context of FP service delivery and uptake in Pakistan and further attempts to understand the barriers and facilitators to the voluntary uptake of FP services and improved health outcomes for women and young girls in this regard. Studies in India and Nigeria report a strong negative correlation between communities that harbor a conservative cultural and religious belief system and low demand and acceptance of contraceptive services and uptake [36]. These belief systems present a spectrum of factors, including women's restrictions on movement, religious reservations regarding the use of contraceptives, and denial of women's rights to decision-making and choices related to their health and well-being [36,37]. Many of these factors are present in the context of Pakistan and serve as barriers to FP access and uptake, specifically for women residing within communities in remote rural areas [21].

Therefore, identifying the local grassroots context and synchronizing it with FP intervention design and implementation strategies holds the key to addressing barriers to voluntary FP uptake faced by women in their day-to-day lives and achieving meaningful and effective coverage of FP interventions. Furthermore, studies conducted in Nepal and Bangladesh that reflect an FP demographic and landscape similar to that of Pakistan emphasize the efficiencies of adopting an integrated approach toward improving FP service delivery, quality, and uptake [30,38]. Strategically integrating FP services with services provided along the MNCH continuum of care has yielded significant impact in terms of increasing the prevalence of MCMs and their sustained use [39]. More specifically, the delivery of FP services to women at crucial points in the delivery of ANC, postpartum care, and PNC is highlighted in several studies in South and sub-Saharan Africa as a highly effective and feasible strategy to motivate new users of both short-term and long-term contraceptive methods [40,41].

Moreover, a plethora of quasi-experimental studies conducted within low- and middle-income countries, including Pakistan, concluded that FP intervention designs that cater to strengthening FP service delivery at both the facility and community levels are best equipped for success in the South Asian context [42-44]. Factors identified in the literature at the facility level include the following: capacity building of FP health care providers; strengthening supply chain and stock management of essential FP commodities and equipment; extending the scope of technical FP services and products, including surgical application of MCMs; and providing counseling services to women and family members at the facility level [45]. At the community level, the mobilization of community-based workers to generate awareness and demand for FP uptake is instituted through visiting households; educating and promoting FP among various community forums and groups; counseling women, men, couples, and families with regard to the importance and uptake of FP and MNCH; and establishing improved systems for referral and linkages [43,46]. This study, through its qualitative and quantitative components, will attempt to gauge both facility- and community-level factors that can be strengthened as part of the strategic implementation of the intervention to increase uptake of MCMs.

Therefore, this project provides a prime opportunity to generate evidence on effective interventions and strategies that are contextually relevant and sensitive. These strategies would improve access to FP information and services for the women of reproductive age and their family members who come into contact with the health system. Such care-seeking presents a premise of missed opportunities that can otherwise be effectively used by simply integrating FP information and services and making them available at the health facility level, in those facilities that provide MNCH services, and at the community level through LHWs and community worker mobilization.

Thus, the learnings from this study will provide a credible and robust foundation on which to design an effective intervention model that can eventually be scaled up to other districts, regions, and provinces across Pakistan. The lessons learned and the best practices emanating from this model will pave the way for evidence-informed FP program and policy making at the broader national level.

Acknowledgments

This study is funded by Global Affairs Canada through the United Nations Population Fund. We also acknowledge the support and close collaboration of the FP2030 Secretariat, the Population Welfare Department, and the Department of Health of the Government of Sindh regarding the implementation of the project.

Authors' Contributions

ZB led the funding acquisition for the main Sehatmand Khandan (healthy families) project. ZAM conceptualized the study design with guidance from HS and ZB. THL provided input on public sector service delivery platforms. SR provided input on the theoretical framework. ZAM drafted the first version of the manuscript and incorporated feedback from all coauthors. ZAM, SR, RS, WA, THL, and ZB reviewed subsequent drafts of the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample sizes of focus group discussions and in-depth interviews.
[DOCX File , 16 KB - resprot_v11i3e35291_app1.docx]

Multimedia Appendix 2
Propensity score matching for selection of control district.
[DOCX File , 18 KB - resprot_v11i3e35291_app2.docx]

Multimedia Appendix 3
Operational definitions of the primary and secondary outcomes.
[DOCX File , 20 KB - resprot_v11i3e35291_app3.docx]

Multimedia Appendix 4
Project timeline.
[DOCX File , 32 KB - resprot_v11i3e35291_app4.docx]

References

1. FP2020: Women at the Center 2018-2019. Progress Report. Washington, DC: Family Planning 2020; 2019. URL: http://progress.familyplanning2020.org/sites/default/files/FP2020_2019Report_FINAL_110819.pdf [accessed 2021-03-31]
2. Taylor HC, Berelson B. Comprehensive family planning based on maternal-child health services: A feasibility study for a world program. Stud Fam Plann 1971 Feb;2(2):1-54. [Medline: 5113345]
3. Bhutta ZA, Hafeez A, Rizvi A, Ali N, Khan A, Ahmad F, et al. Reproductive, maternal, newborn, and child health in Pakistan: Challenges and opportunities. Lancet 2013 Jun 22;381(9884):2207-2218. [doi: 10.1016/S0140-6736(12)61999-0] [Medline: 23684261]
4. Kunii C. How integration of family planning and maternal and child health should be initiated and developed. JOICFP Rev 1984(7):15-19. [Medline: 12313056]
5. Cates W, Abdool Karim Q, El-Sadr W, Haffner DW, Kalema-Zikusoka G, Rogo K, et al. Global development. Family planning and the Millennium Development Goals. Science 2010 Sep 24;329(5999):1603. [doi: 10.1126/science.1197080] [Medline: 20847233]
6. Khan AA, Khan A, Javed W, Hamza HB, Orakzai M, Ansari A, et al. Family planning in Pakistan: Applying what we have learned. J Pak Med Assoc 2013 Apr;63(4 Suppl 3):S3-S10. [Medline: 24386723]
7. National Institute of Population Studies (NIPS) [Pakistan] and ICF. Pakistan Demographic and Health Survey 2017-18. Islamabad, Pakistan, and Rockville, MD: NIPS and ICF; 2019. URL: http://dhsprogram.com/pubs/pdf/FR354/FR354.pdf [accessed 2021-03-31]
8. Ewerling F, Victora CG, Raj A, Coll CVN, Hellwig F, Barros AJD. Demand for family planning satisfied with modern methods among sexually active women in low- and middle-income countries: Who is lagging behind? Reprod Health 2018 Mar 06;15(1):42 [FREE Full text] [doi: 10.1186/s12978-018-0483-x] [Medline: 29510682]
9. Population Council. Landscape Analysis of the Family Planning Situation in Pakistan: Brief Summary of Findings. Islamabad, Pakistan: Population Council; 2016 Sep. URL: https://www.popcouncil.org/uploads/pdfs/2016RH_LandscapeAnalysisFP-Pakistan_summary.pdf [accessed 2022-02-16]
10. Agha S, Fitzgerald L, Fareed A, Rajbhandari P, Rahim S, Shahid F, et al. Quality of labor and birth care in Sindh Province, Pakistan: Findings from direct observations at health facilities. PLoS One 2019;14(10):e0223701 [FREE Full text] [doi: 10.1371/journal.pone.0223701] [Medline: 31622382]
11. de Jongh TE, Gurol-Urganci I, Allen E, Jiayue Zhu N, Atun R. Barriers and enablers to integrating maternal and child health services to antenatal care in low and middle income countries. BJOG 2016 Mar;123(4):549-557 [FREE Full text] [doi: 10.1111/1471-0528.13898] [Medline: 26861695]
12. Hameed W, Azmat S, Ali M, Sheikh M, Abbas G, Temmerman M, et al. Women's empowerment and contraceptive use: The role of independent versus couples' decision-making, from a lower middle income country perspective. PLoS One 2014;9(8):e104633 [FREE Full text] [doi: 10.1371/journal.pone.0104633] [Medline: 25119727]
13. Sathar ZA. Family planning: A missing priority in Pakistan's health sector? Lancet 2013 Jun 22;381(9884):2140-2141. [doi: 10.1016/S0140-6736(13)60763-1] [Medline: 23684258]
14. Tappis H, Kazi A, Hameed W, Dahar Z, Ali A, Agha S. The role of quality health services and discussion about birth spacing in postpartum contraceptive use in Sindh, Pakistan: A multilevel analysis. PLoS One 2015;10(10):e0139628 [FREE Full text] [doi: 10.1371/journal.pone.0139628] [Medline: 26485524]

15. Mahmood A. Birth Spacing and Family Planning Uptake in Pakistan: Evidence From FALAH. Islamabad, Pakistan: Population Council; 2012. URL: https://knowledgecommons.popcouncil.org/departments_sbsr-rh/122/ [accessed 2021-03-31]
16. Javed R, Mughal M. Preference for boys and length of birth intervals in Pakistan. Res Econ 2020 Jun;74(2):140-152. [doi: 10.1016/j.rie.2020.04.001]
17. Kamran I, Tasneem Z, Parveen T, Niazi RM. Family Planning Through the Lens of Men: Readiness, Preferences, and Challenges. Washington, DC: Population Council, The Evidence Project; 2015. URL: https://knowledgecommons.popcouncil.org/departments_sbsr-rh/878/ [accessed 2021-03-31]
18. Zaidi B, Hussain S. Reasons for Low Modern Contraceptive Use—Insights From Pakistan and Neighboring Countries. Literature Review. Islamabad, Pakistan: Population Council; 2015. URL: https://knowledgecommons.popcouncil.org/departments_sbsr-rh/908/ [accessed 2021-03-31]
19. Casterline JB, Sathar ZA, ul Haque M. Obstacles to contraceptive use in Pakistan: A study in Punjab. Stud Fam Plann 2001 Jun;32(2):95-110. [doi: 10.1111/j.1728-4465.2001.00095.x] [Medline: 11449867]
20. Cooper CM, Wille J, Shire S, Makoko S, Tsega A, Schuster A, et al. Integrated family planning and immunization service delivery at health facility and community sites in Dowa and Ntchisi Districts of Malawi: A mixed methods process evaluation. Int J Environ Res Public Health 2020 Jun 24;17(12):4530 [FREE Full text] [doi: 10.3390/ijerph17124530] [Medline: 32599688]
21. Ataullahjan A, Mumtaz Z, Vallianatos H. Family planning in Pakistan: A site of resistance. Soc Sci Med 2019 Jun;230:158-165. [doi: 10.1016/j.socscimed.2019.04.021] [Medline: 31015107]
22. Nelson AR, Cooper CM, Kamara S, Taylor ND, Zikeh T, Kanneh-Kesselly C, et al. Operationalizing integrated immunization and family planning services in rural Liberia: Lessons learned from evaluating service quality and utilization. Glob Health Sci Pract 2019 Sep 26;7(3):418-434. [doi: 10.9745/ghsp-d-19-00012]
23. Dulli LS, Eichleay M, Rademacher K, Sortijas S, Nsengiyumva T. Meeting postpartum women's family planning needs through integrated family planning and immunization services: Results of a cluster-randomized controlled trial in Rwanda. Glob Health Sci Pract 2016 Feb 22;4(1):73-86. [doi: 10.9745/ghsp-d-15-00291]
24. Ringheim K, Gribble J, Foreman M. Integrating Family Planning and Maternal and Child Health Care: Saving Lives, Money, and Time. 2011 Aug. URL: https://www.prb.org/resources/integrating-family-planning-and-maternal-and-child-health-care-saving-lives-money-and-time/ [accessed 2021-03-31]
25. Tsegaye D, Tehone M, Yesuf S, Berhane A. Research Square. Preprint posted online on February 18, 2021 [FREE Full text] [doi: 10.21203/rs.3.rs-207643/v1]
26. Mwaikambo L, Speizer I, Schurmann A, Morgan G, Fikree F. What works in family planning interventions: A systematic review. Stud Fam Plann 2011 Jun;42(2):67-82 [FREE Full text] [doi: 10.1111/j.1728-4465.2011.00267.x] [Medline: 21834409]
27. Hackett K, Henry E, Hussain I, Khan M, Feroz K, Kaur N, et al. Impact of home-based family planning counselling and referral on modern contraceptive use in Karachi, Pakistan: A retrospective, cross-sectional matched control study. BMJ Open 2020 Sep 23;10(9):e039835 [FREE Full text] [doi: 10.1136/bmjopen-2020-039835] [Medline: 32967886]
28. Kurji Z, Premani ZS, Mithani Y. Analysis of the health care system of Pakistan: Lessons learnt and way forward. J Ayub Med Coll Abbottabad 2016;28(3):601-604 [FREE Full text] [Medline: 28712245]
29. Karim MS. The 2017 Census of Pakistan: Analyses of Provisional Results. Research Report No. 101. Karachi, Pakistan: Social Policy and Development Centre; 2018. URL: https://www.academia.edu/43991823/2017_CENSUS_OF_PAKISTAN_SPDC_REPORT_NUMBER_101 [accessed 2021-03-31]
30. Baqui AH, Ahmed S, Begum N, Khanam R, Mohan D, Harrison M, Projahnmo Study Group in Bangladesh. Impact of integrating a postpartum family planning program into a community-based maternal and newborn health program on birth spacing and preterm birth in rural Bangladesh. J Glob Health 2018 Dec;8(2):020406 [FREE Full text] [doi: 10.7189/jogh.08.020406] [Medline: 30023053]
31. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement Sci 2017 Jun 21;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
32. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
33. Dykema J, Basson D, Schaeffer NC. Face-to-face surveys. In: Donsbach W, Traugott MW, editors. The SAGE Handbook of Public Opinion Research. Thousand Oaks, CA: SAGE Publications Inc; 2008:240-248.
34. Caliendo M, Kopeinig S. Some practical guidance for the implementation of propensity score matching. J Econ Surv 2008 Feb;22(1):31-72. [doi: 10.1111/j.1467-6419.2007.00527.x]
35. Sindh Multiple Indicator Cluster Survey 2018-19: Survey Findings Report. Karachi, Pakistan: Bureau of Statistics Sindh, Planning & Development Board, Government of Sindh; 2020. URL: http://sindhbos.gov.pk/wp-content/uploads/2021/03/Sindh-MICS-2018-19-Final-SFR.pdf [accessed 2022-02-16]

36. Sinai I, Omoluabi E, Jimoh A, Jurczynska K. Unmet need for family planning and barriers to contraceptive use in Kaduna, Nigeria: Culture, myths and perceptions. Cult Health Sex 2020 Nov;22(11):1253-1268. [doi: 10.1080/13691058.2019.1672894] [Medline: 31662042]
37. Muttreja P, Singh S. Family planning in India: The way forward. Indian J Med Res 2018 Dec;148(Suppl):S1-S9 [FREE Full text] [doi: 10.4103/ijmr.IJMR_2067_17] [Medline: 30964076]
38. Mathur S, Mehta M, Malhotra A. Youth Reproductive Health in Nepal: Is Participation the Answer?. Washington, DC and New York, NY: International Center for Research on Women and EngenderHealth; 2004 Jan. URL: https://www.icrw.org/wp-content/uploads/2016/10/Youth-Reproductive-Health-in-Nepal-Is-Participation-the-Answer.pdf [accessed 2022-02-16]
39. Mpunga D, Lumbayi J, Dikamba N, Mwembo A, Ali Mapatano M, Wembodinga G. Availability and quality of family planning services in the Democratic Republic of the Congo: High potential for improvement. Glob Health Sci Pract 2017 Jun 06;5(2):274-285 [FREE Full text] [doi: 10.9745/ghsp-d-16-00205]
40. Kuhlmann AS, Gavin L, Galavotti C. The integration of family planning with other health services: A literature review. Int Perspect Sex Reprod Health 2010 Dec;36(4):189-196. [doi: 10.1363/3618910] [Medline: 21245025]
41. Mutisya R, Wambua J, Nyachae P, Kamau M, Karnad S, Kabue M. Strengthening integration of family planning with HIV/AIDS and other services: Experience from three Kenyan cities. Reprod Health 2019 May 29;16(Suppl 1):62 [FREE Full text] [doi: 10.1186/s12978-019-0715-8] [Medline: 31138271]
42. Saeed G, Fakhar S, Rahim F, Tabassum S. Change in trend of contraceptive uptake--Effect of educational leaflets and counseling. Contraception 2008 May;77(5):377-381. [doi: 10.1016/j.contraception.2008.01.011] [Medline: 18402856]
43. Daniel E, Masilamani R, Rahman M. The effect of community-based reproductive health communication interventions on contraceptive use among young married couples in Bihar, India. Int Fam Plan Perspect 2008 Dec 1;34(04):189-197 [FREE Full text] [doi: 10.1363/3418908]
44. Azmat S, Hameed W, Hamza H, Mustafa G, Ishaque M, Abbas G, et al. Engaging with community-based public and private mid-level providers for promoting the use of modern contraceptive methods in rural Pakistan: Results from two innovative birth spacing interventions. Reprod Health 2016 Mar 17;13:25 [FREE Full text] [doi: 10.1186/s12978-016-0145-9] [Medline: 26987368]
45. Silumbwe A, Nkole T, Munakampe M, Milford C, Cordero J, Kriel Y, et al. Community and health systems barriers and enablers to family planning and contraceptive services provision and use in Kabwe District, Zambia. BMC Health Serv Res 2018 May 31;18(1):390 [FREE Full text] [doi: 10.1186/s12913-018-3136-4] [Medline: 29855292]
46. Rahman S, Choudhury A, Khanam R, Moin S, Ahmed S, Begum N, Projahnmo Study Group in Bangladesh. Effect of a package of integrated demand- and supply-side interventions on facility delivery rates in rural Bangladesh: Implications for large-scale programs. PLoS One 2017;12(10):e0186182 [FREE Full text] [doi: 10.1371/journal.pone.0186182] [Medline: 29073229]

Abbreviations

ANC: antenatal care
CPR: contraceptive prevalence rate
FGD: focus group discussion
FP: family planning
LHW: Lady Health Worker
MCM: modern contraceptive method
mCPR: modern contraceptive prevalence rate
MNCH: maternal, newborn, and child health
MWRA: married women of reproductive age
NGO: nongovernmental organization
PNC: postnatal care
RevMan: Review Manager
TDF: Theoretical Domains Framework

Edited by G Eysenbach; submitted 30.11.21; peer-reviewed by S Soofi; comments to author 07.01.22; revised version received 07.01.22; accepted 11.01.22; published 08.03.22.

Please cite as:
Memon ZA, Reale S, Ahmed W, Spencer R, Lashari TH, Bhutta Z, Soltani H
Effects of Integrating Family Planning With Maternal, Newborn, and Child Health Services on Uptake of Voluntary Modern Contraceptive Methods in Rural Pakistan: Protocol for a Quasi-experimental Study
JMIR Res Protoc 2022;11(3):e35291
URL: https://www.researchprotocols.org/2022/3/e35291
doi:10.2196/35291
PMID:35258461

Protocol

Evaluating the Effect of Supported Systematic Work Environment Management During the COVID-19 Pandemic: Protocol for a Mixed Methods Study

Patrik Haraldsson[1,2], MSc; Axel Ros[1,3], MD, PhD; Dirk Jonker[1,2], PhD; Kristina Areskoug Josefsson[1,4,5], PhD

[1]School of Health and Welfare, Jönköping Academy for Improvement of Health and Welfare, Jönköping University, Jönköping, Sweden

[2]Occupational Safety and Health Care, Region Jönköping County, Jönköping, Sweden

[3]Futurum–Academy for Healthcare, Region Jönköping County, Jönköping, Sweden

[4]Faculty of Health Studies, VID Specialized University, Sandnes, Norway

[5]Department of Behavioural Science, Oslo Metropolitan University, Oslo, Norway

Corresponding Author:
Patrik Haraldsson, MSc
School of Health and Welfare
Jönköping Academy for Improvement of Health and Welfare
Jönköping University
Gjuterigatan 5
Jönköping, 553 18
Sweden
Phone: 46 10 2447892
Email: patrik.haraldsson@ju.se

Abstract

Background: The work environment is a complex phenomenon in which many factors interact. Scientific research indicates a relation between the work environment and employee health, staff turnover, patient satisfaction, and patient safety. There is a great need for knowledge on how to conduct work environment interventions and practical work environment management to maximize benefits to the employees.

Objective: The aim of this study is to explore how Occupational Health Service (OHS) support will affect the work environment, sick leave, staff turnover, patient satisfaction, and patient safety during and following the COVID-19 pandemic in a medical ward setting.

Methods: A mixed methods evaluation of a concurrent work environment quality improvement project at the Department of Internal Medicine and Geriatrics in a local hospital in the south of Sweden will be performed.

Results: The mixed methods evaluation of the quality improvement project received funding from Futurum–Academy for Health and Care, Jönköping County Council and Region Jönköping County, and the study protocol was approved by the Swedish Ethical Review Authority. The work environment quality improvement project will continue between May 2020 and December 2021.

Conclusions: The study might contribute to increased knowledge of how work environment interventions and practical work environment management can impact the work environment, and employee health, staff turnover, patient satisfaction, and patient safety. There is a need for knowledge in this area for OHS management to provide increased benefits to employees, employers, and society as a whole.

International Registered Report Identifier (IRRID): DERR1-10.2196/34152

(JMIR Res Protoc 2022;11(3):e34152) doi:10.2196/34152

KEYWORDS

occupational health interventions; implementation; mixed methods; COVID-19 pandemic; COVID-19; pandemic; occupational health; health interventions; health care; support services; employee health

Introduction

Work Environment

"Work environment" is a broad term that implies everything workers are exposed to when working. The work environment can relate to work tools; air, noise, and light; psychological aspects; work organization; and well-being at work [1].

The work environment can be affected by both internal work environment factors as well as external societal factors [2]. Internal work environment aspects must be considered to understand the complexity of the work environment and its effects on human health [3-5]. The physical, environmental, and organizational/social aspects of the work environment must be understood separately [3], and the interaction between these work environment factors must be considered to understand the work environment. The complexity of internal work environment factors is exemplified in a meta-analysis showing associations among workload, job control, decision authority, and social support at work and chronic low back pain [5]. The exposure of work environment factors must be understood in relation to its exposure intensity, duration, and frequency. This is exemplified by a meta-analysis evaluating the effects on lower back pain, showing that both intensity and frequency are important aspects of the exposure [4]. External societal factors such as income inequality, social trust, and public health must also be taken into consideration, since these affect the work environment and occupational health [2].

The COVID-19 pandemic is another example of an external societal factor affecting the work environment and occupational health. Problems related to the work environment, such as high demands and a low degree of support at work, were present in the health care sector before the pandemic [6]. Research shows a high degree of mental health problems such as stress, depression, and sleep disorders in health care personnel as a consequence of the pandemic [7].

Scientific research shows conflicting results regarding work environment interventions and the effect on human health, with some studies showing no effect [8-11] and other studies showing positive effects [12-14]. In addition, there is a lack of both quantitative and qualitative evidence about what type of interventions will most improve the resilience and mental health of frontline workers during epidemics and pandemics [15]. Research to improve knowledge in this area is a high priority, since scientific evidence supports a link between the work environment and health [4,16], staff turnover [17,18], patient satisfaction [19], and patient safety outcomes [20].

Improvement Science

Improvement science might be a suitable approach to deal with the complexity of the work environment [21]. Improvement science has been defined as "a data-driven change process that aims to systematically design, test, implement, and scale change toward systemic improvement, as informed and defined by the experience and knowledge of subject matter experts" [22]. Two central parts of improvement science are the implementation of actions and ongoing evaluations during the actions [22].

An improvement science approach is somewhat supported by a recent systematic review of interventions to improve the work environment in health care, showing that a participatory approach, continuous ongoing improvement projects, and tailoring interventions to the workplace needs are important aspects of interventions to improve nurses' work environments [23].

Systematic Work Environment Management

These previous results support the use of systematic work environment management (SWEM) and Occupational Health Service (OHS) support in SWEM. SWEM is a provision from the Swedish Work Environment Authority that describes mandatory work by the employer to minimize ill health and accidents at work, and promote a satisfactory working environment [24]. SWEM is conducted through risk assessment, measures, and follow-up as a continuous process. When competence within the employer's own activity is insufficient for SWEM, the OHS can be contacted for support [24].

The OHS has a key role in supporting the health and work ability of employees in many settings [25]. The OHS has been endorsed by both the World Health Organization and the International Labor Offices as a means of ensuring a safer, healthier, happier, and more productive workforce [26]. The OHS is characterized as experts who can deliver high-quality services aimed at the working environment, and that are not offered by the employers [25]. Requirements for effective collaboration between employers and the OHS include flexible long-term contracts, effective collaboration with shared goals, frequent contact, trust, and the OHS strategically shifting from being curative to preventive [25]. The SWEM provisions include a requirement for employers without adequate internal competence to contact the OHS for support in the process. However, this is rarely enforced by labor inspectors since they find that the OHS often lacks competence in SWEM [27].

Structured Multidisciplinary Work Evaluation Tool

The Structured Multidisciplinary Work Evaluation Tool (SMET) was developed through action research [28], as no method for OHS support in SWEM was found. SMET consists of four parts, performed in continuous iterations: (1) start-up discussions with the workplace, (2) risk assessment, (3) tailored measures, and (4) evaluation.

The risk assessment has two parts. The first part is the SMET questionnaire by which the employees evaluate the work environment. The second part is an objective in-depth analysis of the workplace, performed by the OHS. The SMET questionnaire consists of 30 items divided into three domains: physically demanding factors, environmentally demanding factors, and psychosocially demanding factors. Each domain consists of a few self-estimating items where the employees rate the degree of work-related problems on a 1-10 scale. Each domain also has one item where the employee will rate which of the previous self-estimating items constitutes the worst problem, and finally an open-ended item. The SMET questionnaire has been evaluated regarding content validity [28], intrarater reliability in the analysis of the open-ended items, and test-retest reliability of the self-estimating items [29], and

has been shown to reflect the true physical workload of nursing assistants in a medical ward setting [30]. The objective in-depth analysis of the workplace is a deeper evaluation of the result in the SMET questionnaire, and entails visiting the workplace and, for example, measuring noise levels and narrow spaces or collecting sick leave data [28,29].

It is of great importance to increase the knowledge of how OHS support, with SMET in SWEM, should be conducted to benefit employers and employees, and what types of benefits can be achieved.

The aim of this study is to explore how OHS support, with SMET in SWEM, will affect the work environment, sick leave, staff turnover, patient satisfaction, and patient safety during and following the COVID-19 pandemic in a medical ward setting.

Methods

Design

The study involves an action research approach, with evaluation of a concurrent work environment quality improvement project using an interactive mixed methods design [31].

Sample

Ongoing evaluation is conducted with all employees at the Department of Internal Medicine and Geriatrics in a local hospital in the south of Sweden based on a work environment project at the department. Involved professions are specialist physicians, resident physicians, intern physicians, nurses, nursing assistants, and care administrators. Inclusion criteria are individuals working at the studied department who want to participate in the research project. Exclusion criteria are hourly employees and employees on sick leave and parental leave.

Clinical Intervention

The work environment project started at the Department of Internal Medicine and Geriatrics in May 2020 to promote the work environment and health during and following the COVID-19 pandemic. Mapping and reporting the work environment will be conducted in three steps at the department, every quarter from June 2020 to December 2021: (1) mapping the work environment with a questionnaire to all employees, (2) meeting with a reference group, and (3) meeting with the management.

The reference group consists of six employees, from different parts of the department and with different professions. Participants in the reference group were based on a pragmatic sample according to willingness to participate.

Results regarding the work environment from the questionnaire and the meeting with the reference group are compiled and presented to the management. Based on the results, the management can initiate tailored interventions in the work environment. Due to the COVID-19 pandemic, no objective in-depth analysis in the actual workplace is conducted, since this would have meant additional staff being present physically at the department.

Data Collection

Research data will be collected quarterly through interactive acquisition and evaluation of qualitative and quantitative data regarding the work environment, leadership qualities, and work environment interventions within the work environment project. The work environment quality improvement project started in May 2020. Quantitative data will be collected with the SMET questionnaire for evaluation of the work environment, and with three questions regarding leadership qualities from the Copenhagen Psychosocial Questionnaire 3 (COPSOQ III). Both questionnaires have been tested and have shown good psychometric properties [28,29,32,33]. The use of the three items of leadership quality in COPSOQ III in isolation is supported by good internal consistency and floor/ceiling effects in these items [33].

Qualitative data will be collected through meetings with a reference group and manager interviews. Meetings with the reference group, lasting approximately 1.5 hours, will be held every quarter and will be led by the first author. The results from the questionnaires will be presented to the reference group, who will then discuss the result based on three reflection topics: *Is the result correct? Causes? Solutions?* The task of the reference group is to support the OHS in interpreting the results based on their work context. The results from the reference group will be compiled using written meeting diaries (by the first author), with these topics.

Monthly interviews will be conducted by the first author with the Deputy Head of the department for continuous evaluation of occupancy rate, ongoing work environment interventions, and other organizational interventions at the department. The interviews will be conducted by telephone and will last 20-30 minutes. The results from these interviews will be compiled using written meeting diaries (by the first author) with the topics occupancy rate, ongoing work environment interventions, and other organizational interventions at the department.

Data on sick leave and staff turnover will be collected from the human resources organization, from January 2015 to the end of the project in December 2021. Data on patient satisfaction will be collected from national registers at the Swedish Association of Local Authorities and Regions, and data on patient safety will be collected from the national quality register Senior Alert.

Data Analysis

The Department of Internal Medicine and Geriatrics has approximately 240 employees. Power analysis showed that to identify a moderate effect size (Cohen W>0.30) in the SMET questionnaire with an α value of .05, power of 0.80, and 2 degrees of freedom, a sample of at least 54 individuals/measurement opportunities is needed. Qualitative data in the SMET questionnaire will be analyzed by content analysis, as described in Haraldsson et al [29]. Qualitative data from the monthly interviews with the Deputy Head of the department and from meetings with the reference group will be compiled using written meeting diaries (by the first author). The use of research notes such as written meeting diaries has been shown to offer data of good quality, being less time-consuming and more cost-efficient than verbatim

transcription of interviews [34]. Data on sick leave, staff turnover, patient satisfaction, and patient safety will be collected and divided into two parts: the time before the COVID-19 pandemic (January 2015 to February 2020), and during and following the COVID-19 pandemic (March 2020 to December 2021) for comparison with a reference group consisting of several other departments of internal medicine. A mixed methods approach, with integration of both quantitative and qualitative data, will be used to increase the understanding of the results [31].

Ethical Considerations

This study is a mixed methods intervention study, where the interactive research follows an ongoing quality improvement project. Questionnaire data on the work environment and leadership might be considered sensitive personal data, and data on sick leave are definitely considered as such. The collected data might be regarded as sensitive concerning personal integrity and will therefore be presented only at the group level. Conducting a study at only one department in a hospital might have the risk that an individual employee will be indirectly identified. This risk is considered insignificant since the department has 240 employees. Email addresses of all the employees will be acquired from the department. The questionnaires will be sent by the online questionnaire program esMaker. The questionnaires will be anonymized in esMaker. This means that the anonymization will be conducted by the esMaker system and not by the researchers. The results from the reference group will be compiled at the group level, which implies that these data will be anonymized as well.

Workplace data on sick leave and staff turnover will be used, as these data are continuously collected as routine practice by the regions in Sweden. Regarding data on sick leave and staff turnover, the personal integrity will be secured by obtaining the data at the group level by the human resource department. No individual data will be acquired, handled, or presented by the research group. The collection of these data is important to be able to achieve the scientific aims, and to evaluate if the quality project was successful in promoting the work environment and health.

Informed written consent will be collected from the head of the department and from the participants in the reference group meetings before the study is started. All employees at the involved department will receive oral and written information about the follow-up research project. Informed consent will be collected from all 240 employees, through the questionnaire, regarding the work environment, but it will not be collected regarding data on sick leave and staff turnover, since these data will only be presented at the group level and collection of such is part of an ongoing quality project at the clinic. Data from the questionnaire, diaries from the reference group, sick leave, staff turnover, and diaries from the monthly interviews with the Deputy Head of the department will be stored on a safe hard drive in Region Jönköping County in accordance with General Data Protection Regulation. No individual data, only data at the group level, will be stored. The study protocol was developed in accordance with the Helsinki Declaration [35] and the Swedish Ethical Review Act [36], and was approved by the Ethical Review Authority on October 26, 2020 (Dnr: 2020-03891).

Results

The mixed methods evaluation of the quality improvement project received funding from Futurum–Academy for Health and Care, Jönköping County Council and Region Jönköping County, and the study protocol was approved by the Swedish Ethical Review Authority. The work environment quality improvement project started at the Department of Internal Medicine and Geriatrics in May 2020. Mapping and reporting the work environment to the management have been conducted every quarter from June 2020 to September 2021. On the basis of the results, the management has conducted tailored interventions to promote the work environment and employee health continuously during and following the COVID-19 pandemic. A scientific evaluation of the work environment quality improvement project will continue when the project ends in December 2021.

Discussion

Principal Results

There is great complexity in assessing a work environment. The conflicting results in work environment intervention studies and the suboptimal use of OHS in SWEM might be an outcome of this complexity. The results from this study will contribute to bridging the knowledge gap between SWEM and effective interventions in this field. A deeper understanding of factors linked to practical work environment management in hospitals can benefit employee health, staff turnover, patient satisfaction, and patient safety.

Central parts of SWEM and improvement science are ongoing evaluations and implementation of actions. Considering ongoing evaluations, valid and reliable tools for measurement are central to increasing the understanding and learning in the process. Our previous research has shown the SMET questionnaire to be a valid and reliable method for evaluation of the work environment [28,29], which ensures good quality of the risk assessment in SMET.

The study is a mixed methods evaluation of a concurrent work environment quality improvement project conducted before, during, and following the COVID-19 pandemic. The COVID-19 pandemic has shown that there is a great need for increased knowledge about how to protect health care personnel, regarding work environment stress, during disease epidemics and pandemics [7]. A recent Cochrane report showed a lack of both quantitative and qualitative evidence with regard to how resilience and mental health can be increased in frontline workers during and after epidemics and pandemics. The authors state that research to determine the effectiveness of such interventions is a high priority [15].

The results from this study will be used to improve the work environment in the regional context but will also contribute to knowledge in work environment interventions from a wider perspective. The results will be disseminated through national and international conferences as well as scientific journals.

Conclusions

The study might add knowledge about work environment management and intervention studies with SMET, and how to conduct work environment interventions with a systems approach, a topic in great need of increased knowledge [37]. The study might also contribute to increased knowledge of how work environment interventions and practical work environment management can impact other factors linked to the work environment. Increased knowledge in this area is of great importance, since scientific research indicates a relation between the work environment and employee health [38-40], staff turnover [17,18], patient satisfaction [19], and patient safety [18,20]. There is a need for knowledge in this area for OHS management to increasingly benefit employees, employers, and society as a whole.

Acknowledgments

The research has been funded by Futurum–Academy for Health and Care, Jönköping County Council and Region Jönköping County (FUTURUM-941684). The funder did not, in any way, participate in the research.

Authors' Contributions

All authors planned the study and contributed to the study design. PH conducted specific literature reviews and wrote the manuscript. AR, DJ, and KAJ conducted continuous proofreading and participated in discussions regarding the manuscript. KAJ acted as scientific supervisor.

Conflicts of Interest

None declared.

References

1. Working Environment. Umeå University. URL: https://www.aurora.umu.se/en/my-employment/work-environment-health-and-equal-opportunities/work-environment/ [accessed 2021-01-21]
2. Torp S, Reiersen J. Globalization, work, and health: a Nordic perspective. Int J Environ Res Public Health 2020 Oct 21;17(20):7661 [FREE Full text] [doi: 10.3390/ijerph17207661] [Medline: 33096630]
3. van der Beek AJ, Dennerlein JT, Huysmans MA, Mathiassen SE, Burdorf A, van Mechelen W, et al. A research framework for the development and implementation of interventions preventing work-related musculoskeletal disorders. Scand J Work Environ Health 2017 Nov 01;43(6):526-539 [FREE Full text] [doi: 10.5271/sjweh.3671] [Medline: 28945263]
4. Coenen P, Gouttebarge V, van der Burght ASAM, van Dieën JH, Frings-Dresen MHW, van der Beek AJ, et al. The effect of lifting during work on low back pain: a health impact assessment based on a meta-analysis. Occup Environ Med 2014 Dec;71(12):871-877. [doi: 10.1136/oemed-2014-102346] [Medline: 25165395]
5. Buruck G, Tomaschek A, Wendsche J, Ochsmann E, Dörfel D. Psychosocial areas of worklife and chronic low back pain: a systematic review and meta-analysis. BMC Musculoskelet Disord 2019 Oct 25;20(1):480 [FREE Full text] [doi: 10.1186/s12891-019-2826-3] [Medline: 31653249]
6. Gohar B, Larivière M, Nowrouzi-Kia B. Sickness absence in healthcare workers during the COVID-19 pandemic. Occup Med 2020 Jul 17;70(5):338-342 [FREE Full text] [doi: 10.1093/occmed/kqaa093] [Medline: 32449751]
7. Rossi R, Socci V, Pacitti F, Di Lorenzo G, Di Marco A, Siracusano A, et al. Mental health outcomes among frontline and second-line health care workers during the coronavirus disease 2019 (COVID-19) pandemic in Italy. JAMA Netw Open 2020 May 01;3(5):e2010185 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10185] [Medline: 32463467]
8. Comper MLC, Dennerlein JT, Evangelista GDS, Rodrigues da Silva P, Padula RS. Effectiveness of job rotation for preventing work-related musculoskeletal diseases: a cluster randomised controlled trial. Occup Environ Med 2017 Aug 01;74(8):545-552. [doi: 10.1136/oemed-2016-104077] [Medline: 28250047]
9. Gupta N, Wåhlin-Jacobsen CD, Abildgaard JS, Henriksen LN, Nielsen K, Holtermann A. Effectiveness of a participatory physical and psychosocial intervention to balance the demands and resources of industrial workers: A cluster-randomized controlled trial. Scand J Work Environ Health 2018 Jan 01;44(1):58-68 [FREE Full text] [doi: 10.5271/sjweh.3689] [Medline: 29095478]
10. Brandt M, Madeleine P, Samani A, Ajslev JZ, Jakobsen MD, Sundstrup E, et al. Effects of a participatory ergonomics intervention with wearable technical measurements of physical workload in the construction industry: cluster randomized controlled trial. J Med Internet Res 2018 Dec 19;20(12):e10272 [FREE Full text] [doi: 10.2196/10272] [Medline: 30567694]
11. Petrie K, Crawford J, Baker STE, Dean K, Robinson J, Veness BG, et al. Interventions to reduce symptoms of common mental disorders and suicidal ideation in physicians: a systematic review and meta-analysis. Lancet Psychiatry 2019 Mar;6(3):225-234. [doi: 10.1016/S2215-0366(18)30509-1] [Medline: 30744997]
12. Rasmussen CDN, Sørensen OH, van der Beek AJ, Holtermann A. The effect of training for a participatory ergonomic intervention on physical exertion and musculoskeletal pain among childcare workers (the TOY project) - a wait-list cluster-randomized controlled trial. Scand J Work Environ Health 2020 Jul 01;46(4):429-436 [FREE Full text] [doi: 10.5271/sjweh.3884] [Medline: 31945165]

13. Panagioti M, Panagopoulou E, Bower P, Lewith G, Kontopantelis E, Chew-Graham C, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. JAMA Intern Med 2017 Feb 01;177(2):195-205. [doi: 10.1001/jamainternmed.2016.7674] [Medline: 27918798]
14. De Simone S, Vargas M, Servillo G. Organizational strategies to reduce physician burnout: a systematic review and meta-analysis. Aging Clin Exp Res 2021 Apr 09;33(4):883-894. [doi: 10.1007/s40520-019-01368-3] [Medline: 31598914]
15. Pollock A, Campbell P, Cheyne J, Cowie J, Davis B, McCallum J, et al. Interventions to support the resilience and mental health of frontline health and social care professionals during and after a disease outbreak, epidemic or pandemic: a mixed methods systematic review. Cochrane Database Syst Rev 2020 Nov 05;11:CD013779 [FREE Full text] [doi: 10.1002/14651858.CD013779] [Medline: 33150970]
16. Joyce K, Pabayo R, Critchley JA, Bambra C. Flexible working conditions and their effects on employee health and wellbeing. Cochrane Database Syst Rev 2010 Feb 17(2):CD008009 [FREE Full text] [doi: 10.1002/14651858.CD008009.pub2] [Medline: 20166100]
17. Kuo H, Lin K, Li I. The mediating effects of job satisfaction on turnover intention for long-term care nurses in Taiwan. J Nurs Manag 2014 Mar;22(2):225-233. [doi: 10.1111/jonm.12044] [Medline: 23465339]
18. Cho H, Steege LM. Nurse fatigue and nurse, patient safety, and organizational outcomes: a systematic review. West J Nurs Res 2021 Dec;43(12):1157-1168. [doi: 10.1177/0193945921990892] [Medline: 33554767]
19. Copanitsanou P, Fotos N, Brokalaki H. Effects of work environment on patient and nurse outcomes. Br J Nurs 2017 Feb 09;26(3):172-176. [doi: 10.12968/bjon.2017.26.3.172] [Medline: 28185485]
20. Duffield C, Diers D, O'Brien-Pallas L, Aisbett C, Roche M, King M, et al. Nursing staffing, nursing workload, the work environment and patient outcomes. Appl Nurs Res 2011 Nov;24(4):244-255. [doi: 10.1016/j.apnr.2009.12.004] [Medline: 20974086]
21. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. BMJ Qual Saf 2014 Apr 11;23(4):290-298 [FREE Full text] [doi: 10.1136/bmjqs-2013-001862] [Medline: 24025320]
22. Lemire S, Christie CA, Inkelas M. The methods and tools of improvement science. In: Christie CA, Inkelas M, Lemire S, editors. Improvement science in evaluation: Methods and uses: new directions for evaluation. Hoboken, NJ: Wiley; 2017.
23. Paguio JT, Yu DSF, Su JJ. Systematic review of interventions to improve nurses' work environments. J Adv Nurs 2020 Oct 08;76(10):2471-2493. [doi: 10.1111/jan.14462] [Medline: 32770584]
24. Systematic Work Environment Management (Provision, AFS 2001). The Swedish Work Environment Authority. URL: https://www.av.se/en/work-environment-work-and-inspections/publications/foreskrifter/systematic-work-environment-management-afs-20011-provisions/ [accessed 2021-01-21]
25. Halonen JI, Atkins S, Hakulinen H, Pesonen S, Uitti J. Collaboration between employers and occupational health service providers: a systematic review of key characteristics. BMC Public Health 2017 Jan 05;17(1):22 [FREE Full text] [doi: 10.1186/s12889-016-3924-x] [Medline: 28056910]
26. Lim MK. Health and economic impact of occupational health services. Scand J Work Env Health 2005(Suppl 1):38-42 [FREE Full text]
27. Frick K. The 50/50 implementation of Sweden's mandatory systematic work environment management. Policy Pract Health Saf 2016 Jan 05;12(2):23-46. [doi: 10.1080/14774003.2014.11667802]
28. Haraldsson P, Jonker D, Strengbom E, Areskoug-Josefsson K. Structured Multidisciplinary Work Evaluation Tool: development and validation of a multidisciplinary work questionnaire. Work 2016;55(4):883-891. [doi: 10.3233/WOR-162454] [Medline: 28059819]
29. Haraldsson P, Jonker D, Rolander B, Strengbom E, Areskoug-Josefsson K. Structured Multidisciplinary Work Evaluation Tool (SMET): reliability testing of a multidisciplinary/multifactorial work questionnaire. Work 2019;62(2):287-297. [doi: 10.3233/WOR-192863] [Medline: 30829639]
30. Haraldsson P, Areskoug-Josefsson K, Rolander B, Strengbom E, Jonker D. Comparing the Structured Multidisciplinary work Evaluation Tool (SMET) questionnaire with technical measurements of physical workload in certified nursing assistants in a medical ward setting. Appl Ergon 2021 Oct;96:103493 [FREE Full text] [doi: 10.1016/j.apergo.2021.103493] [Medline: 34116412]
31. Bowen P, Rose R, Pilkington A. Mixed methods-theory and practice. Sequential, explanatory approach. Int J Quant Qual Res Methods 2017;5(2):10-27.
32. Berthelsen H, Westerlund H, Bergström G, Burr H. Validation of the Copenhagen Psychosocial Questionnaire Version III and establishment of benchmarks for psychosocial risk management in Sweden. Int J Environ Res Public Health 2020 May 02;17(9):3179 [FREE Full text] [doi: 10.3390/ijerph17093179] [Medline: 32370228]
33. Burr H, Berthelsen H, Moncada S, Nübling M, Dupret E, Demiral Y, International COPSOQ Network. The Third Version of the Copenhagen Psychosocial Questionnaire. Saf Health Work 2019 Dec;10(4):482-503 [FREE Full text] [doi: 10.1016/j.shaw.2019.10.002] [Medline: 31890332]
34. Eaton K, Stritzke W, Ohan J. Using scribes in qualitative research as an alternative to transcription. Qual Rep 2019 Mar 18;24(3):586-605. [doi: 10.46743/2160-3715/2019.3473]

35. Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. World Medical Association. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2020-08-20]
36. Värnar människan i forskning. Swedish Ethical Review Authority. URL: https://etikprovningsmyndigheten.se/ [accessed 2021-03-03]
37. Work environment, organization and health - how is it connected and why? Sveriges Företagshälsor. URL: http://www.foretagshalsor.se/sv/rapporter [accessed 2020-12-16]
38. Andersen LL, Clausen T, Persson R, Holtermann A. Perceived physical exertion during healthcare work and risk of chronic pain in different body regions: prospective cohort study. Int Arch Occup Environ Health 2013 Aug;86(6):681-687. [doi: 10.1007/s00420-012-0808-y] [Medline: 22878558]
39. Andersen LL, Fallentin N, Thorsen SV, Holtermann A. Physical workload and risk of long-term sickness absence in the general working population and among blue-collar workers: prospective cohort study with register follow-up. Occup Environ Med 2016 Apr;73(4):246-253. [doi: 10.1136/oemed-2015-103314] [Medline: 26740688]
40. Magnusson Hanson LL, Theorell T, Oxenstierna G, Hyde M, Westerlund H. Demand, control and social climate as predictors of emotional exhaustion symptoms in working Swedish men and women. Scand J Public Health 2008 Sep;36(7):737-743. [doi: 10.1177/1403494808090164] [Medline: 18684778]

Abbreviations

COPSOQ III: Copenhagen Psychosocial Questionnaire 3
OHS: Occupational Health Service
SMET: Structured Multidisciplinary Work Evaluation Tool
SWEM: systematic work environment management

Edited by T Leung; submitted 08.10.21; peer-reviewed by D Madhusudhan, B Knezevic; comments to author 06.02.22; revised version received 11.02.22; accepted 11.02.22; published 10.03.22.

Please cite as:
Haraldsson P, Ros A, Jonker D, Areskoug Josefsson K
Evaluating the Effect of Supported Systematic Work Environment Management During the COVID-19 Pandemic: Protocol for a Mixed Methods Study
JMIR Res Protoc 2022;11(3):e34152
URL: https://www.researchprotocols.org/2022/3/e34152
doi:10.2196/34152
PMID:35234649

Protocol

Creating Effective, Evidence-Based Video Communication of Public Health Science (COVCOM Study): Protocol for a Sequential Mixed Methods Effect Study

Jo Røislien[1], MSc, PhD; Jane K O'Hara[2], MSc, PhD; Ionica Smeets[3], MSc, PhD; Kolbjørn Brønnick[1], MSc, PhD; Siv Hilde Berg[1], MSc, PhD; Marie Therese Shortt[1], MA, PhD; Daniel Adrian Lungu[1], MSc, PhD; Henriette Thune[1], MA, PhD; Siri Wiig[1], MSc, PhD

[1]SHARE - Centre for Resilience in Healthcare, Faculty of Health Sciences, University of Stavanger, Stavanger, Norway

[2]School of Healthcare, Faculty of Medicine & Health, University of Leeds, Leeds, United Kingdom

[3]Science Communication and Society, Institute of Biology, Leiden University, Leiden, Netherlands

Corresponding Author:
Jo Røislien, MSc, PhD
SHARE - Centre for Resilience in Healthcare
Faculty of Health Sciences
University of Stavanger
Kjell Arholms gate 41
Stavanger, 4021
Norway
Phone: 47 51 83 10 00
Email: jo.roislien@uis.no

Abstract

Background: The nonlinear nature of contagious diseases and the potential for exponential growth can be difficult to grasp for the general public. This has strong implications for public health communication, which needs to be both easily accessible and efficient. A pandemic is an extreme situation, and the accompanying strict societal measures are generally easier to accept if one understands the underlying reasoning behind them. Bringing about informed attitude change and achieving compliance to strict restrictions requires explanations of scientific concepts and terminologies that laypersons can understand.

Objective: The aim of the project is to develop effective, evidence-based modes of video communication for translating complex, but important, health messages about pandemics to both the general population and decision makers. The study uses COVID-19 as a case to learn and prepare society for handling the ongoing and future pandemics, as well as to provide evidence-based tools for the science communication toolbox.

Methods: The project applies a mixed methods design, combining qualitative methods (eg, interviews, observational studies, literature reviews) and quantitative methods (eg, randomized controlled trials [RCTs]). The project brings together researchers from a wide range of academic fields, as well as communication industry professionals.

Results: This study has received funding from the Trond Mohn Foundation through the Research Council of Norway's "COVID-19 Emergency Call for Proposals" March 2020. Recruitment and data collection for the exploratory first phase of the project ran from February 2021 to March 2021. Creative communication work started in May 2021, and the production of videos for use in the RCTs in the final phase of the project started in September 2021.

Conclusions: The COVCOM project will take on several grand challenges within the field of communicating science and provide evidence-based tools to the science communication toolbox. A long-term goal of the project is to contribute to the creation of a more resilient health care system by developing communication responses tailormade for different audiences, preparing society for any future pandemic.

International Registered Report Identifier (IRRID): DERR1-10.2196/34275

(JMIR Res Protoc 2022;11(3):e34275) doi:10.2196/34275

KEYWORDS

pandemics; risk; public health; science communication; mixed methods; evidence-based medicine; COVID-19

Introduction

A pandemic is an extreme situation, and extreme measures are needed to combat it. Although similar numbers of people die annually in traffic or from cancer, the number of people who catch a contagious disease can grow at a rapidly increasing rate. This nonlinear nature of contagious diseases and the potential for exponential growth can be difficult to grasp for the general public. At the same time, this feature of contagious diseases brings about a need for measures that are not only rapid but also often radical. This has strong implications for public health communication, which needs to be both easily accessible and efficient.

During the COVID-19 pandemic, scientific knowledge from various fields of health research has been at the heart of decision-making. Strong preventive and societal measures—with substantial implications on peoples' lives—have been central to combating the pandemic. Such measures are generally easier to accept if one understands the underlying reasoning behind them [1-3], and bringing about informed attitude change and achieving compliance to strict restrictions thus require explanations of scientific concepts and terminologies that laypersons can understand. Therefore, in order to effectively tackle pandemics, the provision of large-scale information outreach is imperative.

However, how to do this—or how *not* to do it—is poorly understood, with several grand challenges, sparse research within the field, and limited available empirical evidence for making generalizations [4,5]. Systematic reviews [6-9] have found multiple knowledge gaps in pandemic risk communication, for example how risk perception is framed by competing narratives, the lack of in-depth learning from past experiences [8], and slim evidence of the effectiveness of pandemic risk communication [6].

Guaranteeing the flow of information to engage stakeholders is important [9]. Different stakeholders (eg, policy makers, infection control bodies, primary and specialized health care providers) need a common understanding of risk, risk implications, and consequences. This implies that the understanding of key scientific information, and the lack thereof, might facilitate understanding of why national and international societies have not been able to learn from the past and prepare for predicted catastrophic pandemics like COVID-19 [10].

Although crisis communication is an established research field within leadership [11], little is known about the communication of the science upon which leadership and decision-making are based. In the COVID-19 pandemic, this typically entails health science topics related to contagious diseases, including mathematics, epidemiology, risk, medicine, and health. Establishing evidence-based communication that helps facilitate understanding, attitudes, and risk behaviors, and what works for whom, is essential.

The first empirical gap analysis in the field of science communication—including ~3000 papers—identified several grand challenges [4]. First, the field is mostly limited to one-off studies, and there is a need for longitudinal, experimental, comparative, and wider systemic research to understand how contents and channels, actors, and audiences interrelate. Second, the field is caught in disciplinary structures, and the opportunity for interdisciplinary integration has not been seized. Third, there is a lack of transfer and collaboration between researchers and practitioners, and getting practitioners involved in the research has been suggested to bridge the divide. Randomized controlled trials (RCTs) in science communication are few, with retrospective analyses being the norm. The move toward evidence-based science communication, inspired by the impact of evidence-based medicine over the past decades, has been called for [5]. Experimental studies, capable of invalidating previously accepted practices and replacing them with new ones that are more accurate and effective [12], is an integral part of achieving this.

All communication takes place within a continuously changing culture, and narratives, images, and metaphors that worked yesterday might not work today. This must be taken into account when creating health science communication. *How* you say something is as important as *what* you say. Arts is central for outreach [13], and in March 2020, the World Health Organization (WHO) sent out an open global call to creatives to help with communication in the COVID-19 pandemic [14,15]: Creativity in health science communication is a necessity. It is however poorly understood and underused.

The population's media habits have changed rapidly in recent years. Driven by social media and smartphones, each citizen can now choose on what to spend their time. Video consumption is increasing, and video makes up an estimated 80% of all internet traffic [16]. YouTube, the world's largest video site, has 2 billion monthly users, with >500 hours of content uploaded every minute [17]. In the United States, YouTube on mobile alone reaches more 18- to 49-year-olds than any cable TV network [18]. However, despite the scale of video consumption, there is little extant evidence to guide the effective use of video for relaying complex health messages. It is well-known in risk and public health communication that trust is key [19,20], and there is a growing body of research on narrative-based methods and digital storytelling within medicine and health [21,22]. However, much research remains before we fully understand how to best utilize the video format for effective public health communication.

Film is a powerful communication format, frequently used in advertising and popular culture. It is a collective process in which screenwriters, directors, cinematographers, set designers, and professional performers come together to tell a story using moving pictures. Artistic choices are central to filmmaking but are rarely considered when communicating science through film. In an analysis of 400 science videos on YouTube, videos generally fell into 1 of 4 categories: video-blog, voice-over animation, recorded presentation, or interview [23]. Notably, none of the most viewed videos on YouTube in 2019 fell into any of these 4 categories. That is, the video styles most often used by science communicators are not the video styles that tend to attract large audiences, indicating that there is something to be gained by exploring other creative means when producing videos in order to attract larger—and other—audiences when communicating scientific knowledge. In order to create effective

science communication, interdisciplinary collaboration of experts from disciplines with different norms and practices is needed [24]. Achieving this is challenging, and the field remains immature.

To tackle the grand challenges within the field of science communication [5], the current project takes an interdisciplinary approach to developing effective communication for pandemics. Creating effective science communication requires collaboration between not only scientists with expert knowledge in the subject matter and in communication [25] but also actual communicators. This study brings together experts from a wide range of fields, including researchers from health studies, humanities, risk, societal safety, medicine, nursing, public health studies, psychology, visual communication, epidemiology and statistics, as well as mass media professionals, communicators, and filmmakers, in order to ensure first-hand cultural know-how for contemporary communication. This is done order to move from experience to evidence-based communication.

The primary objective of the study is to use video to develop effective, evidence-based modes of communication for translating complex, but important, health messages about pandemics. The study uses COVID-19 as a case to learn and prepare society for handling the ongoing and future pandemics, as well as to provide evidence-based tools for the science communication toolbox. This will be achieved through the following secondary objectives:

1. Identify communication strategies and key topics about pandemics that public health scientists and officials need to communicate
2. Explore communication strategies and artistic dimensions in filmmaking for the creation of effective science communication videos aimed at lay viewers, focusing primarily on the adult part of the Norwegian speaking population with a general primary school level understanding of science
3. Test the effect of these videos through controlled experiments, coupling communication to learning outcome and individual differences such as attitude(s) and compliance towards the topic(s), exploring also sociodemographic variables
4. Identify the features of the most effective videos on a mass communication scale in collaboration with national mass media broadcasters

We hypothesize that shorter, props-driven videos will outperform longer, more static and scientifically focused videos in terms of comprehension of the topics communicated, while trust will be higher in the latter. For intensions, behavior, and behavior change, we hypothesize that narratives and metaphors will be more effective than the factual scientific information itself. Further, optimizing communication for accessibility by reducing scientific precision will not significantly reduce comprehension, intensions, or active behavior. Finally, observed effects will vary according to the demographic characteristics of the receiver.

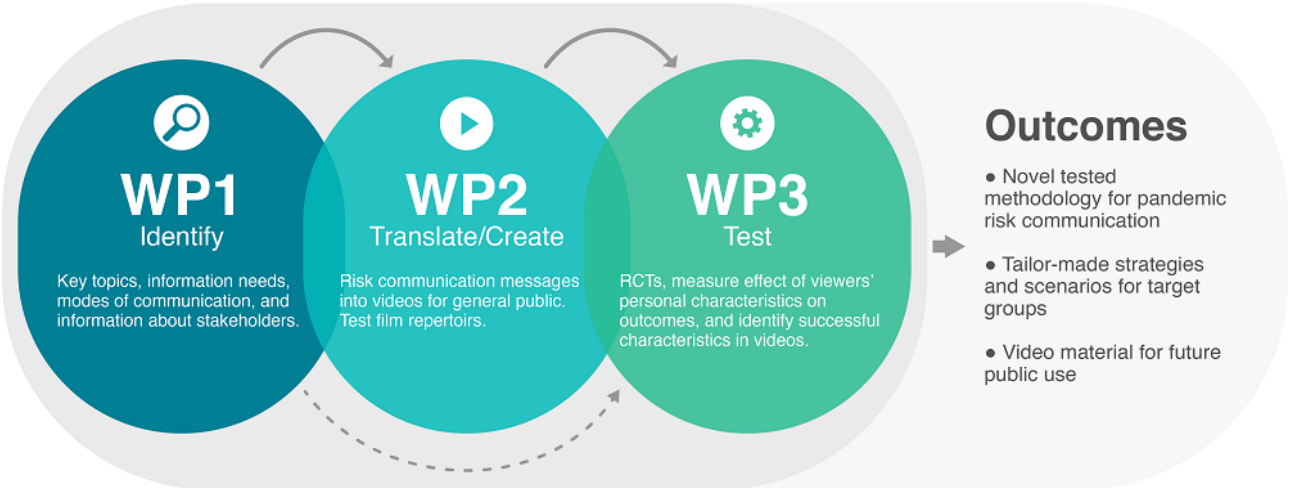
Methods

Study Design

The study applies a sequential mixed methods design [26], combining various qualitative and quantitative research methods.

Research in science education [27,28], health communication [29], and cognitive anthropology and psychology [30,31] suggests that people interpret new information in light of their existing beliefs. Rather than relying on scientists' opinions only, the communication of scientific knowledge should thus be based on evidence of the audience's relevant beliefs and what they are still missing [32,33]. Aligned with this, we apply a so-called mental models approach to developing communications as our theoretical anchor for the sequential design to communicating scientific knowledge [34]: First, identify what people need to know to make more informed decisions; second, identify what they already know and how they make decisions; third, create the communication; fourth, test its effectiveness. This approach resonates with the call for an evidence-based approach to communicating science [5,35]. Correspondingly, the project is operationalized into 3 work packages (WPs) that build on each other (Figure 1). These 3 phases of the project are described in detail in the following sections.

Figure 1. Study design and workflow of work packages (WPs). RCTs: randomized controlled trials.



Recruitment

A large part of the study will run throughout the course of a worldwide pandemic, with the accompanying societal restrictions setting the context of the project. Most of the data collection will take place digitally, with reduced need for travel and physical meetings, using online surveys, video interviews, and interactive methods (eg, stakeholder analyses and workshops). The study population will include both experts and representatives from the general public as part of a holistic approach to how health risk and consequences are understood and communicated on all levels. More detail about the individual WPs is provided in the following sections.

Phase 1: Establishing What Needs to Be Communicated (WP1)

WP1 will apply multiple qualitative methods to identify communication strategies and key topics related to pandemic risks to be communicated to the public. The views from both communicators and receivers of pandemic information are included. WP1 will inform WP2 and WP3 on what people already know, how they make decisions, and what they need to know to make more informed decisions when facing pandemic risks [32,34].

Rapid Scoping Review

The WP will involve a rapid scoping review [36,37] to obtain an overview of the evidence pertaining to diverse modes of communication used by health authorities in health risk communication with the public during a pandemic. The databases MEDLINE and EMBASE will be searched for publications from 2009 to 2020. This will provide information about key scientific concepts, types of outcomes, and research gaps related to diverse modes of communication.

Mental Models

Employing semistructured individual interviews, WP1 will involve the creation of 2 mental models. First, a public mental model will explore how the public perceives difficult scientific concepts and acts on the public health risk communication related to COVID-19. Second, an expert mental model will identify key topics and communication strategies related to pandemic risk [32]. Analyses will be guided by the mental model framework by de Bruin and Bostrom [34]. Mental models are representations of how something works in the real world, and the framework by de Bruin and Bostrom [34] was developed to assess what to address in science communication.

Rather than recruiting a large representative sample, it is recommended to recruit a small but diverse sample of participants (~10-15 participants) when exploring mental models [34]. For the public mental model, we apply a purposive sample of Norwegian citizens between 18 years and 80 years old, with various levels of education, gender, and a range of geographic regions in Norway. For the expert mental model, we include experts with various levels of knowledge of issues regarding COVID-19 risk and mitigation at different system levels (eg, municipality, hospital, national level, research). The identified topics will be analyzed in a directed content analysis, focusing on scientific concepts for public risk communication and the identification of possible new scientific concepts for pandemic risk communication [38].

Stakeholder Analysis

Finally, in WP1, we will undertake a stakeholder analysis to identify various stakeholder groups, assessing their different roles and values [39]. This to gain insight into the responsibility of different stakeholder groups, the perspectives of those involved in public health risk communication, and possible trade-offs in public risk communication. Here, we define stakeholders as individuals who represent a unit or organization that participates in public risk communication related to COVID-19 in Norway.

First, we will identify key stakeholder groups for pandemic risk communication at the local municipal level, at the hospital level, and at the policy level by reviewing publicly available policy documents describing the stakeholder's roles and responsibility in pandemic management. Our sampling strategy is to include stakeholders with key roles in risk communication across these levels in the health care system. We then assess their different perspectives and values by interviewing a sample of ~10 stakeholders involved in communication of pandemic risk information representing different system levels. The document and participant interview data will be analyzed in a directed qualitative content analysis with predetermined categories for micro, macro, and meso levels and inductive category development to analyze values and perspectives [40]. Here, micro-level research refers to the examination of individuals and individual-level interactions (eg, intentions, feelings, and beliefs), the meso-level examines groups (including teams, units, and organizations), and macro-level research examines the political-administrative environment (including national systems, regulation, and cultures).

Recruitment and Sampling

User representatives, nongovernmental organizations, partners in the consortium from Stavanger University Hospital, Haukeland University Hospital, and the Center for Developing Institutional and Home Care Services Sogn and Fjordane will contribute to the recruitment in the various WP1 substudies. Participants will also be recruited by inviting eligible participants identified by the COVCOM team.

Phase 2: Creating Communication (WP2)

Creative choices are central to filmmaking but are rarely considered when communicating science. In this WP, we will study the creative processes that underlie effective translation of scientific information into understandable communication.

Video Review

To explore how videos created by health authorities measure up to contemporary video content, WP2 will include a video review of existing COVID-19 video communication. Online sites for Norwegian health authorities, including health entities at both national and regional levels, will be searched for video content, and entities with a dedicated YouTube channel featuring COVID-19–related video content will be included. These videos will be compared with COVID-19–related videos created by the WHO, as well as the most watched COVID-19 videos on

YouTube. Aiming for a comparable number of videos, we will select, for example, the top 10 to 20 videos on YouTube for each of the search terms "covid 19" and "corona virus" using the YouTube search engine. Press briefings, live videos, and news reports will be excluded: We are interested in purposely produced video content rather than mere recordings of ongoing events. A content analysis of video formats and creative means utilized will be carried out to identify how health authorities measure up to contemporary video communication, both creatively and in reaching video consumers, identifying potential shortcomings and potential for improvement.

The Communicators' Views

Aiming to uncover structural differences and similarities in how communicators from different fields approach the creation of health science communication, we will recruit 2 participants from each of 6 different professions: public health communication, health communication, science TV and film production, video journalism, creative advertising, and social media. First, we will conduct semistructured individual interviews, identifying their approaches to video communication, their thoughts on existing COVID-19 videos, and their take on interdisciplinary collaboration. Participants will then be paired profession-wise and observed while discussing ideas for new pandemic-related video communication. Through content analysis gaps and differences in the weighting, sequencing and overall approach when tasked with creating pandemic video communication will be identified.

Turning Scientific Information Into Accessible Videos

Informed by WP1, professional audiovisual science communicators will create videos with explanations of key topics related to pandemics aimed at laypersons. This entails a focus on factors like choice of sender, ethos of the messenger, semantics, properties of context, metaphors, use of props, visual language, cinematic techniques, and editing. The output will be multiple research-based videos developed to translate pandemic scientific information in different ways, ready for experimental testing and assessment of the effect of creative choices on outcomes related to learning and attitudes toward the topic (in WP3).

This creative work will be observed, with an estimated 10 group sessions involving 1 hour of creative concept development and 10 hours of observation of the production, supplemented with 10 half-hour individual interviews. These time estimates are pragmatically chosen as estimates of how much—and how closely—the creative process and the production should be followed in order to gain sufficient insight into the process. These times are flexible and can be adjusted if needed.

Following the full video production throughout the WP will help untangle the creative process as performed by industry professionals when developing effective health science communication. The study will be guided by the creative process stages (CPS) model with its 17 stages that make up the artistic creative process [41]. This enables the study of the creative process at the macro level (the main stages) and micro level (the thinking underlying the stages) and takes an ecological approach: observation in the natural environment while the creative work is unfolding. This enables a direct, rich, and inexpensive assessment of the creative process, and the method has high ecological validity [41]. Based on the CPS model, WP2 will involve deductive development of a methodology for visualizing creative processes, so as to be able to capture, analyze, and visually present the multidimensional and intertwined aspects and stages involved in creative communication work.

Recruitment and Sampling

Participants will be recruited through the project's consortium members and its wider network, as well as through project members' own networks, including Stavanger University Hospital, Center for Developing Institutional and Home Care Services Sogn and Fjordane, Norwegian Institute of Public Health, Dagbladet, Anorak, Bulldozer Film, and Nordic Screens. Gender balance will be taken into account in the recruitment process, along with age, educational background, and occupation. All interviews and observation studies will be carried out virtually using Zoom.

Phase 3: Evaluating the Communication's Effect (WP3)

Building on the work in WP1 and WP2, WP3 will involve quantitative experiments to assess the effect of various factors related to audiovisual public communication of health science in general and pandemics in particular.

Scoping Review

First, in WP3, we will conduct a scoping review focusing on the recipients of the communication. We will search 3 main databases for public health, social sciences, and biomedical studies: PubMed, Scopus, and Embase. The search string will be designed based on the PCC (Population-Concept-Context) framework as recommended by the Johanna Briggs Institute Manual for Evidence Synthesis and scoping reviews [42] as a less restrictive alternative to the PICO (population, intervention, comparator, and outcome) mnemonic recommended for systematic reviews. The scoping review is expected to provide a comprehensive view of the state of the art and give insight about characteristics of recipients and how they can impact the outcomes of health video communication. The knowledge will contribute to the interpretation of findings from the WP's experimental studies described in the following sections.

Randomized Controlled Trials

In WP3, we will perform 2 RCTs with a factorial between-subjects experimental design including various factors related to the video communication. The first RCT will assess nonvisual factors in health communication videos such as the effect of the video super of the presenter (ie, whether the person on screen is labeled as "Professor" or "Citizen" in the text at the bottom of the screen) or whether the message presented is neutral or includes a "call to action" (ie, whether the presenter merely delivers factual information or also encourages the viewer to act). The final choice of nonvisual factors will be informed by work in the preliminary phases of the project. Assessing the effect of such nonvisual factors will be done by having an actor recite different scripts to the camera. The second RCT will use the videos developed by professional video creators in WP2.

The videos will be screened for representativeness of the population. After consent, participants will be randomized to different videos and between-factors conditions. A link will lead them to a web page where the procedure is explained. For the first RCT, this web page differs according to ascribed Source and Messenger, Content, and Engaging factors. Participants will then be shown the video(s), followed by the questionnaires. Here, we focus not only on learning outcomes and the understanding of risk at a conceptual level but also on attitudes toward the subject, behavioral change, and compliance with the message. Gender, cultural background, age, education, and work variables will be entered as covariates in the statistical analyses and in targeted subgroup analyses.

All data collection in WP3 will be conducted online, using online surveys (Survey Monkey) and collaborative tools. With a fully digital and completely automated random-number generator and allocation procedure, video link submission and measurement instruments can be performed automatically, keeping the process free from experimenter input.

Both RCTs foresee the possibility of a longitudinal study through a follow-up survey aiming to assess whether communication outcomes (eg, learning, attitudes) persist over time, evaluate the motivation of participants to rejoin the study based on the video alternative they have been exposed to, and explore whether communication outcomes are influenced by the evolution of the COVID-19 pandemic (eg, vaccination coverage, recent infection trends).

Real-Life Observational Study

Based on the RCTs, the videos used in the experiments will go through mild editing and then be screened for audiences on a mass scale, tracking spread and exploring engagement metrics such as likes and comments. This will be done in collaboration with consortium partners (eg, the national newspaper, Dagbladet, and video influencer network, Nordic Screens). Here, success criteria will be assessed through strategies such as tracking of how much of a video the viewers watch before skipping, likes, sharing, and response rates to “click here for more” links and coupled with demographic data.

Recruitment and Sampling

Participants for the RCTs will be recruited through invites from consortium partners such as the Norwegian Air Ambulance Foundation (NAAF). Statistical power depends on a number of factors. For the first RCT, a power analysis in G*Power [43] assuming a medium effect-size ($f^2 \geq 0.25$) and a full factorial design with 5 covariates in an analysis of covariance (ANCOVA) yields a conservative sample size of 401 to achieve a statistical power of 95%. For the second RCT, a smaller sample size is needed, as there are fewer covariates. The NAAF database of financial supporters consists of approximately 150,000 people from the general public, and recruitment of only a small fraction of them (1%) will thus suffice for the intended power.

Ethical Considerations

The study is approved by the Norwegian Centre for Research Data (WP1 Ref number 583192, WP2 Ref number 703372) and exempted from ethical approval from the Regional Ethical Committee.

Results

The COVCOM study is supported by the Trond Mohn Foundation (TMS) grant number TMS2020TMT10 and the University of Stavanger, Norway.

Recruitment and interviews for Phase 1 of the project ran from February 2021 to March 2021. Creative communication work in Phase 2 started in May 2021, and video production for use in the RCTs started in September 2021. Preparation for the RCTs in Phase 3 started in January 2021, with recruitment and data collection planned for 2021/2022.

Discussion

The COVCOM project will take on several grand challenges within the field of communicating science and provide evidence-based tools to the science communication toolbox. A long-term goal of the project is to contribute to the creation of a more resilient health care system by developing tailor-made communication responses for different audiences, preparing society for any future pandemic [44].

The different backgrounds and perspectives of the individuals in the COVCOM project operate along different axes in terms of norms and what constitutes a project’s success. Getting all involved individuals to pull in the same direction makes the project challenging. The serial structure of the WPs, where unexpected results and delays in one WP might affect other WPs, puts high demand on collaborative efforts in the group. Clear communication and close follow-up between WP leaders and other key personnel are central.

As the project has expected outcomes beyond the scientific community for stakeholders in terms of communicators (eg, Norwegian Institute of Public Health, Norwegian Directorate of Health, and the government) as well as recipients (eg, the general population, hospitals, and representatives of vulnerable groups) of health information, dissemination will not only be through academic publications but also focus on communicating research results to communication practitioners, health professionals, and the general public in adequate trade magazines, newspapers, and online fora. Popular science articles to disseminate results beyond academia will be encouraged and facilitated by the project manager.

Acknowledgments

We wish to acknowledge the contributions from Martha Strandos, Stig Harthug, Svein Skeie, Elisabeth Jakobsen, and Stephen Sollied in the development of the grant application to the Research Council of Norway. We also wish to acknowledge Anja Schou

Lindman, Jeffrey Braithwaite, Morten Sodemann, Johnny Advocaat-Vedvik, and Janne Espevalen who took on roles in the Strategic Advisory Board.

Authors' Contributions

JR was responsible for the application to the Research Council of Norway for funding for the COVCOM study in collaboration with SW, KB, and HT. JR, SW, KB, and HT set up the initial study design and study protocol. JR drafted the manuscript, with substantial input from SW, and revised it based on comments from all coauthors. SHB (work package [WP]1), MTS (WP2), and DAL (WP3) contributed to the study design and development of data collection tools in their respective WPs. JO and IS contributed to the study design and development of data collection tools and commented on the draft. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychological Inquiry 2000 Oct;11(4):227-268. [doi: 10.1207/s15327965pli1104_01]
2. Wong PTP, Fry PS. The human quest for meaning: A handbook of psychological research and clinical applications. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 1998:A.
3. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist 2000;55(1):68-78. [doi: 10.1037/0003-066x.55.1.68]
4. Gerber A, Broks P, Gabriel M, Lorenz L, Lorke J, Merten W, et al. Science Communication Research: an Empirical Field Analysis. Institute for Science and Innovation Communication. 2020. URL: https://sciencecomm.science/app/uploads/2020/05/Research_Field_Analysis__Science_Communication__2020__public.pdf [accessed 2022-02-21]
5. Jensen EA, Gerber A. Evidence-based science communication. Front. Commun 2020 Jan 23;4:1. [doi: 10.3389/fcomm.2019.00078]
6. Bradley DT, McFarland M, Clarke M. The effectiveness of disaster risk communication: a systematic review of intervention studies. PLoS Curr 2014 Aug 22;6:1 [FREE Full text] [doi: 10.1371/currents.dis.349062e0db1048bb9fc3a3fa67d8a4f8] [Medline: 25642365]
7. Klemm C, Das E, Hartmann T. Swine flu and hype: a systematic review of media dramatization of the H1N1 influenza pandemic. Journal of Risk Research 2014 Jun 20;19(1):1-20. [doi: 10.1080/13669877.2014.923029]
8. Barrelet C, Bourrier M, Burton-Jeangros C, Schindler M. Unresolved issues in risk communication research: the case of the H1N1 pandemic (2009-2011). Influenza Other Respir Viruses 2013 Sep 27;7 Suppl 2:114-119 [FREE Full text] [doi: 10.1111/irv.12090] [Medline: 24034495]
9. Jha A, Lin L, Short SM, Argentini G, Gamhewage G, Savoia E. Integrating emergency risk communication (ERC) into the public health system response: Systematic review of literature to aid formulation of the 2017 WHO Guideline for ERC policy and practice. PLoS One 2018 Oct 31;13(10):e0205555 [FREE Full text] [doi: 10.1371/journal.pone.0205555] [Medline: 30379900]
10. A World at Risk: Annual report on global preparedness for health emergencies - Global Preparedness Monitoring Board. Relief Web. 2019. URL: https://reliefweb.int/report/world/world-risk-annual-report-global-preparedness-health-emergencies-global-preparedness [accessed 2022-02-21]
11. Coombs WT, Holladay SJ, editors. The Handbook of Crisis Communication. Hoboken, NJ: Blackwell Publishing Ltd; 2010.
12. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996 Jan 13;312(7023):71-72 [FREE Full text] [doi: 10.1136/bmj.312.7023.71] [Medline: 8555924]
13. Lesen AE, Rogan A, Blum MJ. Science communication through art: objectives, challenges, and outcomes. Trends Ecol Evol 2016 Sep;31(9):657-660. [doi: 10.1016/j.tree.2016.06.004] [Medline: 27377601]
14. Global Call to Creatives: An Open Brief from the United Nations. Google Docs. 2020 Mar 30. URL: https://docs.google.com/presentation/d/1uR9gq1nC_ZYlGHkyaU7bRivTJ0gmtaxqZ5_4BvzcGnM/preview?slide=id.p [accessed 2022-02-21]
15. United Nations Global Call Out To Creatives - help stop the spread of COVID-19. TalentHouse. URL: https://www.talenthouse.com/i/united-nations-global-call-out-to-creatives-help-stop-the-spread-of-covid-19 [accessed 2022-02-21]
16. Cisco Annual Internet Report (2018–2023) White Paper. Cisco. 2020 Mar 09. URL: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html [accessed 2022-02-21]
17. Ceci L. Statista. 2022 Feb 07. URL: https://www.statista.com/topics/2019/youtube/ [accessed 2022-02-21]
18. 37 Mind Blowing YouTube Facts, Figures and Statistics. Merch Dope. 2021. URL: https://merchdope.com/youtube-stats/ [accessed 2021-12-10]
19. Balog-Way D, McComas K, Besley J. The evolving field of risk communication. Risk Anal 2020 Nov 20;40(S1):2240-2262 [FREE Full text] [doi: 10.1111/risa.13615] [Medline: 33084114]

20. Siegrist M. Trust and risk perception: a critical review of the literature. Risk Anal 2021 Mar 02;41(3):480-490. [doi: 10.1111/risa.13325] [Medline: 31046144]
21. Gubrium A. Digital storytelling: an emergent method for health promotion research and practice. Health Promot Pract 2009 Apr 04;10(2):186-191. [doi: 10.1177/1524839909332600] [Medline: 19372280]
22. Miller-Day M, Hecht ML. Narrative means to preventative ends: a narrative engagement framework for designing prevention interventions. Health Commun 2013 Oct;28(7):657-670 [FREE Full text] [doi: 10.1080/10410236.2012.762861] [Medline: 23980613]
23. Welbourne DJ, Grant WJ. Science communication on YouTube: Factors that affect channel and video popularity. Public Underst Sci 2016 Aug 19;25(6):706-718. [doi: 10.1177/0963662515572068] [Medline: 25698225]
24. Fischhoff B. Evaluating science communication. Proc Natl Acad Sci U S A 2019 Apr 16;116(16):7670-7675 [FREE Full text] [doi: 10.1073/pnas.1805863115] [Medline: 30478063]
25. Fischhoff B. The sciences of science communication. Proc Natl Acad Sci U S A 2013 Aug 20;110 Suppl 3(Supplement_3):14033-14039 [FREE Full text] [doi: 10.1073/pnas.1213273110] [Medline: 23942125]
26. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. New York, NY: SAGE Publications; 2017.
27. Smelser NJ, Baltes PB, editors. International Encyclopedia of the Social & Behavioral Sciences. Amsterdam, Netherlands: Elsevier; 2001.
28. Gentner D, Stevens AL. Mental Models. Milton Park, Abingdon-on-Thames, Oxfordshire United Kingdom: Taylor & Francis; 2014.
29. Meyer D, Leventhal H, Gutmann M. Common-sense models of illness: The example of hypertension. Health Psychology 1985;4(2):115-135. [doi: 10.1037/0278-6133.4.2.115]
30. Kempton W. Two theories of home heat control. Cognitive Science 1986;10(1):75-90. [doi: 10.1207/s15516709cog1001_3]
31. Giere RN, editor. Cognitive Models of Science. Minneapolis, MN: University of Minnesota Press; 1992.
32. Morgan MG. Risk Communication: A Mental Models Approach. Cambridge, MA: Cambridge University Press; 2005.
33. Jungermann H, Schütz H, Thüring M. Mental models in risk assessment: informing people about drugs. Risk Anal 1988 Mar;8(1):147-155. [doi: 10.1111/j.1539-6924.1988.tb01161.x] [Medline: 3375502]
34. Bruine de Bruin W, Bostrom A. Assessing what to address in science communication. Proc Natl Acad Sci U S A 2013 Aug 20;110 Suppl 3(Supplement_3):14062-14068. [doi: 10.1073/pnas.1212729110] [Medline: 23942122]
35. van der Sanden MCA, Meijman FJ. Evidence-based science communication. Science Communication 2016 Aug 18;25(3):272-287. [doi: 10.1177/1075547003262662]
36. Peters M, Godfrey C, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
37. Tricco AC, Langlois EV, Straus SE. Rapid reviews to strengthen health policy and systems: a practical guide. World Health Organization. 2017. URL: https://www.who.int/alliance-hpsr/resources/publications/rapid-review-guide/en/ [accessed 2022-02-21]
38. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov 01;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
39. Reed MS, Graves A, Dandy N, Posthumus H, Hubacek K, Morris J, et al. Who's in and why? A typology of stakeholder analysis methods for natural resource management. J Environ Manage 2009 Apr;90(5):1933-1949. [doi: 10.1016/j.jenvman.2009.01.001] [Medline: 19231064]
40. Mayring P. Qualitative Content Analysis: Theoretical Background and Procedures. In: Bikner-Ahsbahs A, Knipping C, Presmeg N, editors. Approaches to Qualitative Research in Mathematics Education. Advances in Mathematics Education. Dordrecht, Netherlands: Springer Publishing Company; 2015:365-380.
41. Botella M, Nelson J, Zenasni F. It is time to observe the creative process: how to use a creative process report diary. J Creat Behav 2017 Mar 06;53(2):211-221. [doi: 10.1002/jocb.172]
42. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco A, Khalil H. Chapter 11: Scoping Reviews. JBI Manual for Evidence Synthesis. 2020. URL: https://jbi-global-wiki.refined.site/space/MANUAL/3283910770/Chapter+11%3A+Scoping+reviews [accessed 2022-02-21]
43. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
44. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? The Lancet 2020 Apr;395(10231):1225-1228. [doi: 10.1016/s0140-6736(20)30627-9]

Abbreviations

ANCOVA: analysis of covariance
COVCOM: COVID communication – Fighting a pandemic through translating science
CPS: creative process stages
NAAF: Norwegian Air Ambulance Foundation

PCC: Population-Concept-Context
PICO: population, intervention, comparator, and outcome
RCT: randomized controlled trial
TMS: Trond Mohn Foundation
WHO: World Health Organization
WP: work package

Edited by T Leung; submitted 14.10.21; peer-reviewed by S Poduval, K Bredbenner; comments to author 08.12.21; revised version received 21.01.22; accepted 06.02.22; published 11.03.22.

Please cite as:
Røislien J, O'Hara JK, Smeets I, Brønnick K, Berg SH, Shortt MT, Lungu DA, Thune H, Wiig S
Creating Effective, Evidence-Based Video Communication of Public Health Science (COVCOM Study): Protocol for a Sequential Mixed Methods Effect Study
JMIR Res Protoc 2022;11(3):e34275
URL: https://www.researchprotocols.org/2022/3/e34275
doi:10.2196/34275
PMID:35147500

Protocol

Assessing Determinants of Programmatic Performance of Community Management of Malaria, Pneumonia, and Diarrhea in Children in Africa: Protocol and Data Collection for a Mixed Methods Evaluation of Integrated Community Case Management

Aliya Karim[1,2], MPH; Don de Savigny[1,2], PhD; Jean Serge Ngaima[3], MPH, MD, PhD; Daniel Mäusezahl[1,2], PhD; Daniel Cobos Muñoz[1,2], MPH, MD, PhD; Antoinette Tshefu[3], MD, PhD

[1]University of Basel, Basel, Switzerland

[2]Swiss Tropical and Public Health Institute, University of Basel, Basel, Switzerland

[3]Université de Kinshasa, Kinshasa, the Democratic Republic of the Congo

Corresponding Author:
Aliya Karim, MPH
University of Basel
Petersplatz 1
Basel, 4001
Switzerland
Phone: 41 061 207 31 11
Email: aliya.karim@unibas.ch

Abstract

Background: Integrated community case management (iCCM) is a child health program designed to provide integrated community-based care for children with pneumonia, malaria, or diarrhea in hard-to-reach areas of low- and middle-income countries. The foundation of the intervention is service delivery by community health workers (CHWs) who depend on reliable provision of drugs and supplies, consistent supervision, comprehensive training, and community acceptance and participation to perform optimally. The effectiveness of the program may also depend on a number of other elements, including an enabling policy environment, financing mechanisms from the national to the local level, data transmission systems, and appropriate monitoring and evaluation. The extent to which these factors act upon each other to influence the effectiveness and viability of iCCM is both variable and challenging to assess, especially across different implementation contexts.

Objective: In this paper, we describe a mixed methods systems-based study protocol to assess the programmatic components of iCCM that are associated with intervention effectiveness and report preliminary results of data collection.

Methods: This protocol uses a mixed qualitative and quantitative study design based on a systems thinking approach within four iCCM programs in Malawi, Democratic Republic of the Congo, and Niger State and Abia State in Nigeria. Routine monitoring data are collected to determine intervention effectiveness, namely testing, treatment, and referral outcomes. Surveys with CHWs, supervisors, and caregivers are performed to collect quantitative data on their demographics, activities, and experiences within the program and how these relate to the areas of intervention effectiveness. Focus group discussions are conducted with these stakeholders as well as local traditional leaders to contextualize these data. Key informant interviews are undertaken with national- and district-level program stakeholders and officers knowledgeable in critical program processes.

Results: We performed 3836 surveys and 45 focus group discussions of 379 participants with CHWs, supervisors, caregivers, and traditional leaders, as well as 120 key informant interviews with district- and national-level program managers, health officers, and ministry officials. Policy and program documents were additionally collected for review.

Conclusions: We expect that evidence from this study will inform child health programs and practice in low- and middle-income settings as well as future policy development within the iCCM intervention.

International Registered Report Identifier (IRRID): DERR1-10.2196/33076

(JMIR Res Protoc 2022;11(3):e33076) doi:10.2196/33076

KEYWORDS

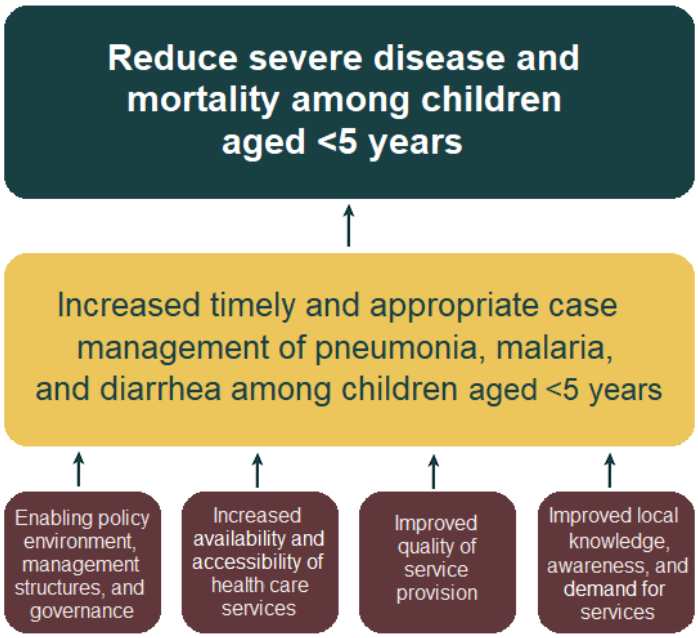
iCCM; integrated community case management; systems thinking; malaria; study design; systems methods; child health; program evaluation

Introduction

Background

Integrated community case management (iCCM) is a child health strategy that aims to provide equitable access to care for children with pneumonia, malaria, and diarrhea in some of the world's poorest and low-access areas [1]. It is focused on a community-centered model of service delivery, where trained community health workers (CHWs) provide diagnostics and treatment to children aged <5 years in assigned catchment areas. iCCM has become recognized in the global public health sphere as a critical approach to reducing child mortality, reinforced by the World Health Organization and United Nations Children's Fund in a joint statement published in 2012 [2]. Figure 1 provides a logical framework of iCCM, its pillars, and intended outputs [3].

Figure 1. Logical framework of integrated community case management.



As an integrated package of services, the intervention relies on a host of elements to ensure programmatic success. Generally, CHWs should be adequately supplied with quality commodities; receive comprehensive training, supervision, and oversight; be embedded in an environment of enabling policy, supervision, and support; and treat communities that are sufficiently knowledgeable and mobilized, among other requisites [4-7]. Although some of these aspects have been explored in isolated detail, the extent to which these act upon each other as an ecology of interdependent elements and how this influences the outcomes of the program is not completely understood. Moreover, the magnitude of their impact on effective coverage and their interactions within the health system at large are not well known; therefore, a more profound exploration of these intervention dynamics is warranted.

Because of the intricate relationships of these elements and the various ways in which different contexts shape their impact, we require an approach that elevates these properties from program isolates acting in silos to interconnected determinants behaving in a complex system. We therefore apply a systems thinking lens to our assessment of the iCCM intervention. Systems thinking is an approach to scientific inquiry that emphasizes the interrelated nature of the composite parts of a system; how these operate within a specific setting; and how these integrate, relate to, and are embedded within the surrounding environment [8].

This protocol builds upon existing systems thinking theory and a methodological approach to assess complexity in community health interventions [9]. This study protocol deals specifically with steps 2, 3, and 4 of the systems approach, namely process mapping, quantitative analysis, and qualitative analysis. In this paper, we outline the specific study design and preliminary results from the data collection process.

Study Aim

The overall aim of this study is to identify gaps and bottlenecks within critical areas and processes of 4 iCCM programs in Africa, assess factors that are attributable to these, and analyze determinants of programmatic success. Specifically, we are interested in investigating how program design, provider demographics, intervention context, service delivery practices, care recipient experiences and support, and other key health systems areas influence, drive, or are otherwise associated with program outputs of effective coverage. Effective coverage is a

metric designed to evaluate the health gain that is experienced by a population, given a certain set of conditions of the health service delivery model [10].

Study Design

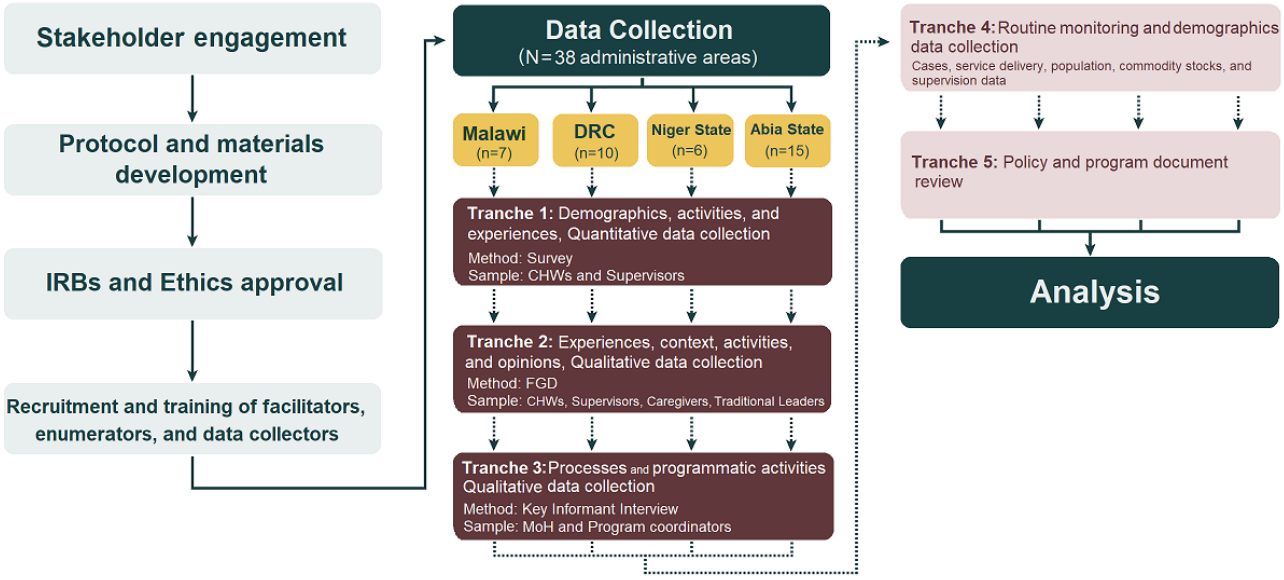
We applied a mixed methods concurrent triangulation design with an intersecting systems theoretical lens [9-12]. Data collection occurs in five discrete tranches:

- Tranche 1: surveys conducted with service providers and recipients of the iCCM intervention (quantitative)
- Tranche 2: focus group discussions (FGDs) conducted with service providers and recipients of the iCCM intervention (qualitative)
- Tranche 3: key informant interviews with program managers, ministry of health (MoH) officers, and other stakeholders at national and district level involved in iCCM activities (qualitative)
- Tranche 4: routine data on cases, service delivery, population and demographics, commodity stocks, reporting, and supervision (quantitative)
- Tranche 5: a policy and document review

An illustration of these tranches is provided in Figure 2.

The study was performed recognizing the existence of different levels of outputs and subsequent outcomes for analysis that will be emergent during the data collection and preliminary analysis period. For example, it is generally accepted that frequent supervision of CHWs is a precondition to successful implementation of community-based programs and thus may serve as an endogenous factor or a process in itself for analysis [13]. Analogously, we may also wish to examine the extent to which the frequency, duration, or other aspect of supervision may affect other activities, elements, or outcomes related to iCCM, such as downstream availability of drug stocks or community ownership of the intervention. The study appreciates that these elements that comprise iCCM do not occur in a vacuum and iteratively influence each other, necessitating a holistic approach to the analysis of the intervention.

Figure 2. Schematic representation of the research process. CHW: community health worker; DRC: Democratic Republic of the Congo; FGD: focus group discussion; IRB: institutional review board.



Methods

Study Setting

The study was implemented in 4 large, geographically distinct territories of 3 sub-Saharan African countries: Malawi, Democratic Republic of the Congo (DRC), and Niger State and Abia State in Nigeria. Within these countries, the iCCM program was administered by the respective state and national ministries of health, with implementation support from local nongovernmental organizations (NGOs) and oversight by the World Health Organization. The program was rolled out across 7 districts in the central and northern regions of Malawi; 10 health zones in the province of Tanganyika in DRC; 6 local government areas (LGAs) in Niger State, Nigeria; and 15 LGAs in Abia State, Nigeria. At the time of the study, iCCM implementation had been ongoing for at least two years. Data were expected to be collected across all these administrative areas where iCCM was implemented.

Data

The study required both qualitative and quantitative primary and secondary data from different stakeholders and experts within key thematic areas. Primary data were collected from three sources: (1) surveys with CHWs and their supervisors; (2) FGDs with CHWs, supervisors, caregivers and traditional leaders; and (3) key informant interviews with a variety of program partners and country-level stakeholders. An additional survey of caregivers was conducted in Abia State, Nigeria.

We also obtained secondary data to inform the analysis, namely routine monitoring records; household surveys of care-seeking behavior; routine health facility and hospital data (where possible); catchment area demographics, geographical data, and population statistics for communities, health facility areas, and

districts; and documentation on program guidelines, manuals, and country child health policies.

Tool Content and Development

We began with protocol development and drafting of data collection and training materials, specifically survey instruments, FGD guides, and key informant interview templates. We developed 9 survey instruments, 12 FGD guides, and 75 interview guides, varying according to country context. The survey instruments contained between 73 and 137 questions each, the FGD tools contained between 66 and 108 questions per stakeholder group, and the key informant interview guides contained between 25 and 69 questions per informant. These were reviewed for quality and comprehensiveness by funding partners and country team members, which included local academic institutions, country technical representatives, program managers, and MoH members within the country's child health network. These individuals were identified by the research team through a snowball approach. Each item and country protocol was developed in English and translated into 1 of 10 languages by at least two trained translators recruited either through institutional or local contacts. Translations for the surveys were then subject to a roundtable in-country participatory translation workshop consisting of at least 10 country actors, including program managers or stakeholders, to verify accuracy, relevance, audience appropriateness, and adherence to original question intent.

Surveys and FGDs conducted with CHWs and supervisors focused on their personal profiles, core iCCM processes, inputs, and activities; the caregivers and communities they served; and motivations and opinions. Those conducted with caregivers centered on their knowledge, care-seeking behavior, and adherence, whereas the FGDs conducted with traditional leaders discussed community support, buy-in, and local ownership. The key informant guides collected data on specific iCCM processes, policy, and management of the program, including but not limited to: procurement and distribution of key iCCM commodities; health management information systems; child health policy and program financing processes; human resources; local mobilization; and program coordination, integration, and sustainability at different administrative levels.

Survey Participants and Eligibility Criteria

The targeted survey participants were CHWs and supervisors. An additional survey with caregivers in Abia State was warranted by the local interest in a statistical measure of the perceived applicability of the intervention in a setting that was geographically considerably smaller and more densely populated than those of the other 3 programs. Eligible CHWs were those who had been trained in iCCM and had participated in the implementation of activities in their communities, whether belonging to a pre-existing health worker cadre or recruited for the specific purpose of this program. Eligible supervisors were health facility attachés who were charged with the supervision of at least one CHW throughout the life of the program. Caregivers in Abia State were eligible participants if they were the primary caretaker of a child aged 2-59 months and resided within the catchment areas of iCCM implementation. Stratified systematic sampling was used for caregiver participant selection. For the CHW and supervisor surveys, bias was mitigated by enumerators cycling through a complete list of all CHWs and supervisors operating in the 38 administrative areas, rather than relying upon local actors to choose participants for the survey.

Sampling Population

Survey

A single-stage cluster sampling scheme was used for the surveys. All administrative divisions where the iCCM program was implemented in each of the 4 country areas were expected to be sampled. This treated each administrative area as its own independent sampling frame, allowing for full saturation across implementation areas and enabling robust analysis of district-specific effects. The sampling frame for supervisors was at the country area level, with even distribution across all administrative areas. CHW and supervisor population data were sourced from national statistics offices or program records. Survey sample sizes were calculated from these based on probability proportional to size sampling, with aimed representation within a CI of 95% [14]. Figure 3 illustrates the formula for the sample size calculation used to obtain a statistically viable sample of the population within the expected CI. The expected sample sizes and achieved participation for each stakeholder are listed in the *Results* section.

Figure 3. Sample size calculation for surveys.

$$Sample\ Size = \frac{Z^2 p(1-p)}{C^2}$$

Where:

Z= 1.96 (for 95% confidence level)

p= .5 (expected proportion)

C= .05 (margin of error)

Correcting for finite populations:

$$New\ Sample\ Size = {}^{ss}\Big/\left(\frac{ss-1}{population}+1\right)$$

FGD Procedures

The FGDs aimed for an average of 8 participants per focus group, which is considered optimal for clarity and representative participation [15]. Saturation of themes is generally expected to occur after 2-3 focus groups; we therefore aimed for 4 FGDs per stakeholder group for the 3 stakeholder groups, where each trine was derived from 1 administrative area [16]. In some country areas, we aimed to hold 1 FGD with traditional leaders to supplement community context.

Team Recruitment, Enumerator, and Facilitator Training

Enumerator and FGD facilitator guidelines, protocols, and other materials were developed as part of the enumerator training package. We recruited 8-10 survey enumerators per country area according to predetermined criteria with the support of local research teams. These enumerators participated in a round of training resulting in the final selection process; this was followed by a 10-day, 80-hour training course in question content, survey strategy, tablet use, and surveying etiquette. At the end of the course, each enumerator performed 2 test runs of all questions with the survey director from the Swiss Tropical and Public Health Institute and codirectors, usually from the MoH, the supporting NGO, and local university partner. All survey questions were presented with multiple response scenarios, and enumerators were marked based on performance. This also allowed survey supervisors to address any shortcomings in survey structure and design. FGD facilitators were recruited and trained in a similar fashion, where teams consisted of a timekeeper, notetaker, and discussion facilitator.

We adhered to strict compensation guidelines that aim to ensure fair and advantageous remuneration set forth by the International Labor Organization's fundamental conventions [17-19]. As the recruits were usually young public health professionals, they were also given the opportunity to participate in various capacity-building sessions held as part of the research program.

Surveys and FGD Rollout

The surveys were piloted over the course of a week, after which final adjustments were made to the questionnaires and enumeration procedures. The official survey data collection lasted on average 2 months per country area. Surveys were carried out with CHWs and their supervisors by trained enumerators by telephone for practicality and cost-effectiveness. CHWs and supervisors were alerted to the study by district health officers at the behest of the investigators and country program managers. Contact information of all CHWs and supervisors was thereafter provided for participant recruitment. Recruitment for the caregiver survey (Abia State) was facilitated by village heads and randomly selected individuals from identified communities. The survey was conducted with electronic data collection using Open Data Kit (ODK), which was programmed on tablets during the tool development stage. Phone credit and headphones were provided to survey participants to carry out the survey. Data were uploaded daily to the secure institutional server. Surveys were conducted in all administrative areas of the iCCM country sites participating in the program.

Establishing contact with CHWs and supervisors proved challenging in some settings as a result of ubiquitous network issues, respondent availability, and general electricity outages. Furthermore, some CHWs had either moved, transferred, resigned, or were deceased, and many had been added who were not updated in the program files. We attempted to control for bias associated with poor network coverage by ensuring a proportional distribution of hard-to-reach (HTR) and easily reached respondents. HTR respondents, meaning those with whom contact was unsuccessful even after 3 attempts, were registered as HTR and were targeted specifically through their supervisor or through the administrative area focal person, who encouraged them to move toward areas of reception or regroup directly at the health facility for the survey. Some enumerators in Abia State traveled to their assigned areas to conduct interviews in person because LGAs were clustered closely together.

The survey interviews typically lasted for 70 minutes each for both survey types. The length of the interview did not deter participants during piloting; therefore, the length was retained. Targeted sample sizes were reached across all country areas, and survey participation was high; 100% of all CHWs and supervisors reached consented to participate. Participants expressed appreciation in having the opportunity to voice their opinions, concerns, and experiences.

For each administrative area, a group of caregivers, supervisors, and CHWs each was chosen to participate in FGDs, averaging 3 FGDs per administrative area and 12 FGDs per country area with these stakeholders. In Niger State and Abia State, additional FGDs were held with local development committees and traditional leaders because of specific interest in these country sites. Participants were recruited with assistance from district health officers and NGO program officers, with a recruitment strategy based on a random sample of participants across communities, health facilities, and districts. The discussions typically lasted between 1½ hours and 2½ hours and up to 3 hours. They were intensive, and hourly breaks were necessary; a stipend, drinks, and food were provided to participants and facilitators. Transport to and from data collection areas was organized either independently, in collaboration with the local ministry or NGO, or through the local research team.

Quality Control

After the pilots and official launch of the survey, further steps were taken to ensure survey quality. In addition to daily supervision, the principal investigator held triweekly individual meetings with each of the enumerators to discuss results, obtain feedback, and check survey entries. Enumerators were also instructed to obtain correct administrative area, community, and population information for all CHWs and supervisors to revise existing records that were often misaligned with project areas. These data points were corrected in a master registry on institutional servers. Each survey that was conducted was checked against its original entry in the master database, where any errors were rectified. In addition, various roundtable discussions were held and recorded among all enumerators to review emerging themes, outlier anecdotes, concerns, shortcomings, or administrative tasks. Enumerators were also

required to take notes and report on specific questions, which were shared with the survey director and recorded in the master file as auxiliary information.

Ethics Approval and Considerations

This study was approved in 2016 by the institutional review boards of Malawi (NHSRC 16/6/1610), DRC (ESP/CE/046/2016), and Nigeria (NHREC 01012007-15122016), as well as the Ethics Commission of Northwest and Central Switzerland (EKNZ REQ-2016-00478). All subjects provided informed consent before participating in interviews, discussions, or surveys. All individual data used were anonymized upon collection or extraction; routine monitoring data used existed within a deidentified format. All iCCM programs operated under the supervision of each country's respective MoH and aligned with national standards of care.

Results

Overview

Data collection was undertaken from 2016 to 2017. The following sections outline the experience of preliminary data collection and cleaning in detail. Results and analysis of the implementation of this protocol from each country are in preparation.

Surveys and FGDs

We successfully collected survey data from CHWs and supervisors across all 38 administrative divisions where iCCM was implemented within the 4 country sites. A total of 3836 surveys were completed. FGDs were conducted in 3-5 administrative areas of each country site resulting in 45 FGDs comprising 379 participants. Statistics on survey and FGD implementation are provided in Table 1.

Table 1. Data collection for surveys and focus group discussions (FGDs).

Country site	Administrative area[a]	Survey participants, N=3836			FGD participants, N=378 (45 groups)			
		CHW[b]	Supervisor	Caregiver	CHW	Supervisor	Caregiver	Traditional leader
Malawi								
	Dedza[c]	118	32	N/A[d]	8	8	10	N/A
	Lilongwe	128	42	N/A	__[e]	—	—	N/A
	Mzimba North[c]	81	20	N/A	8	8	9	N/A
	Nkhatabay[c]	29	42	N/A	8	8	8	N/A
	Ntcheu[c]	94	8	N/A	8	8	9	N/A
	Ntchisi	88	8	N/A	—	—	—	N/A
	Rumphi	67	11	N/A	—	—	—	N/A
Total		605	163	N/A	32	32	36	N/A
Sample size, 95% CI		305[f], 598[g]	139[f]	—	—	—	—	—
Democratic Republic of the Congo								
	Ankoro	7	9	N/A	—	—	—	N/A
	Kabalo	5	14	N/A	—	—	—	N/A
	Kalemie[c]	75	13	N/A	7	7	12	N/A
	Kansimba[c]	21	13	N/A	8	5	9	N/A
	Kongolo	48	17	N/A	—	—	—	N/A
	Manono	12	10	N/A	—	—	—	N/A
	Mbulula	8	12	N/A	—	—	—	N/A
	Moba[c]	41	11	N/A	8	9	10	N/A
	Nyemba[c]	53	11	N/A	8	8	10	N/A
	Nyunzu	12	7	N/A	—	—	—	N/A
Total		282	117	N/A	31	29	41	N/A
Sample size, 95% CI		281[f]	114[f]	—	—	—	—	—
Niger State, Nigeria								
	Edati	89	15	N/A	—	—	—	—
	Lapai[c]	112	25	N/A	8	8	9	—
	Mariga	176	26	N/A	—	—	—	—
	Paikoro	84	15	N/A	—	—	—	—
	Rafi[c]	141	26	N/A	8	8	9	—
	Rijau[c]	233	23	N/A	8	8	8	9
Total		835	130	N/A	24	24	26	9
Sample size, 95% CI		298[f], 813[g]	117[f]	—	—	—	—	—
Abia State, Nigeria								
	Arochukwu	41	12	—	—	—	—	—
	Bende[c]	101	16	—	8	8	9	—
	Ikwuano[c]	109	15	—	8	8	10	—
	Isialangwa North	24	3	—	—	—	—	—

Country site	Administrative area[a]	Survey participants, N=3836			FGD participants, N=378 (45 groups)			
		CHW[b]	Supervisor	Caregiver	CHW	Supervisor	Caregiver	Traditional leader
	Isialangwa South	99	11	—	—	—	—	—
	Isuikwuato	54	10	—	—	—	—	—
	Obingwa[c]	118	14	—	—	—	—	9
	Ohafia	101	13	—	—	—	—	—
	Osisioma[c]	105	16	—	8	8	9	—
	Ugwanagbo	3	1	—	—	—	—	—
	Ukwa East[c]	12	3	—	—	—	9	—
	Ukwa West	10	1	—	—	—	—	—
	Umuahia North	39	11	—	—	—	—	—
	Umuahia South	40	7	—	—	—	—	—
	Umunneochi	70	5	—	—	—	—	—
Total		926	138	640	24	24	37	9
Sample size, 95% CI		294[f], 892[g]	112[f]	385[f]	—	—	v	—

[a]Administrative areas maintain different nomenclature across countries; accordingly, these subdivisions are *districts* in Malawi, *zones* in Democratic Republic of the Congo, and *local government areas* in Nigeria.

[b]CHW: community health worker.

[c]Administrative areas where FGDs were also conducted.

[d]N/A: not applicable.

[e]Focus group discussions were only conducted in 3 to 4 administrative areas per country site. Caregiver surveys were only collected in Abia State with a representative sample distribution computed at state level.

[f]Minimum sample sizes displayed are national or state level (according to territory distinction); achieved across all territories.

[g]Minimum sample sizes displayed are administrative area level (district, local government area, or zone); specific to CHWs, achieved for Malawi, Niger State, and Abia State.

Key Informant Interviews

We conducted 120 key informant interviews with a variety of actors fulfilling different roles related to iCCM, including district and zonal government stakeholders; national ministry representatives; national program managers, coordinators, and NGO officers; data managers, pharmacists, and central medical stores affiliates; and field implementers. Program managers and local teams assisted in identifying and requesting participation of key informants. Some follow-up key informant interviews were necessary because of scheduling conflicts. A list of these is provided in supplementary Tables S1-S4 (Multimedia Appendix 1).

Routine Monitoring, Demographic and Geographic Data, and Document Review

Routine monitoring data were extracted from independent databases held by the iCCM implementing agent (NGO) or the MoH. Country-owned data on care-seeking behavior before and after program implementation were also provided. Permission was requested to access the country's national health system or to obtain all records and data used, which were usually submitted in digitized format. Country demographic data were obtained from national statistics offices directly or from programmatic registers. These entities also provided access to policy and program documentation.

Analysis

Statistical data were extracted, cleaned, coded, and analyzed using Stata software (version 15.0; StataCorp LLC). We used nonparametric tests, specifically the Fisher exact test, to assess change and magnitude of effective coverage at critical areas of the iCCM program. Bivariate and multivariate logistic and linear regression analyses were performed to assess factors associated with these critical areas, as well as other emergent programmatic factors. Structural equation modeling was used to measure relationships between variables and latent constructs. Comparative descriptive analyses on routine monitoring data and forms were also performed to evaluate data strength, systems performance, and health provider activities.

Interviews with key informants and FGDs were facilitated, recorded, transcribed, and translated by trained personnel. Transcripts were coded for emergent themes and analyzed according to the Framework Method [20]. Process mapping of key iCCM processes and their bottlenecks was performed using Bizagi software. Causal loop diagrams were used to explore and illustrate programmatic relationships. Full details of the analysis plan have been published separately [8].

Discussion

Strengths and Limitations

This systems evaluation protocol benefits from design elements that give ample consideration to the complexity inherent to health systems and programs in low- and middle-income countries. It steps outside the collection of isolated indicators or siloed phenomena to allow exploration of relationships among tangible and intangible facets of the iCCM intervention. As it concentrates not only on program performance, but also on systems performance, it inherently takes into account appropriateness, applicability, and sustainability of the intervention, enabling simultaneously a focused study and an extensive evaluation.

The study protocol includes certain limitations. Data collection of this order is time consuming and can require considerable human resources in both fieldwork and analyses. In addition, because systems studies are necessarily large in scope, parsing apart emergent health systems themes for evaluation can be challenging, while introducing a data burden to the research plan. Potential solutions to address these issues could be using a 2-stage cluster sampling method and samples representative of the national or state level as opposed to administrative areas, a lower threshold for qualitative data saturation (ie, fewer FGDs per key actor and a smaller sample of administrative areas), a narrower focus on specific systems thematic areas, condensed surveys and discussion guides, and remote interviews for those key informants with stable internet access.

Conclusions

Our proposed study, which capitalizes on the application of the same child health program at wide scale in different contexts, has the potential to advance our understanding of how iCCM programs and other community-based health interventions can operate more effectively and sustainably, what unintended consequences might emerge from their application, and what factors drive the key outputs and impacts of the program. These are essential to achieving Sustainable Development Goal 3.2 of reducing child mortality.

Acknowledgments

This study was reviewed and funded by the World Health Organization through a grant from the government of Canada. The authors wish to thank Prof Christian Schindler and Prof Mitchell Weiss who reviewed the design for merging quantitative and qualitative methods. The authors wish to thank the reviewers for their thoughtful and thorough commentary and suggestions.

Authors' Contributions

AK led the research project and developed the protocol and methods with DDS. AK led data collection with JSN and AT. DDS, DCM, and DM supported the data collection and review processes. All listed authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Tables S1-S4: Key informant interviews conducted in Malawi, Democratic Republic of the Congo, and Nigeria.
[PDF File (Adobe PDF File), 290 KB - resprot_v11i3e33076_app1.pdf]

References

1. Marsh D, Hamer D, Pagnoni F, Peterson S. Introduction to a special supplement: evidence for the implementation, effects, and impact of the integrated community case management strategy to treat childhood infection. Am J Trop Med Hyg 2012 Nov;87(5 Suppl):2-5 [FREE Full text] [doi: 10.4269/ajtmh.2012.12-0504] [Medline: 23136271]
2. Young M, Wolfheim C, Marsh DR, 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]
3. CCM DRC Program Brief. URL: https://www.childhealthtaskforce.org/sites/default/files/2019-07/DRC%20iCCM%20of%20PMD%20Program%20Brief%28PSI%29.pdf [accessed 2022-01-26]
4. Theodoratou E, Al-Jilaihawi S, Woodward F, Ferguson J, Jhass A, Balliet M, et al. The effect of case management on childhood pneumonia mortality in developing countries. Int J Epidemiol 2010 Apr;39 Suppl 1:i155-i171 [FREE Full text] [doi: 10.1093/ije/dyq032] [Medline: 20348118]
5. Chinbuah M, Kager P, Abbey M, Gyapong M, Awini E, Nonvignon J, et al. Impact of community management of fever (using antimalarials with or without antibiotics) on childhood mortality: a cluster-randomized controlled trial in Ghana. Am J Trop Med Hyg 2012 Nov;87(5 Suppl):11-20 [FREE Full text] [doi: 10.4269/ajtmh.2012.12-0078] [Medline: 23136273]
6. Bosch-Capblanch X, Marceau C. Training, supervision and quality of care in selected integrated community case management (iCCM) programmes: a scoping review of programmatic evidence. J Glob Health 2014 Dec;4(2):020403 [FREE Full text] [doi: 10.7189/jogh.04.020403] [Medline: 25520793]

7. Rasanathan K, Bakshi S, Rodriguez DC, Oliphant NP, Jacobs T, Brandes N, et al. Where to from here? Policy and financing of integrated community case management (iCCM) of childhood illness in sub-Saharan Africa. J Glob Health 2014 Dec;4(2):020304 [FREE Full text] [doi: 10.7189/jogh.04.020304] [Medline: 25520790]
8. Wilkinson J, Goff M, Rusoja E, Hanson C, Swanson RC. The application of systems thinking concepts, methods, and tools to global health practices: an analysis of case studies. J Eval Clin Pract 2018 Jun;24(3):607-618. [doi: 10.1111/jep.12842] [Medline: 29152819]
9. Karim A, Cobos Munoz D, Mäusezahl D, de Savigny D. A systems approach to assessing complexity in health interventions: an effectiveness decay model for integrated community case management. Glob Health Action 2020 Dec 31;13(1):1794106 [FREE Full text] [doi: 10.1080/16549716.2020.1794106] [Medline: 32772891]
10. Ng M, Fullman N, Dieleman JL, Flaxman AD, Murray CJ, Lim SS. Effective coverage: a metric for monitoring Universal Health Coverage. PLoS Med 2014 Sep;11(9):e1001730 [FREE Full text] [doi: 10.1371/journal.pmed.1001730] [Medline: 25243780]
11. Montiel-Overall P. Designing and Conducting Mixed Methods Research. By John W. Creswell and Vicki L. Plano Clark. Thousand Oaks, CA: Sage Publications, 2007. 274 pp. ISBN: 1-4129-2791-9; ISBN: 1-4129-2792-7. Library & Information Science Research 2007 Sep;29(3):434-436. [doi: 10.1016/j.lisr.2007.02.003]
12. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, California: SAGE Publications; 2018.
13. Ameha A, Karim A, Erbo A, Ashenafi A, Hailu M, Hailu B, et al. Effectiveness of supportive supervision on the consistency of integrated community cases management skills of the health extension workers in 113 districts of Ethiopia. Ethiop Med J 2014 Oct;52 Suppl 3:65-71. [Medline: 25845075]
14. Innocenti F, Candel MJ, Tan FE, van Breukelen GJ. Optimal two-stage sampling for mean estimation in multilevel populations when cluster size is informative. Stat Methods Med Res 2021 Feb;30(2):357-375 [FREE Full text] [doi: 10.1177/0962280220952833] [Medline: 32940135]
15. Nyumba TO, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: insights from two decades of application in conservation. Methods Ecol Evol 2018 Jan 11;9(1):20-32 [FREE Full text] [doi: 10.1111/2041-210x.12860]
16. Hennink MM, Kaiser BN, Weber MB. What influences saturation? Estimating sample sizes in focus group research. Qual Health Res 2019 Aug;29(10):1483-1496 [FREE Full text] [doi: 10.1177/1049732318821692] [Medline: 30628545]
17. Hovary C. ILO Convention 95 The Protection of Wages Convention. In: International and European Labour Law. Baden-Baden, Germany: Nomos; 1949.
18. Teklè T. ILO Convention 100 Equal Remuneration Convention. In: International and European Labour Law. Baden-Baden, Germany: Nomos; 1951.
19. International Labour Office. Equality at Work: The Continuing Challenge : Global Report Under the Follow-up to the ILO Declaration on Fundamental Principles and Rights at Work. Geneva, Switzerland: International Labour Office; 2011.
20. 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 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]

Abbreviations

CHW: community health worker
DRC: Democratic Republic of the Congo
FGD: focus group discussion
HTR: hard-to-reach
iCCM: integrated community case management
LGA: local government area
MoH: ministry of health
NGO: nongovernmental organization
ODK: Open Data Kit

Edited by G Eysenbach; submitted 25.08.21; peer-reviewed by A Safarnejad, J Carter; comments to author 01.11.21; revised version received 11.11.21; accepted 16.11.21; published 14.03.22.

Please cite as:
Karim A, de Savigny D, Ngaima JS, Mäusezahl D, Cobos Muñoz D, Tshefu A
Assessing Determinants of Programmatic Performance of Community Management of Malaria, Pneumonia, and Diarrhea in Children in Africa: Protocol and Data Collection for a Mixed Methods Evaluation of Integrated Community Case Management
JMIR Res Protoc 2022;11(3):e33076
URL: https://www.researchprotocols.org/2022/3/e33076
doi:10.2196/33076
PMID:35285813

Protocol

The Impact of Hypoglycemia on Productivity Loss and Utility in Patients With Type 2 Diabetes Treated With Insulin in Real-world Canadian Practice: Protocol for a Prospective Study

Veronique Lambert-Obry[1], MSc; Jean-Philippe Lafrance[2], MD, MSc; Michelle Savoie[1], MBA, MPH, LLM, PhD; Jean Lachaine[1], PhD

[1]Faculty of Pharmacy, Université de Montréal, Montreal, QC, Canada
[2]Faculty of Medicine, Université de Montréal, Montreal, QC, Canada

Corresponding Author:
Veronique Lambert-Obry, MSc
Faculty of Pharmacy
Université de Montréal
2940 chemin de Polytechnique
Montreal, QC, H3T 1J4
Canada
Phone: 1 514 343 6111 ext 4873
Email: veronique.lambert-obry@umontreal.ca

Abstract

Background: Type 2 diabetes mellitus (T2DM) imposes a substantial burden owing to its increasing prevalence and life-threatening complications. In patients who do not achieve glycemic targets with oral antidiabetic drugs, the initiation of insulin is recommended. However, a serious concern regarding insulin is drug-induced hypoglycemia. Hypoglycemia is known to affect quality of life and the use of health care resources. However, health economics and outcomes research (HEOR) data for economic modelling are limited, particularly regarding utility values and productivity losses.

Objective: This real-world prospective study aims to assess the impact of hypoglycemia on productivity and utility in insulin-treated adults with T2DM from Ontario and Quebec, Canada.

Methods: This noninterventional, multicenter, 3-month prospective study will recruit patients from 4 medical clinics and 2 endocrinology or diabetes clinics. Patients will be identified using appointment lists and enrolled through consecutive sampling during routinely scheduled consultations. To be eligible, patients must be aged ≥18 years, diagnosed with T2DM, and treated with insulin. Utility and productivity will be measured using the EQ-5D-5L questionnaire and Institute for Medical Technology Assessment Productivity Cost Questionnaire, respectively. Questionnaires will be completed 4, 8, and 12 weeks after recruitment. Generalized estimating equation models will be used to investigate productivity losses and utility decrements associated with incident hypoglycemic events while controlling for individual patient characteristics. A total of 500 patients will be enrolled to ensure the precision of HEOR estimates.

Results: This study is designed to fill a gap in the Canadian evidence on the impact of hypoglycemia on HEOR outcomes. More specifically, it will generate productivity and utility inputs for the economic modeling of T2DM.

Conclusions: Insulin therapy is expensive, and hypoglycemia is a significant component of economic evaluation. Robust HEOR data may help health technology assessment agencies in future reimbursement decision-making.

International Registered Report Identifier (IRRID): PRR1-10.2196/35461

(JMIR Res Protoc 2022;11(3):e35461) doi:10.2196/35461

KEYWORDS

real-world evidence; work productivity; health-related quality of life; diabetes; hypoglycemia

Introduction

Background

Diabetes imposes a substantial burden because of its increasing prevalence and life-threatening complications. According to Diabetes Canada, 1 in 3 Canadians is living with diabetes or prediabetes, among which type 2 diabetes mellitus (T2DM) accounts for 90% to 95% of cases [1]. Glycemic control is key to diabetes management, and several types of oral therapies have been approved for T2DM. In patients who do not achieve glycemic targets on oral antidiabetic drugs, initiation of insulin therapy is recommended [1]. However, a serious concern regarding insulin is drug-induced hypoglycemia [2]. Hypoglycemic events are a major obstacle to the achievement of glycemic targets and represent a challenge for both patients and physicians. The severity of hypoglycemia is defined by clinical manifestations ranging from mild symptoms to seizure and coma [2].

Current State of Knowledge

Hypoglycemia is known to affect quality of life and the use of health care resources. However, health economics and outcomes research (HEOR) data for economic modelling are limited, particularly regarding utility values and productivity losses. O'Reilly et al [3] recently published data on health care costs associated with hypoglycemia based on the Canadian cohort of the International Hypoglycemia Assessment Tool study. Although relevant data on direct costs were provided, productivity losses were subject to recall bias (1-year recall) and were limited to absenteeism (excluding presenteeism and unpaid work). Regarding utility values, a highly cited literature review by Beaudet et al [4] identified utilities for 20 modeling complications associated with T2DM. For hypoglycemia, the review found only 1 study with utility estimates suitable for economic modeling, that is, considering the number of events (rather than the presence vs absence of hypoglycemia). Economic modeling of T2DM is particularly challenging because multiple interrelated organ systems are involved over a long period, with numerous risk factors. For an adverse event such as hypoglycemia, HEOR data should depend on the number of episodes to fit T2DM economic models, as in the study by Currie et al [5], the current preferred source for utility values for hypoglycemia. Although the study by Currie et al [5] is the source recommended by several health technology assessment (HTA) agencies, it is criticized by the National Institute for Health and Care Excellence (NICE) for being subject to selection bias and recall bias [6,7]. It should be noted that Beaudet et al [4] also highlighted that variability around the utility estimates (eg, SE) is not always reported, thereby limiting future sensitivity analyses in cost-utility analyses (CUAs).

Canadian data on productivity losses due to hypoglycemia include a 2005 survey of 133 patients with T2DM of whom up to 9% and 26% missed work or studies following nonsevere hypoglycemic events (NSHEs) and severe hypoglycemic events (SHEs), respectively [8]. A web-based survey included 150 patients with T2DM having a nocturnal NSHE in the previous month. Among the 87 working respondents, 15 (17%) reported an average of 3.5 hours of lost work [9]. In the Canadian Hypoglycemia Assessment Tool cohort, 6% (8/134), 3.8% (5/134), and 7.5% (10/134) of T2DM workers reported on average 2.9 days taken off (SD 21.2), 3.8 days arriving late (SD 8.8), and 1.7 days leaving early (SD 0.7), respectively, based on a 1-year recall [3]. As for utility, time trade-off (TTO) values were elicited from 51 Canadian respondents with diabetes and 79 respondents from the general population. The mean utility ranged between 0.85 and 0.94, 0.77 and 0.90, and 0.66 and 0.83 for rare hypoglycemic events (quarterly), intermediate (monthly), and frequent (weekly) health states, respectively. In a multivariate linear regression, the estimated utilities associated with a single hypoglycemic event (all types combined) were −0.0033 and −0.0032 for respondents with diabetes and the general population, respectively [10]. Another TTO study estimated utility decrements of 0.006, 0.008, 0.059, and 0.062 for daytime NSHEs, nocturnal NSHEs, daytime SHEs, and nocturnal NSHEs, respectively [11]. A third TTO study including patients with diabetes and respondents from the general population estimated utility decrements ranging from 0.0028 to 0.0056, 0.0076, 0.0592 to 0.0726, and 0.0616 to 0.0826 for daytime NSHEs, nocturnal NSHEs, daytime SHEs, and nocturnal SHEs, respectively [12]. No Canadian studies that met the Canadian Agency for Drugs and Technologies in Health (CADTH) recommendation to use an indirect method for utility measurement (eg, EQ-5D) were identified.

Study Rationale and Relevance

The population at risk of drug-induced hypoglycemia is significant, with nearly 20% to 30% of patients with T2DM requiring insulin [13,14]. In the last 5 years, CADTH has reviewed 9 economic evaluations in T2DM, among which 3 were CUAs for insulin therapies [15-17]. Considering the growing interest in HEOR data in insulin-treated patients with T2DM, there is a need for robust estimates. Data on productivity losses are lacking, which are particularly meaningful in Quebec, where the *Institut national d'excellence en santé et en services sociaux* preconizes economic evaluations by adopting a societal perspective. In addition, in 3 recent pharmacoeconomic reports on insulin therapies for T2DM, CADTH highlighted the uncertainty around utility values for hypoglycemia [15-17]. The committee described the current evidence on the impact of hypoglycemia on utility as limited and of low quality. In CADTH's reanalyses for the 3 CUAs, variation in utility values for hypoglycemia led to a significant change in the estimates of quality-adjusted life-years gains. The incremental cost-effectiveness ratios were very sensitive to any changes in utility decrements for hypoglycemic events and were therefore determined to be an important driver of the results [15-17]. Thus, there is a need for HEOR data that are robust enough to make an informed decision with reasonable uncertainty. The aim of this real-world prospective study is to assess the impact of hypoglycemia on productivity and utility in insulin-treated adults with T2DM from Ontario and Quebec, Canada.

Methods

Research Purpose and Study Design

This noninterventional, multicenter, 3-month prospective study is designed to collect HEOR inputs for future economic

modeling. This study will generate descriptive HEOR estimates that can be incorporated into T2DM models, along with precision parameters (ie, CIs and SEs) to evaluate uncertainty in CUAs in sensitivity analyses. Considering the nature of the disease, a longitudinal design with repeated measures was deemed appropriate to limit confounding and increase the quality of evidence.

Study Population

In Canadian clinical practice, patients with T2DM who are on insulin can be followed up by a family physician or a specialist [18]. Therefore, patients will be recruited from 4 medical clinics and 2 endocrinology or diabetes clinics. To increase generalizability, patients will be recruited from the 2 largest Canadian provinces, Ontario and Quebec. Patients will be identified using appointment lists and enrolled through consecutive sampling during routinely scheduled consultations. To be eligible, patients must be aged ≥18 years, diagnosed with T2DM, and treated with insulin. There are no restrictions on insulin regimens (eg, short-acting, long-acting, or mixed insulin). Given the lack of reimbursement, the use of insulin pumps is limited in the treatment of T2DM [19]. Patients must also be able to understand and read English or French and provide informed consent. Patients will be excluded if they participate in a clinical trial. A screening log will be used to document the participation rates.

Definition of Hypoglycemia

According to the Canadian Diabetes Association and the American Diabetes Association (ADA), hypoglycemia is defined as plasma glucose concentrations of ≤3.9 mmol/L (≤70 mg/dL) [2,20]. To reflect a real-world setting, the definition of hypoglycemia in this study is not restricted to documented episodes but includes any hypoglycemic event as judged by the patient. Symptoms defining hypoglycemia include trembling, palpitations, sweating, anxiety, hunger, nausea, tingling, difficulty concentrating, confusion, weakness, drowsiness, impaired vision, difficulty speaking, headache, and dizziness. Patients will self-report hypoglycemic episodes recorded by either symptoms or blood glucose testing alone or a combination of both. This approach has also been used in several large real-world studies [18,19,21-24] and better represents real-life practice where patients do not always test their blood glucose level for different reasons (eg, forgetting, neglecting, lack of knowledge, or lack of testing materials). Hypoglycemic events will be categorized as either severe or nonsevere. An SHE is defined as an event requiring assistance from another person (to administer carbohydrates, to administer glucagon, or take other corrective actions), whereas an NSHE can be managed by the patient alone, as per the Canadian Diabetes Association and ADA definitions [2,20].

Along with the severity of hypoglycemic events, the frequency of events promotes the fear of future hypoglycemia [2,20]. Fear itself affects patients' quality of life [2,20]. In addition to the physical symptoms due to hypoglycemic events, the negative emotional impact of the fear of hypoglycemia is also a concern. For modeling purposes, HTA agencies recommend that the assessment of the impact of hypoglycemia on utility should consider both the severity and frequency of episodes. They also recommend not applying a utility decrement separately for the fear of hypoglycemia. Therefore, this study was designed to capture the overall impact of hypoglycemia on utility, including the transient effects of events and the resulting fear.

Measurement of Outcomes

Utility

The EQ-5D-5L questionnaire will be used to measure the utility values [25,26]. This validated instrument is widely used and covers five dimensions (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression) that are further divided into five levels (no problems, slight problems, moderate problems, severe problems, and extreme problems). The EQ-5D (EQ-5D) health states may be converted into a single summary index through a crosswalk value set. An EQ-5D summary index of 1 represents complete health, 0 represents death, and negative values represent states worse than death. The questionnaire also includes a visual analog scale that records the respondent's self-rated health using a vertical scale ranging from 0 to 100 with end points labeled as "the best health you can imagine" and "the worst health you can imagine." EQ-5D scores represent patients' health status on the day of questionnaire completion.

Productivity Loss

The Institute for Medical Technology Assessment (iMTA) of Erasmus University Rotterdam recently developed and validated a questionnaire to estimate productivity losses, referred to as the iMTA Productivity Cost Questionnaire (iPCQ) [27]. The *i*PCQ is a generic questionnaire designed to determine the value of productivity losses for economic evaluations adopting a societal perspective. The questionnaire included three modules: lost productivity at paid work because of absenteeism, lost productivity at paid work because of presenteeism (reduced productivity), and lost productivity at unpaid work. A recall period of 4 weeks is used for presenteeism and unpaid work. Absenteeism is divided into two categories: short-term absence (no longer than 4 weeks) and long-term absence (sick leave that started before the recall period).

Study Procedures

Data Collection

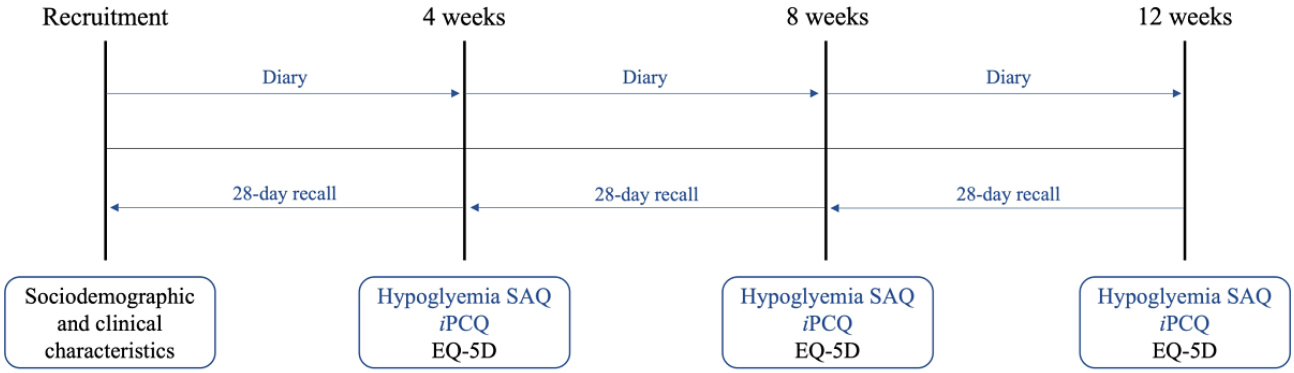
The sociodemographic and clinical characteristics of the patients will be collected at recruitment. Patients will complete a baseline questionnaire to report their age, gender, ethnicity, education, income level, employment status, living status, alcohol consumption, smoking, and physical activity. As a proxy for frailty, the self-reported number of visits to a health care provider over the last 3 months and the number of emergency visits and hospitalizations over the last 6 months will also be collected [28]. As a proxy for lifestyle habits, immunization status (pneumococcal, influenza, and COVID-19) will also be self-reported. Self-care (disease management activities) will be assessed using the Diabetes Self-Management Questionnaire, a validated questionnaire covering behaviors related to glycemic control (diet, blood glucose monitoring, medication adherence, physical activity, and contact with health care professionals) [29]. The type of device used for glucose monitoring will also be recorded (eg, traditional finger-prick monitor and flash

glucose monitoring device). It should be noted that continuous glucose monitoring systems are not reimbursed in T2DM, thereby limiting their use. To describe the history of hypoglycemia, patients will also be asked if they have experienced NSHEs and SHEs over the last 3 months and the last year, respectively. The duration of diabetes, glycated hemoglobin (HbA_{1c}), BMI, therapy, duration of insulin, and vascular complications will be extracted from the patients' medical records.

Assessments and Study Calendar

Patients will complete a self-assessment questionnaire (SAQ) at 4, 8, and 12 weeks to report the number of NSHEs and SHEs experienced in the last 28 days. Patients will also be provided with a diary to prospectively record hypoglycemic events. Both the SAQ and the diary will be used to estimate the number of hypoglycemic events. The EQ-5D and *i*PCQ will also be completed at 4, 8, and 12 weeks. The study calendar is shown in Figure 1. The questionnaires will be completed electronically via a web-based platform. To ensure that patients complete their questionnaires in a timely manner, reminders will be sent to them via email. Paper-based questionnaires can be provided to participants who do not have access to or are less familiar with using the internet (phone call reminders).

Figure 1. Study calendar. iPCQ: institute for Medical Technology Assessment Productivity Cost Questionnaire; SAQ: self-assessment questionnaire.



Ethical Considerations

This study will be conducted in accordance with the Declaration of Helsinki [30]. Approval will be taken from an independent review board before the initiation of the study, and each patient will have to provide written informed consent. After receiving independent review board approval, the protocol will be registered on ClinicalTrials.gov. All study documents, including validated versions of questionnaires, will be made available to participants in English or French, according to their preferences. The patients will not incur any costs for volunteering to participate in this study. Patients will receive a compensation of CAD $20 (US $15.75) for their participation after completion of each questionnaire.

Statistical Analysis

Presentation of Results

Data will be screened for accuracy, and questionnaires completed in paper format (if any) will entail double data entry to minimize transcription errors. Baseline characteristics of all patients participating in the study will be presented. Categorical variables will be summarized as frequencies (number and proportion), and descriptive statistics (mean and SD) will be provided for continuous variables. Hypoglycemia will be measured as hypoglycemic events and treated as a continuous variable, as recommended by HTA agencies for economic modeling of T2DM [6,7]. If hypoglycemic events are reported at a higher rate in the patient's diary than in the SAQ, diary values will be used to calculate the number of events, and SAQ values will be used in the sensitivity analysis. For the EQ-5D-5L, questionnaire scores will be calculated according to the EuroQoL scoring manual [31] using the Canadian crosswalk value set [32]. For the *i*PCQ, productivity losses will be calculated for each module (absenteeism, presenteeism, and unpaid work) using the iMTA scoring manual [33] and summed up to a total number.

Independent Impact of Hypoglycemia

Generalized estimating equation (GEE) models will be used to investigate productivity losses and utility decrements associated with incident hypoglycemic events while controlling for individual patient characteristics. For utility, two separate models will be presented: one for the EQ-5D index and one for the visual analogue scale. For productivity losses, one main analysis will be presented for the total number, and separate models for each module will be presented as a complementary analysis. The potential model-confounding covariates are presented in Table S1 in Multimedia Appendix 1. Independent variables were identified based on existing literature and previous work from the Alliance for Canadian Health Outcomes Research in Diabetes [34]. Diabetes complications (micro- and macrovascular) to be included in the models are those preconized by HTA agencies for CUAs in T2DM [4,6,7]. On the basis of the current reference study for hypoglycemic utility values [5], the number of hypoglycemic episodes will be log-transformed to facilitate model fitting. The original scale will be tested using sensitivity analyses. Moreover, transformations of continuous independent variables will also be tested in the sensitivity analyses (Table S1 in Multimedia Appendix 1). The presence of multicollinearity will be tested, where important collinearity will be defined as a variance inflation factor of >10. Owing to the short duration of the study, all covariates, except hypoglycemia, will be considered time

invariant for model simplicity. If vascular complications occur during the study period (incident cases), a history of macrovascular and microvascular complications will be treated as a time-varying covariate. The main models will not include interaction terms, as large-scale studies evaluating the impact of diabetes complications on utility have not demonstrated any interaction between vascular events and event history and among the different vascular events [35-37]. A sensitivity analysis will be performed to test for the interaction between hypoglycemic events and determinants, namely, a history of hypoglycemia and diabetes complications.

Distributions of outcomes will be examined considering the expected nature of productivity data (right-skewed and excess zeros) and utility data (left-skewed, censored, and ceiling effect). Although utility is commonly left-skewed, the distribution of utility scores may vary greatly between health conditions depending on the severity. Left skewness is more common when the studied condition is associated with mild symptoms, whereas severe health states may even lead to negative values representing conditions worse than death [38]. One major advantage of the GEE is the avoidance of distributional assumptions, where several distributions can be tested. Several large-scale longitudinal studies evaluating the impact of diabetes complications on utility, including the landmark United Kingdom Prospective Diabetes Study, used models fitted under a linear framework [35-37]. Common alternatives to the Gaussian distribution with identity link function include the Gaussian distribution with log link, negative binomial distribution with log link, gamma distribution with identity link, and gamma distribution with log link. The gamma distribution requires nonzero positive continuous data; thus, utility must be transposed into disutility (1−utility value). A beta binomial distribution can also be used with a transformation (linear transformation or rescaling) to fit the restrictive open interval (0,1), which excludes the end points 0 and 1 [39]. For the use of health care resources, nonlinear options mostly include negative binomial distribution (for resource use) and gamma and inverse-Gaussian distribution with log link function (for costs). Raw outcome data will be explored, where distributions will be depicted using histograms and normal probability plots. The selection of the model will be based on predictive performance and goodness of fit, with the lowest values for the root mean square error and the mean absolute error and the highest values for pseudo-R^2. Moreover, the model performance will also be calculated using the quasi-likelihood under the independence model criterion (QIC), where the lowest values are the best. QIC and QIC*u* will be used for correlation structure selection and variable selection, respectively. As no distributional assumptions about the response variable are made, the regression parameters cannot be estimated using the maximum likelihood method. Thus, quasi-likelihood statistics will be used (pseudo-R^2, QIC, and QIC*u*) instead of the well-known likelihood statistics (R^2, Akaike information criterion, and Bayesian information criterion) [40]. Different variance-covariance structures are available to fit the relationship between responses. In health economics, common choices include exchangeable correlation structure (observations from the same cluster are equally correlated), autoregressive (correlation decreases with time), and unstructured (different and complex correlations). According to the Good Research Practices for Retrospective Database Analysis Task Force recommendations by the International Society for Pharmacoeconomics and Outcomes Research, if results are similar with different matrices, an exchangeable matrix will be used [41]. Regression diagnostics will be performed to explore the presence of influential observations and outliers (Cook distance and residual plots). Exclusion of extreme values will be tested in sensitivity analysis. All results will be expressed as coefficients, SEs, 95% CIs, and associated *P* values (2-sided tests at the .05 significance level). Full regression models and adjusted effects will be presented. All statistical analyses will be performed using R (R Foundation for Statistical Computing) [42].

Sample Size

As per good research practices, the sample sizes for utility estimates should be based on precision rather than hypothesis testing [38,43]. Indeed, there is no consensus on the use of minimally important differences in EQ-5D measures for statistical power calculations [31,44]. Therefore, the sample size was determined to ensure reasonable variability around the utility value for SHEs (which also ensures precision for NSHEs as the incidence is significantly higher). However, no value for the expected SD was found in the literature for performing a reliable precision-based sample size calculation. A literature review was performed to identify studies that estimated utility values for hypoglycemia. However, it is often difficult to find an expected SD for utility values because they are sensitive to the type of instrument, country, precise definition of outcome, or timeframe. Furthermore, most studies did not provide uncertainty values around the estimates. As discussed by Beaudet et al [4], utility values for hypoglycemia are limited and CIs are scarce.

Therefore, the required sample size was determined based on the study by Currie et al [5], which is used as a reference for hypoglycemic utility values by CADTH and NICE [6,7]. Currie et al [5] obtained utility estimates from 68 patients who experienced at least one SHE over a 3-month period. According to Canadian real-world studies, the 3-month incidence of SHEs among insulin-treated patients with T2DM was approximately 17% [18,19,45]. With a 3-month incidence rate of 17%, 400 patients will have to be recruited to ensure that 68 patients experience at least one SHE during the study follow-up. To account for attrition, the sample size will be increased by 20% for a target number of 500 enrolled patients. It is assumed that this sample size will also ensure the precision of the productivity estimates.

Missing Data

Missing data will be defined according to outcome-specific guidelines, and a descriptive analysis will be conducted. According to the EQ-5D guidelines, situations that should be treated as missing data include not only unit nonresponse (no completion of questionnaire) but also item nonresponse (incomplete questionnaire) and ambiguous values (eg, 2 boxes are ticked for a single dimension). When completing the *i*PCQ, some questions can be skipped if they are not applicable (eg,

when the patient is not employed). If the questions necessary to score a module are incomplete, the module is defined as missing.

If ≤5% of the data are missing and there is no significant difference between completers and noncompleters, missing data will be assumed to be missing completely at random (MCAR). Under the MCAR assumption, the available case analysis performed with a GEE model yields valid estimates [40,46-48]. If >5% of the data are missing, the missing data pattern will be categorized as monotonic (ie, no return after dropout) or nonmonotonic (ie, intermittent missing data). Moreover, the mechanism by which data are missing will be investigated by examining which baseline covariates and previous measurements predict the missingness of a given outcome. On the basis of the results of the regression analysis, a specific variable will be determined to be a predictor of missingness based on statistical significance and clinical plausibility [49]. In the presence of predictors of missingness, data will be considered as not MCAR, which may bias the GEE estimates. This would lead to the use of the weighted generalized estimating equation (WGEE), a recently published extension of the GEE that incorporates an inverse probability weight matrix [50]. WGEE is a valid approach for dropout missingness (monotonic missing pattern), which is the typically observed pattern in longitudinal HEOR studies [51]. In addition, HEOR outcomes are most commonly missing as unit nonresponses (no completion of questionnaire) rather than item nonresponses (incomplete questionnaire) [49,51]. Therefore, the overall questionnaire scores (eg, EQ-5D index score) will be used for missing data analysis. If the use of the WGEE is required, a sensitivity analysis will be conducted to compare the WGEE results with the GEE estimates for the available case analysis.

Unmeasured Confounding

The presence of an unknown, unmeasured confounder will be explored using *E* values, which is a new standardized approach for sensitivity analysis [52,53]. For effect estimates, the *E* value is the minimum strength of association on the risk ratio scale that an unmeasured confounder would need to have with both the exposure and outcome, above and beyond the measured covariates, to fully explain the observed association of exposure with the outcome. For the 95% CI limit, the *E* value denotes the minimum strength of association on the risk ratio scale that an unmeasured confounder would need to have with both the exposure and the outcome, above and beyond the measured covariates, to shift the CI to include the null value. This sensitivity analysis will assess the robustness of the associations to unmeasured confounding.

Results

This study is designed to fill a gap in Canadian evidence on the impact of hypoglycemia on HEOR outcomes. More specifically, it will generate productivity and utility inputs for economic modeling of T2DM. Insulin therapy is expensive, and hypoglycemia is a significant component of economic evaluation. Robust HEOR data may help HTA agencies in future reimbursement decision-making.

Discussion

Strengths and Limitations

Insulin-induced hypoglycemia is a burden to patients with diabetes, and this study will collect HEOR estimates reflecting how SHEs and NSHEs negatively affect patients' productivity and utility. To our knowledge, this would be the first Canadian real-world study to attempt to longitudinally measure the impact of hypoglycemia on utility and productivity loss, including absenteeism, presenteeism, and unpaid work, in insulin-treated patients with T2DM. In this study, hypoglycemia will be categorized into SHEs and NSHEs without further subgrouping by severity (eg, mild or moderate) or time of occurrence (daytime or nocturnal). This classification is preconized by HTA agencies for CUAs in T2DM [6,7]. Subjective measurement of hypoglycemia may overestimate the number of episodes, which may underestimate the outcome values by event (conservative approach). In addition, the recently published guidelines by the International Hypoglycaemia Study Group recommend that the hypoglycemia threshold be lowered to <3.0 mmol/L (<54 mg/dL) in clinical trials [54]. This threshold was suggested because it is sufficiently low to indicate serious, clinically important hypoglycemia and because this level does not occur under physiological conditions in individuals who do not have diabetes. A joint position statement by the ADA and the European Association for the Study of Diabetes was subsequently published, which indicated that the glucose level should be lowered for clinical trials [54]. The use of the official definition of hypoglycemia (≤3.9 mmol/L) instead of a lower threshold might lead to overreporting of nonclinically serious events. Nevertheless, this approach is also conservative as it underestimates the outcome values by event. Defining hypoglycemia based on symptoms or blood glucose measurements is considered a reliable method that best reflects real-world practice [18,19]. The frequency of assessments (recall period) was based on previous Canadian studies [18,19]. The use of diary records (if higher than the number reported in the questionnaire) can compensate for potential recall bias and is a conservative approach as it would decrease the outcome values by event. Considering that longitudinal studies usually have a minimum of 3 measurements and that the frequency of assessments is 4 weeks, the duration of the study is therefore 3 months.

The outcomes will be measured using validated questionnaires. Although several instruments are available to estimate productivity losses, there is no gold standard [55]. In addition, HTA agencies make different recommendations regarding types of productivity losses (absenteeism, presenteeism, and unpaid work) to include in economic evaluations and methods for valuation or monetization (human capital approach and friction approach) [33]. Therefore, productivity losses will be presented as total and type-specific raw scores (number of hours per day) and will not be transposed into monetary value. This approach provides flexibility and allows future economic evaluations to be fit for specific purposes. The *i*PCQ scoring manual presents different valuation methods depending on various scenarios (eg, presence of long-term absences and frequency of measurements) and is fully adaptable to different HTA requirements [33]. It is

recommended that the choice of a work productivity instrument for economic evaluations should be based on purpose, perspective, practicality, population, and psychometrics (5 *P*s) [55]. The *i*PCQ meets the 5 *P*s criteria as it is a validated generic instrument that allows monetization, captures all types of productivity losses, and minimizes the burden to patients (has short completion time and is easy to understand). Owing to its recent development, the use of the *i*PCQ is less extended than older questionnaires, which comes with the advantage that questions are optimized based on 3 previously validated instruments [33]. Moreover, this promising instrument has a recall period of 4 weeks, which matches measurements of hypoglycemia while limiting recall bias. Although there is no gold standard, the Work Productivity and Activity Impairment Questionnaire is the most widely used instrument to assess productivity losses [56]. Similar to the *i*PCQ, the Work Productivity and Activity Impairment Questionnaire covers absenteeism, presenteeism, and unpaid work, but over a shorter period. As the recall period is only 7 days, it may not capture all productivity losses due to hypoglycemia if events occur outside of the recall period. Regarding utility, the use of generic preference-based instruments is recommended, among which the EQ-5D is the most extensively used [57]. The EQ-5D asks patients how they feel on that day without any recall period. Therefore, for complications associated with transient effects only (eg, diarrhea), utility should be measured simultaneously (ie, on the same day). When acute events are followed by persistent fear and anxiety (eg, stroke), different measurement timings can be used. If the utility is measured shortly after the event, the punctual effect of the complication will be captured. If utility is not measured concurrently, then a general effect will be captured. For economic modeling of T2DM, utility values recommended by HTA agencies were all sourced from studies designed to capture the general impact of diabetic complications, including hypoglycemia [6,7]. In a reference study by Currie et al [5], the EQ-5D captured decrements in utility due to hypoglycemic events that occurred over the last 3 months. Furthermore, in several randomized controlled trials, the impact of SHEs on utility was measured for up to 1 year after the event [36,58,59]. A limitation of this approach is the underestimation of acute physical effects related to the event. Indeed, the current design may capture the psychological effects of hypoglycemia more accurately. Nevertheless, this study will estimate the overall impact of this treatment-related adverse effect on utility, providing relevant inputs for T2DM economic modeling.

Confounding is a concern in observational research [28]. Before implementing this study, independent variables were thoroughly identified, and each known variable will be measured to limit confounding. To reflect real-world practice, independent variables that will be extracted from patients' medical files will represent the last available value, which may not reflect the current unknown value. Moreover, some independent variables will be self-reported, potentially leading to residual confounding. Nevertheless, there are no unmeasured known confounders, and the potential impact of an unmeasured unknown confounder will be tested using *E* values [52,53]. There is evidence that a cross-sectional design may overestimate the impact of T2DM complications on utility because of the underlying heterogeneity across patients [37]. Therefore, a longitudinal design with time-varying exposure will help protect against time-invariant confounding (natural heterogeneity) [40,46-48]. A GEE model will be used to account for the correlation associated with repeated measures from the same individual. Although mixed models provide a flexible framework compared with the GEE model, they require a large sample size and may be computationally demanding. Therefore, the simpler GEE method will be used to deal with this noncomplex data set (no large-scale data analysis, single level of clustering, and absence of nonstochastic time-varying covariates, eg, time from baseline). In addition, it is acknowledged that a GEE model is comparable with a random intercept model for continuous outcomes. A drawback of the GEE model is the assumption that the data are MCAR [40,46-48]. Therefore, if the data are not MCAR, the WGEE will be used to ensure the robustness of the estimates [50]. Furthermore, reminders and incentives should help to minimize the rate of missing data.

Patients will be recruited from several regions throughout Quebec and Ontario, Canada, to increase the external validity of the results. However, recruitment sites will be limited to urban areas and may not be representative of rural areas. The real-world design and broad eligibility criteria will ensure that the HEOR estimates are generalizable to a target population for future reimbursement purposes. It is assumed that recruitment through clinical sites only will not affect the representativeness of the sample as patients with diabetes have regular follow-ups with their health care providers, thereby capturing most eligible patients and not only patients in worse condition. Furthermore, enrolling patients from both medical and diabetes clinics will be representative of the target population. Although probability sampling is the gold standard for ensuring sample representativeness, it is often not feasible in Canada because many jurisdictions lack electronic patient databases, particularly family practice. Yet, systematic participant recruitment as consecutive sampling using appointment lists also helps minimize selection bias (including oversampling). Participation rates will be recorded to document the risk of selection bias. It should be noted that the COVID-19 pandemic may affect patient productivity and utility. However, given the use of the GEE model to estimate the independent impact of hypoglycemia on the HEOR estimates, the results are expected to be valid and generalizable.

Conclusions

Robust evidence on the productivity and utility of insulin-induced hypoglycemia is lacking in Canada. Currently, available data on productivity loss have not been estimated using a validated questionnaire, thereby increasing the risk of bias [3,8,9]. A systematic review published in 2021 identified 42 unique instruments for measuring productivity, and the authors recommended the iPCQ for use in economic evaluations [60]. As for the current evidence on utility decrement due to hypoglycemia, Canadian data are limited to vignette studies (ie, bespoke descriptions of impaired health states), which are not the preferred source of utility owing to their inherent drawbacks [10-12]. This study will generate high-quality HEOR estimates for future economic modeling of T2DM.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Potential independent variables for the generalized estimating equations models.
[DOCX File , 18 KB - resprot_v11i3e35461_app1.docx]

References

1. Diabetes Canada Clinical Practice Guidelines Expert Committee, Lipscombe L, Butalia S, Dasgupta K, Eurich DT, MacCallum L, et al. Pharmacologic glycemic management of type 2 diabetes in adults: 2020 update. Can J Diabetes 2020 Oct;44(7):575-591. [doi: 10.1016/j.jcjd.2020.08.001] [Medline: 32972640]
2. Diabetes Canada Clinical Practice Guidelines Expert Committee, Yale JF, Paty B, Senior PA. Hypoglycemia. Can J Diabetes 2018 Apr;42 Suppl 1:104-108. [doi: 10.1016/j.jcjd.2017.10.010] [Medline: 29650081]
3. O'Reilly DJ, Burke N, Tarride J, Hahn J, Nurkanovic L. Direct health-care costs and productivity costs associated with hypoglycemia in adults with type 1 and type 2 diabetes mellitus that participated in the Canadian hypoglycemia assessment tool program. Can J Diabetes 2018 Dec;42(6):659-663. [doi: 10.1016/j.jcjd.2018.01.010] [Medline: 29885881]
4. Beaudet A, Clegg J, Thuresson P, Lloyd A, McEwan P. Review of utility values for economic modeling in type 2 diabetes. Value Health 2014 Jun;17(4):462-470 [FREE Full text] [doi: 10.1016/j.jval.2014.03.003] [Medline: 24969008]
5. Currie CJ, Morgan CL, Poole CD, Sharplin P, Lammert M, McEwan P. Multivariate models of health-related utility and the fear of hypoglycaemia in people with diabetes. Curr Med Res Opin 2006 Jun 28;22(8):1523-1534. [doi: 10.1185/030079906x115757]
6. New drugs for type 2 diabetes: second-line therapy science report. Canadian Agency for Drugs and Technologies in Health (CADTH). 2017. URL: https://www.cadth.ca/dv/new-drugs-type-2-diabetes-second-line-therapy-recommendations-report [accessed 2021-09-14]
7. Type 2 diabetes in adults: management. Appendix F: Full Health Economics Report. National Institute for Health and Care Excellence (NICE). 2015. URL: https://www.nice.org.uk/guidance/ng28 [accessed 2021-09-14]
8. Leiter LA, Yale JF, Chiasson JL, Harris S, Kleinstiver P, Sauriol L. Assessment of the impact of fear of hypoglycemic episodes on glycemic and hypoglycemia management. Can J Diabetes 2005;29:186-192 [FREE Full text]
9. Brod M, Wolden M, Groleau D, Bushnell DM. Understanding the economic, daily functioning, and diabetes management burden of non-severe nocturnal hypoglycemic events in Canada: differences between type 1 and type 2. J Med Econ 2014 Jan 12;17(1):11-20 [FREE Full text] [doi: 10.3111/13696998.2013.857676] [Medline: 24199622]
10. Levy AR, Christensen TL, Johnson JA. Utility values for symptomatic non-severe hypoglycaemia elicited from persons with and without diabetes in Canada and the United Kingdom. Health Qual Life Outcomes 2008 Sep 29;6(1):73 [FREE Full text] [doi: 10.1186/1477-7525-6-73] [Medline: 18823555]
11. Evans M, Khunti K, Mamdani M, Galbo-Jørgensen C, Gundgaard J, Bøgelund M, et al. Health-related quality of life associated with daytime and nocturnal hypoglycaemic events: a time trade-off survey in five countries. Health Qual Life Outcomes 2013 Jun 03;11(1):90 [FREE Full text] [doi: 10.1186/1477-7525-11-90] [Medline: 23731777]
12. Harris S, Mamdani M, Galbo-Jørgensen CB, Bøgelund M, Gundgaard J, Groleau D. The effect of hypoglycemia on health-related quality of life: Canadian results from a multinational time trade-off survey. Can J Diabetes 2014 Feb;38(1):45-52 [FREE Full text] [doi: 10.1016/j.jcjd.2013.09.001] [Medline: 24485213]
13. Al Sayah F, Yeung RO, Johnson JA. Association of depressive symptoms and diabetes distress with severe hypoglycemia in adults with type 2 diabetes. Can J Diabetes 2019 Jul;43(5):316-321. [doi: 10.1016/j.jcjd.2018.11.002] [Medline: 30578165]
14. Basu S, Yudkin JS, Kehlenbrink S, Davies JI, Wild SH, Lipska KJ, et al. Estimation of global insulin use for type 2 diabetes, 2018–30: a microsimulation analysis. Lancet Diabetes Endocrinol 2019 Jan;7(1):25-33. [doi: 10.1016/s2213-8587(18)30303-6]
15. Pharmacoeconomic review report - Insulin degludec and liraglutide injection (Xultophy). Canadian Agency for Drugs and Technologies in Health (CADTH). 2019. URL: https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0599-xultophy-pharmacoeconomic-review-report.pdf [accessed 2021-09-14]
16. Pharmacoeconomic review report - Insulin glargine + lixisenatide (Soliqua). Canadian Agency for Drugs and Technologies in Health (CADTH). 2019. URL: https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/sr0564-soliqua-pharmacoeconomic-report.pdf [accessed 2021-09-14]
17. Pharmacoeconomic review report - Insulin degludec (Tresiba). Canadian Agency for Drugs and Technologies in Health (CADTH). 2017. URL: https://www.cadth.ca/sites/default/files/cdr/pharmacoeconomic/SR0521_Tresiba_PE_Report.pdf [accessed 2021-09-14]
18. Ratzki-Leewing A, Harris SB, Mequanint S, Reichert SM, Belle Brown J, Black JE, et al. Real-world crude incidence of hypoglycemia in adults with diabetes: results of the InHypo-DM Study, Canada. BMJ Open Diabetes Res Care 2018 Apr 24;6(1):e000503 [FREE Full text] [doi: 10.1136/bmjdrc-2017-000503] [Medline: 29713480]

19. Aronson R, Goldenberg R, Boras D, Skovgaard R, Bajaj H. The Canadian Hypoglycemia Assessment Tool Program: insights into rates and implications of hypoglycemia from an observational study. Can J Diabetes 2018 Feb;42(1):11-17 [FREE Full text] [doi: 10.1016/j.jcjd.2017.01.007] [Medline: 28528246]
20. Seaquist ER, Anderson J, Childs B, Cryer P, Dagogo-Jack S, Fish L, et al. Hypoglycemia and diabetes: a report of a workgroup of the American Diabetes Association and the Endocrine Society. Diabetes Care 2013 May 15;36(5):1384-1395 [FREE Full text] [doi: 10.2337/dc12-2480] [Medline: 23589542]
21. Bradley C, Eschwège E, de Pablos-Velasco P, Parhofer KG, Simon D, Vandenberghe H, et al. Predictors of quality of life and other patient-reported outcomes in the PANORAMA multinational study of people with type 2 diabetes. Diabetes Care 2018 Feb 28;41(2):267-276. [doi: 10.2337/dc16-2655] [Medline: 29183910]
22. Davis TM, Bruce DG, Curtis BH, Barraclough H, Davis WA. The relationship between intensification of blood glucose-lowering therapies, health status and quality of life in type 2 diabetes: The Fremantle Diabetes Study Phase II. Diabetes Res Clin Pract 2018 Aug;142:294-302. [doi: 10.1016/j.diabres.2018.05.047] [Medline: 29879496]
23. Brod M, Rana A, Barnett AH. Impact of self-treated hypoglycaemia in type 2 diabetes: a multinational survey in patients and physicians. Curr Med Res Opin 2012 Dec 14;28(12):1947-1958. [doi: 10.1185/03007995.2012.743457] [Medline: 23150950]
24. Östenson CG, Geelhoed-Duijvestijn P, Lahtela J, Weitgasser R, Jensen MM, Pedersen-Bjergaard U. Self-reported non-severe hypoglycaemic events in Europe. Diabet Med 2014 Jan 26;31(1):92-101 [FREE Full text] [doi: 10.1111/dme.12261] [Medline: 23796113]
25. EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9]
26. 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 25;22(7):1717-1727 [FREE Full text] [doi: 10.1007/s11136-012-0322-4] [Medline: 23184421]
27. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van RL. The iMTA productivity cost questionnaire: a standardized instrument for measuring and valuing health-related productivity losses. Value Health 2015 Sep;18(6):753-758 [FREE Full text] [doi: 10.1016/j.jval.2015.05.009] [Medline: 26409601]
28. Antihyperglycemic therapy and cardiovascular risk: design and emulation of a target trial using healthcare databases. Patient-Centered Outcomes Research Institute (2019). URL: https://www.pcori.com/sites/default/files/PCORI-Antihyperglyemic-Therapy-Cardiovascular-Risk-White-Paper-for-T2DM-RFI-052419.pdf [accessed 2021-09-14]
29. Schmitt A, Reimer A, Hermanns N, Huber J, Ehrmann D, Schall S, et al. Assessing diabetes self-management with the diabetes self-management questionnaire (DSMQ) can help analyse behavioural problems related to reduced glycaemic control. PLoS One 2016 Mar 3;11(3):e0150774 [FREE Full text] [doi: 10.1371/journal.pone.0150774] [Medline: 26938980]
30. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. J Am Med Assoc 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
31. Devlin N, Parkin D, Janssen B. Methods for Analysing and Reporting EQ-5D Data. Switzerland: Springer; 2020.
32. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, Canadian EQ-5D-5L Valuation Study Group. A time trade-off-derived value set of the EQ-5D-5L for Canada. Med Care 2016 Jan;54(1):98-105 [FREE Full text] [doi: 10.1097/MLR.0000000000000447] [Medline: 26492214]
33. Productivity costs questionnaire manual. Institute for Medical Technology Assessment (iMTA), Erasmus University Rotterdam. 2020. URL: https://www.imta.nl/questionnaires/ [accessed 2021-09-14]
34. Al Sayah F, Majumdar SR, Soprovich A, Wozniak L, Johnson ST, Qiu W, et al. The Alberta's Caring for Diabetes (ABCD) Study: rationale, design and baseline characteristics of a prospective cohort of adults with type 2 diabetes. Can J Diabetes 2015 Oct;39 Suppl 3:113-119. [doi: 10.1016/j.jcjd.2015.05.005] [Medline: 26243463]
35. Hayes A, Arima H, Woodward M, Chalmers J, Poulter N, Hamet P, et al. Changes in quality of life associated with complications of diabetes: results from the ADVANCE study. Value Health 2016 Jan;19(1):36-41 [FREE Full text] [doi: 10.1016/j.jval.2015.10.010] [Medline: 26797234]
36. Shao H, Yang S, Fonseca V, Stoecker C, Shi L. Estimating quality of life decrements due to diabetes complications in the United States: the health utility index (HUI) diabetes complication equation. Pharmacoeconomics 2019 Jul 19;37(7):921-929 [FREE Full text] [doi: 10.1007/s40273-019-00775-8] [Medline: 30778865]
37. Alva M, Gray A, Mihaylova B, Clarke P. The effect of diabetes complications on health-related quality of life: the importance of longitudinal data to address patient heterogeneity. Health Econ 2014 Apr 11;23(4):487-500 [FREE Full text] [doi: 10.1002/hec.2930] [Medline: 23847044]
38. Wolowacz SE, Briggs A, Belozeroff V, Clarke P, Doward L, Goeree R, et al. Estimating health-state utility for economic models in clinical studies: an ISPOR good research practices task force report. Value Health 2016 Sep;19(6):704-719 [FREE Full text] [doi: 10.1016/j.jval.2016.06.001] [Medline: 27712695]
39. Hunger M, Döring A, Holle R. Longitudinal beta regression models for analyzing health-related quality of life scores over time. BMC Med Res Methodol 2012 Sep 17;12(1):144. [doi: 10.1186/1471-2288-12-144]

40. Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis. Hoboken, New Jersey, United States: John Wiley & Sons, Inc; 2011.
41. Johnson ML, Crown W, Martin BC, Dormuth CR, Siebert U. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report--Part III. Value Health 2009 Nov;12(8):1062-1073 [FREE Full text] [doi: 10.1111/j.1524-4733.2009.00602.x] [Medline: 19793071]
42. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org [accessed 2021-09-14]
43. Ara R, Brazier J, Young T. Recommended methods for the collection of health state utility value evidence in clinical studies. Pharmacoeconomics 2017 Dec 19;35(Suppl 1):67-75. [doi: 10.1007/s40273-017-0549-6] [Medline: 29052159]
44. Minimal clinically important difference in EQ-5D: we can calculate it? But does that mean we should? International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and Avalon Health Economics. 2017. URL: https://www.ispor.org/docs/default-source/presentations/1066.pdf?sfvrsn=25feffd6_1 [accessed 2021-09-14]
45. Khunti K, Alsifri S, Aronson R, Berković MC, Enters-Weijnen C, Forsén T, et al. Rates and predictors of hypoglycaemia in 27 585 people from 24 countries with insulin-treated type 1 and type 2 diabetes: the global HAT study. Diabetes Obes Metab 2016 Sep 20;18(9):907-915 [FREE Full text] [doi: 10.1111/dom.12689] [Medline: 27161418]
46. Griffiths A, Paracha N, Davies A, Branscombe N, Cowie MR, Sculpher M. Analyzing health-related quality of life data to estimate parameters for cost-effectiveness models: an example using longitudinal EQ-5D data from the SHIFT randomized controlled trial. Adv Ther 2017 Mar 15;34(3):753-764 [FREE Full text] [doi: 10.1007/s12325-016-0471-x] [Medline: 28205056]
47. Gardiner JC, Luo Z, Roman LA. Fixed effects, random effects and GEE: what are the differences? Stat Med 2009 Jan 30;28(2):221-239. [doi: 10.1002/sim.3478] [Medline: 19012297]
48. Hubbard AE, Ahern J, Fleischer NL, Van der Laan M, Lippman SA, Jewell N, et al. To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. Epidemiology 2010 Jul;21(4):467-474. [doi: 10.1097/EDE.0b013e3181caeb90] [Medline: 20220526]
49. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics 2014 Dec 29;32(12):1157-1170 [FREE Full text] [doi: 10.1007/s40273-014-0193-3] [Medline: 25069632]
50. Xu C, Li Z, Xue Y, Zhang L, Wang M. An R package for model fitting, model selection and the simulation for longitudinal data with dropout missingness. Commun Stat Simul Comput 2019 Oct 16;48(9):2812-2829 [FREE Full text] [doi: 10.1080/03610918.2018.1468457] [Medline: 32346220]
51. Simons CL, Rivero-Arias O, Yu L, Simon J. Multiple imputation to deal with missing EQ-5D-3L data: should we impute individual domains or the actual index? Qual Life Res 2015 Apr 4;24(4):805-815. [doi: 10.1007/s11136-014-0837-y] [Medline: 25471286]
52. Haneuse S, VanderWeele TJ, Arterburn D. Using the e-value to assess the potential effect of unmeasured confounding in observational studies. J Am Med Assoc 2019 Mar 12;321(6):602-603. [doi: 10.1001/jama.2018.21554] [Medline: 30676631]
53. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the e-value. Ann Intern Med 2017 Aug 15;167(4):268-274. [doi: 10.7326/M16-2607] [Medline: 28693043]
54. International Hypoglycaemia Study Group. Glucose concentrations of less than 3.0 mmol/l (54 mg/dl) should be reported in clinical trials: a joint position statement of the American Diabetes Association and the European Association for the study of diabetes. Diabetes Care 2017 Jan 21;40(1):155-157. [doi: 10.2337/dc16-2215] [Medline: 27872155]
55. Tang K. Estimating productivity costs in health economic evaluations: a review of instruments and psychometric evidence. Pharmacoeconomics 2015 Jan 29;33(1):31-48. [doi: 10.1007/s40273-014-0209-z] [Medline: 25169062]
56. Work productivity and activity impairment questionnaire. Reilly Associates. URL: http://www.reillyassociates.net/WPAI_General.html [accessed 2021-09-14]
57. Brazier J, Ara R, Rowen D, Chevrou-Severac H. A review of generic preference-based measures for use in cost-effectiveness models. Pharmacoeconomics 2017 Dec 19;35(Suppl 1):21-31. [doi: 10.1007/s40273-017-0545-x] [Medline: 29052157]
58. Nauck MA, Buse JB, Mann JF, Pocock S, Bosch-Traberg H, Frimer-Larsen H, LEADER Publication Committee for the LEADER Trial Investigators. Health-related quality of life in people with type 2 diabetes participating in the LEADER trial. Diabetes Obes Metab 2019 Mar 25;21(3):525-532 [FREE Full text] [doi: 10.1111/dom.13547] [Medline: 30260088]
59. Briggs AH, Bhatt DL, Scirica BM, Raz I, Johnston KM, Szabo SM, et al. Health-related quality-of-life implications of cardiovascular events in individuals with type 2 diabetes mellitus: a subanalysis from the Saxagliptin Assessment of Vascular Outcomes Recorded in Patients with Diabetes Mellitus (SAVOR)-TIMI 53 trial. Diabetes Res Clin Pract 2017 Aug;130:24-33 [FREE Full text] [doi: 10.1016/j.diabres.2016.12.019] [Medline: 28554140]
60. Hubens K, Krol M, Coast J, Drummond MF, Brouwer WB, Uyl-de Groot CA, et al. Measurement instruments of productivity loss of paid and unpaid work: a systematic review and assessment of suitability for health economic evaluations from a societal perspective. Value Health 2021 Nov;24(11):1686-1699 [FREE Full text] [doi: 10.1016/j.jval.2021.05.002] [Medline: 34711370]

Abbreviations

ADA: American Diabetes Association
CADTH: Canadian Agency for Drugs and Technologies in Health
CUA: cost-utility analysis
GEE: generalized estimating equation
HAT: hypoglycemia assessment tool
HEOR: health economics and outcomes research
HTA: health technology assessment
iMTA: institute for Medical Technology Assessment
iPCQ: institute for Medical Technology Assessment Productivity Cost Questionnaire
MCAR: missing completely at random
NICE: National Institute for Health and Care Excellence
NSHE: nonsevere hypoglycemic event
QIC: quasi-likelihood under the independence model criterion
SAQ: self-assessment questionnaire
SHE: severe hypoglycemic event
T2DM: type 2 diabetes mellitus
TTO: time trade-off
WGEE: weighted generalized estimating equation

Edited by T Leung; submitted 05.12.21; peer-reviewed by E van den Akker-van Marle, YK Lin; comments to author 17.02.22; revised version received 24.02.22; accepted 25.02.22; published 28.03.22.

Please cite as:
Lambert-Obry V, Lafrance JP, Savoie M, Lachaine J
The Impact of Hypoglycemia on Productivity Loss and Utility in Patients With Type 2 Diabetes Treated With Insulin in Real-world Canadian Practice: Protocol for a Prospective Study
JMIR Res Protoc 2022;11(3):e35461
URL: https://www.researchprotocols.org/2022/3/e35461
doi:10.2196/35461
PMID:35343912

Protocol

Lessons From a Rapid Project Management Exercise in the Time of Pandemic: Methodology for a Global COVID-19 VIRUS Registry Database

Janice R Turek[1], BSN, MAPM; Vikas Bansal[2], MBBS, MPH; Aysun Tekin[3], MD; Shuchita Singh[3], MBBS; Neha Deo[3], BS; Mayank Sharma[3], MBBS; Marija Bogojevic[2], MD; Shahraz Qamar[3], BSc; Romil Singh[3], MBBS; Vishakha Kumar[4], MD, MBA; Rahul Kashyap[3], MBBS, MBA

[1]Department of Management Engineering and Consulting, Mayo Clinic, Rochester, MN, United States
[2]Division of Pulmonary and Critical Care Medicine, Mayo Clinic, Rochester, MN, United States
[3]Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, MN, United States
[4]The Society of Critical Care Medicine, Mount Prospect, IL, United States

Corresponding Author:
Rahul Kashyap, MBBS, MBA
Department of Anesthesiology and Perioperative Medicine
Mayo Clinic
1216 2nd St SW
Rochester, MN, 55902
United States
Phone: 1 507 255 7196
Email: kashyap.rahul@mayo.edu

Abstract

Background: The rapid emergence of the COVID-19 pandemic globally collapsed health care organizations worldwide. Incomplete knowledge of best practices, progression of disease, and its impact could result in fallible care. Data on symptoms and advancement of the SARS-CoV-2 virus leading to critical care admission have not been captured or communicated well between international organizations experiencing the same impact from the virus. This led to the expedited need for establishing international communication and data collection on the critical care patients admitted with COVID-19.

Objective: Developing a global registry to collect patient data in the critical care setting was imperative with the goal of analyzing and ameliorating outcomes.

Methods: A prospective, observational global registry database was put together to record extensive deidentified clinical information for patients hospitalized with COVID-19.

Results: Project management was crucial for prompt implementation of the registry for synchronization, improving efficiency, increasing innovation, and fostering global collaboration for valuable data collection. The Society of Critical Care Medicine Discovery VIRUS (Viral Infection and Respiratory Illness Universal Study): COVID-19 Registry would compile data for crucial longitudinal outcomes for disease, treatment, and research. The agile project management approach expedited establishing the registry in 15 days and submission of institutional review board agreement for 250 participating sites. There has been enrollment of sites every month with a total of 306 sites from 28 countries and 64,114 patients enrolled (as of June 7, 2021).

Conclusions: This protocol addresses project management lessons in a time of crises which can be a precept for rapid project management for a large-scale health care data registry. We aim to discuss the approach and methodology for establishing the registry, the challenges faced, and the factors contributing to successful outcomes.

Trial Registration: ClinicalTrials.gov NCT04323787; https://clinicaltrials.gov/ct2/show/NCT04323787

(JMIR Res Protoc 2022;11(3):e27921) doi:10.2196/27921

KEYWORDS

COVID-19; critical care; global; program management; registry

Introduction

The COVID-19 pandemic has introduced unprecedented challenges to health care systems worldwide [1]. Because of the effects of COVID-19 on the respiratory system [2], geographic areas affected by the pandemic have experienced large surges in critically ill patients who require intensive care and multiple organ system support [3]. The recognition for the necessity of conception of a COVID-19 global critical care database rapidly developed with the growing crisis hospitals experienced from the spread of the disease [4,5]. This would allow for near real-time data collection, analysis, and display. The design of the registry would be consistent with data analytic requirements [4,5]. Rapid formation of partnerships, gaining needed resources, identifying and enrolling participants, and developing the tool to collect data are crucial and time sensitive due to the rapid progression of the virus.

Implementation of a global registry during a pandemic requires extensive project management, which includes planning, initiation, execution, monitoring, and eventually closing of the registry [6,7]. Early steps require (1) research approvals for human patients/participants and data use agreements; (2) development of electronic case report forms; (3) defining common data standards and terminology; (4) development of standard operating procedure (SOP) and training for data entry; (4) coordination with other studies; (5) data quality control, automation, and validation; and finally (6) planning for diverse methods of knowledge dissemination through the registry dashboard and publications.

Project management, in a methodical but agile approach, is fundamental to release a data collection tool that would benefit health care organizations worldwide. The shared registry would enable organizations to review and analyze data in accelerated time frames on a continuous ongoing spectrum. Rapid project management and successfully implementing an initiative of this magnitude and complexity require strong leadership and the right partnerships [8,9]. An iterative, team-based approach without any lag time between different phases to get tasks done would be ideal in contrast to the waterfall methodology.

The Viral Infection and Respiratory Illness Universal Study (VIRUS): COVID-19 Registry [4,5] is a first of its kind, enrolling patients from 25 countries and more than 275 participating sites aimed at sharing real-time information on hospital care and intensive care unit admissions, which could expand the scope of SARS-CoV-2 research. The objective of this protocol was to report the project management activities, timelines, and various steps for implementation of the COVID-19 global registry. We aim to discuss the outcomes, factors which led to successful implementation, and challenges faced which could act as a guide for future establishment of large-scale registries in times of acute need.

Methods

Ethics Approval

The VIRUS Registry was approved by the Mayo Clinic Institutional Review Board (IRB; approval no. 20-002610). The Mayo Clinic IRB waived the need to obtain consent for the collection, analysis, and publication of the anonymized data for this non-interventional study and determined to be exempt from the requirement for IRB approval (45 CFR 46.104d, category 4).

Overview of the Society of Critical Care Medicine Discovery VIRUS: COVID-19 Registry

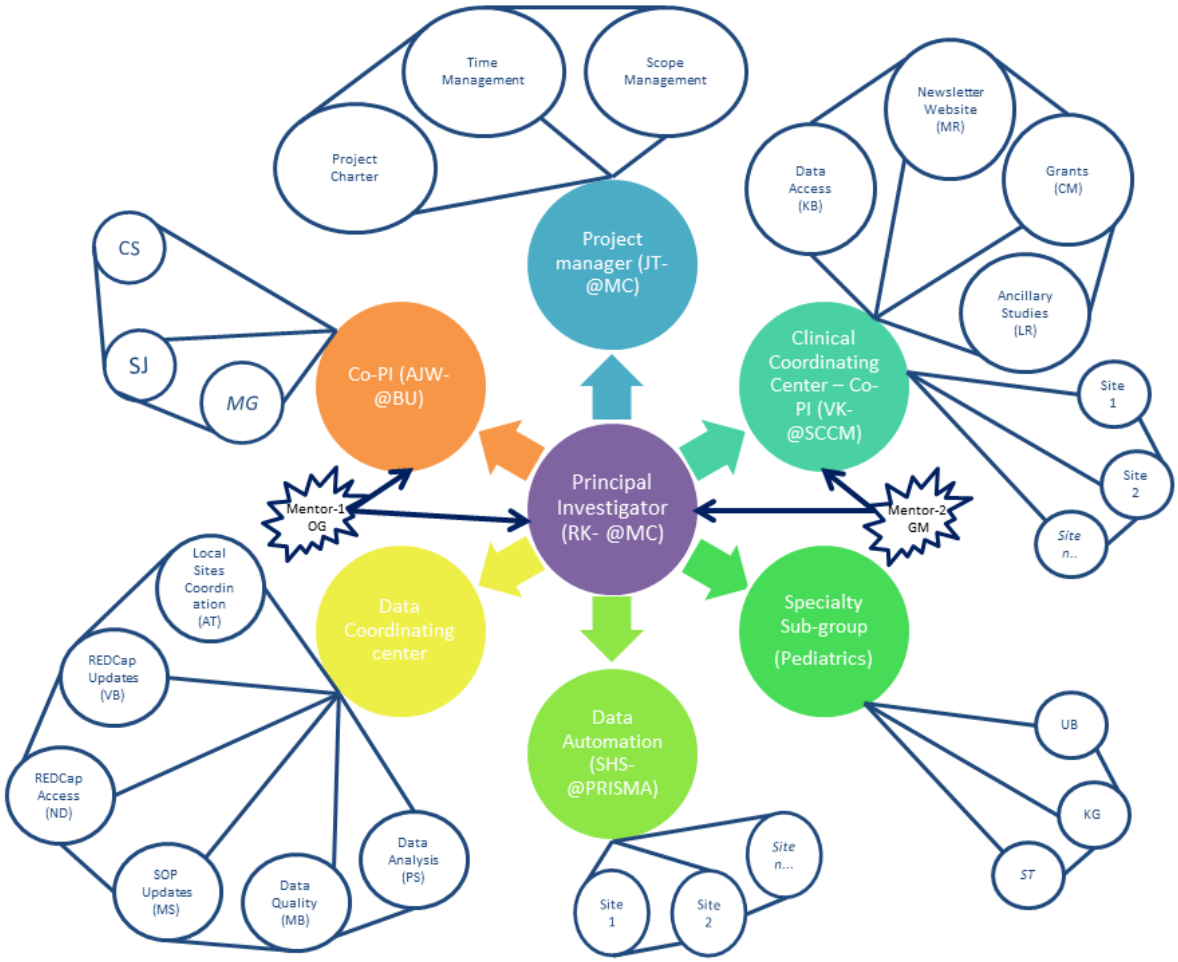
The detailed methodology and aims of the registry have been published elsewhere [4,5]. In brief, the VIRUS Registry [4,5] is a prospective, observational global registry database. In this registry, except 25 sites in the United States with data automation, all other sites collect data manually on a voluntary basis. Case report forms were constructed using Research Electronic Data Capture (REDCap) [10], to record deidentified clinical information as well as daily physiologic, laboratory, and treatment information for patients hospitalized with COVID-19.

Project Management Review

Team Organization Structure

The team organization works based on the “hub and spoke” model, where the centralization around the principle investigator (primary contact) with primary team leads promotes stability within the organization. Members are often cross-trained and learn how to work with others outside of their immediate department. This team organization style was most suited for our needs and effective communications with each node (secondary contacts). Everyone is granted equal access to information, which promotes resource sharing and the “hub and spoke” distributed style was utilized at team member level (tertiary contacts; Figure 1). The model was generated by the tool provided from [11] (also see [12]). A blank template is provided in Multimedia Appendix 1.

Figure 1. VIRUS Global Pandemic Registry Team Organizations. AJW: Allan J Walkey. AT: Aysun Tekin. BU: Boston University. CM: Colleen McNamara. CS: Christopher R Sheldrick. GM: Gregory Martin. JT: Janice Turek. KB: Karen Boman. KG: Katja Gist. LR: Lynn Retford. MB: Marija Bogojevic. MC: Mayo Clinic. MG: Michael Garcia. MR: Mary Reidy. MS: Mayank Sharma. ND: Neha Deo. OG: Ognjen Gajic. PS: Phillip Schulte. RK: Rahul Kashyap. SHS: Smith Heavner. SJ: Shelsey Johnson. ST: Sandeep Tripathi. UB: Utpal Bhalala. VB: Vikas Bansal. VK: Vishakha Kumar.



Project Management

The first and imperative step in project management is quickly establishing a manageable scope [7]. The scope (Table 1) determines the parameters of the data collection, enabling the project team to better focus efforts on setting up the registry.

The project governance is identified by the project manager and project proponents. A rapid stakeholder analysis is crucial for alliance with organizations with the same interest and familiarity with international collaboration and identification of executive leadership [12].

Table 1. Scope description of the VIRUS[a] registry.

Scope	Description
In scope description and vision	• Scope: Hospitalized patients with COVID-19 who have positivity in PCR (or other SARS-CoV-2) test or test result pending or only observational clinical data. • Vision: Real-time COVID-19 registry of current intensive care unit and hospital care patterns to allow evaluation of safety and observational effectiveness of COVID-19 practices.
Out of scope description	• Non-COVID-19–related admissions, COVID outpatients, any intervention, any biospecimen collection.

[a]VIRUS: Viral Infection and Respiratory Illness Universal Study.

Project Plan

Creating a project plan led by the project manager assures that predecessor and dependent tasks are accounted for [13]. The plan lists the key and ongoing tasks and reviewing these on a regular basis assures tasks are not being missed and resources are not over allocated. The plan accounts for using the agile project framework utilizing the sprint methodology, which is necessary to quickly create a plan for global enrollment and education.

Developing a project timeline outlining milestones ensures a quick progress review, and helps with identifying roadblocks [14]. The timeline sums up the progression of the project.

Project Charter

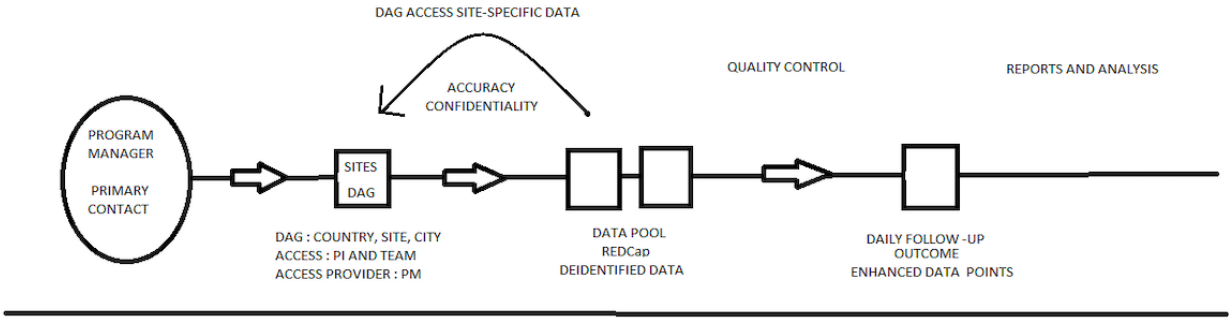
Creating a project charter is necessary but not a priority at the start of the project. The charter will serve as a valuable tool for recording methodology, history of the project and project goals, and sustainment activities. It aids in creating and supporting a budget and obtaining financial assistance. The charter for this project was developed approximately 2 months after the initiation. Project information, data, and timelines had all been recorded and saved to allow for a charter to be efficiently created and remain an agile document with additions and modifications as the project progresses.

Steps Included in Registry Implementation for Obtaining Data Inputs

Defining Data Elements and Weekly REDCap Update for Patient Data Enrollment

Focusing on data standards by using all existing standard data elements and definitions whenever possible is crucial for the development of interoperability and globalization of a registry aimed at consolidating data during a pandemic (Figure 2). This will be increasingly important as the use of electronic medical records is becoming widely available around the globe. It is also important to note that adopting standard data variables not only improves the efficiency in establishing registries but also promotes effective sharing, combining, or linking of data sets from different sources and institutions globally.

Figure 2. Workflow of the VIRUS: COVID-19 International registry. DAG: data access group, PI: program investigator, PM: program manager.



To achieve this, our registry adopted applicable data elements and definitions in accordance with (1) Critical Care Data Dictionary developed by the Data Definitions and Outcomes workgroup within the Society of Critical Care Medicine's (SCCM) Discovery, the Critical Care Research Network; (2) available published data on COVID-19 standards; and (3) case report forms, data elements, and definitions from the World Health Organization (WHO) International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) COVID-19 core case report form.

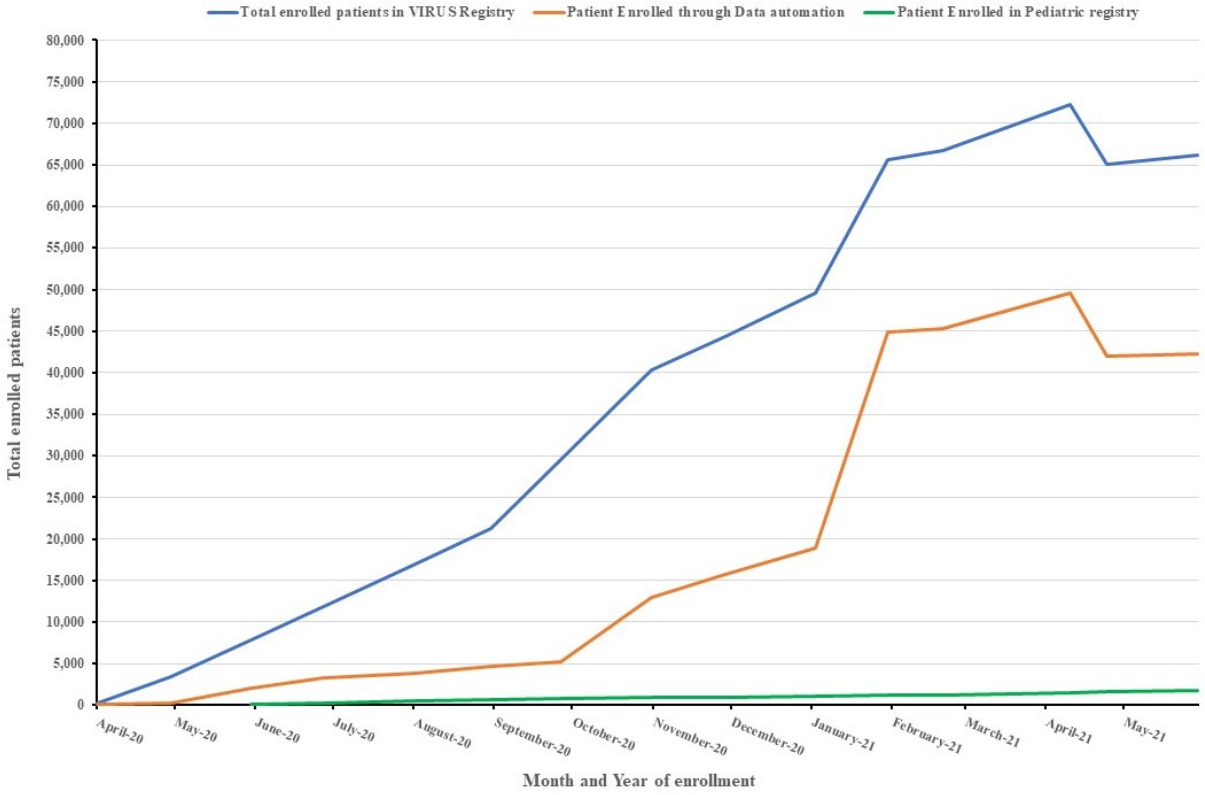
Data Collection and Research Tools Management

Collecting and sharing data in a secure manner with numerous collaborators across academic departments or even institutions remain a formidable challenge. To address this challenge, the discovery and diagnosis phase of the project were combined and REDCap [10] was recognized as an optimal data collection tool for a large registry. REDCap is an electronic data capture system that allows electronic data securely while expediting the research process and ensuring data reusability. In addition, REDCap [10] allows real-time data validation, integrity checks, and other mechanisms to ensure data quality (eg, double-data entry options) and provides central data storage and backups.

REDCap Access Management

REDCap access is necessary in order for each site to enter data for their patients. First, each hospital system needs to be inputted into REDCap as a data access group (DAG). A DAG only stores data for patients of that site. Each DAG is noted in the following format: country, site name, and city. Each site being labeled this way allows for easy organization and tracking of user accounts. To gain access to the respective DAG, either the principal investigator or members of the team request access to the data access manager. After requesting this, the data access manager requests an account creation for each respective user to REDCap administrators and then assigns them to the correct DAG. This ensures that the users in each team can only access patient data specific to their DAG, and not the rest of the patient database. This ensures accuracy and confidentiality of patient information (Figure 3).

Figure 3. The one-year (April 2020-April 2021) trend of patient enrollment in VIRUS Registry.



Data Entry

Before entering data into the registry, it is important to define inclusion and exclusion criteria for the patients. Patients admitted to a participating hospital with COVID-19 were evaluated by the local institutional research teams according to the inclusion and exclusion criteria (Table 1). The process does not differ between floor or intensive care unit admissions. When included, each patient is assigned a study number and the real-time deidentified daily data collection begins thereafter. The data that are collected for each patient can be classified as baseline characteristics, daily follow-up, outcome, and enhanced data points. Detailed information regarding collected data points is outlined in Table 2. Patients were followed during the entire hospitalization period, regardless of any transfers between different divisions in the same institution. If a patient is transferred to another hospital, reenrollment does not occur, even if the new institution is participating in the registry. The outcome data are collected at the time of discharge from the hospital. On the 28th day after hospital admission, follow-up information regarding COVID-19 readmissions and 28-day mortality is recorded.

Table 2. Specifics of collected clinical data.

Data groups	Details
Baseline characteristics	• Demographic characteristics • Symptoms and history • Comorbidities and premedications • COVID-19 manifestations • Microbiological test details
Daily follow-up	• Imaging results • Respiratory support details • Medications and other interventions • Employment of best care practices
Outcome	• Requirement of oxygenation methods • Complications during the follow-up period • Hospital and intensive care unit length of stay • In-hospital mortality • 28-day mortality
Enhanced data point	• Daily vital signs • Comprehensive physical examination findings • Detailed laboratory results • Sequential Organ Failure Assessment (SOFA) and Pediatric Logistic Organ Dysfunction (PELOD) scores
Full data	• Acute Physiological Assessment and Chronic Health Evaluation (APACHE) II score • Electrocardiogram and echocardiography findings and deidentified image upload
Data regarding co-venting patients	• Demographics of both patients • Detailed ventilator setting information • Outcome of both patients
Pediatric-specific data	• Functional Status Score (FSS) and Pediatric Risk of Mortality (PRISM) scores

Personnel Training for Data Collection and Development of Standard Operating Procedure

Performing semistructured trainings with clinical research coordinators from participating study sites globally is very important to ensure good quality assurance. The goal of these meetings would be to provide a general overview of the registry database, while concurrently identifying specific factors which could compromise the integrity of data collection. In addition to structured training, a global registry requires a good SOP.

SOP is a set of detailed instructions that define and standardize research procedures in clinical registries. SOPs describe each data variable, standard definition and step of the research process, and the actions to be taken for data collection. It provides autonomy and improves the quality of the data collected, thereby improving the science of the study. SOP can be utilized as a reference and guideline as to how research will be conducted for new study sites, and to ensure the whole process is well described, comprehensive, predictable, less prone to error, and serves as an initial training source. Several iterations may be required before the needs and concerns of all are met. SOPs are a “living” instrument. They can only work if they reflect the actual process on the ground rather than what should ideally be happening. The document should be regularly updated to reflect changes or improvements in processes over time. Primary sites should retain and store older versions of an SOP to track what were done and what data were collected at what time and place.

To ensure a standardized and consistent data collection procedure, we developed a live SOP on Google Drive specifically related to the task of the primary data collectors. This SOP provides a description of all data elements collected as well as the sources used to obtain the data.

It is also crucial to have an oversight on data entry by experienced investigators. After the training process, the data entry activities of clinical research coordinators are closely monitored by the core team to assess whether data collection was conducted according to the study protocol.

Data Automation Management

Collaboration with health care centers to optimize electronic health record (EHR) data collection is a key in making data collection more efficient. During a pandemic, the primary concern of health care workers is to care for the patients, and therefore great effort needs to be made to make the process as smooth and efficient as possible. To manage data automation, we formed a working group called “Practical EHR Export Pathways” (PEEP). This group developed methods to automate the uploading of data from EHRs to reduce workloads from sites that were struggling with the heavy clinical burdens that had resulted because of the pandemic.

Recognizing data automation as an "intervention," PEEP workgroup participating sites began an iterative process, with multiple sites producing and rapidly distributing resources (eg, MS Excel spreadsheet templates, SAS code, EHR–EPIC Workbench), incorporating feedback to describe successful tools, to optimize fit to practice settings and the ecological system, and to identify key constructs by consensus. PEEP workgroup–prioritized resources specific to the EPIC platform leverage standard search queries to obtain structured data from common EHR-generated relational databases and facilitate peer-to-peer coaching on the development and execution of end-user EPIC reporting functions. The participating sites reported implementation progress and validity through an online form and in biweekly team meetings.

Data Quality Management

The primary goal of the data quality checks is to ensure correct and consistent data entry [15]. The first step would be to analyze large volumes of data with attention to detail, accuracy, and data quality. This will help in creating acceptable data quality reports for any sort of registry. Troubleshooting to find root causes will also lead to quality improvement in the data collection tool. Preparation and consolidation of the data will be instrumental in periodic review and generation of data quality reports. The data collection tool (REDCap) also supports performing root cause analysis, investigating any data errors or anomalies, and assisting in implementing solutions to correct data problems. Meeting with study sites weekly and sharing inconsistencies and missing data for their sites will help in developing and publishing a set of quality metrics. It will facilitate capturing data trends on a weekly, monthly, and quarterly basis to ensure that data quality programs are working effectively.

Dissemination of Information Management

Partnering with a widely known medical organization or specialty society (for VIRUS registry, it was SCCM) with an imprint in the global health arena will prove to be a strategic measure enabling communication with worldwide organizations that enroll in the Registry. Here are a few highlights of communication management of the Registry:

- A communication plan needs to be put in place, identifying the standard content for each communication along with standard reporting.
- View of registry information needs to be granted to all participants so they could view information in real time.
- A public-facing webpage with registry information, contacts, and an intake form would be essential to manage the early barrage of inquiries.
- A weekly newsletter is the fastest and cheapest communication tool between the core team and participating sites.
- A public-facing data dashboard [16] would be key for simple descriptive data dissemination for participating investigators and other health care communities.
- Virtual meetings, weekly with all investigators and as needed with sites with various needs, would keep the communication lines open; details are described below.

Virtual Meetings

A pandemic usually results in a suspension of normal day-to-day activities. One of these activities is the group meeting among the researchers. Good communication between team members is key to having success in creating a registry. Therefore, it is essential to set up regular virtual meetings with clear agendas. To substitute for in-person meetings, since the beginning of the VIRUS: COVID-19 Registry project, 1-hour-long structured weekly meetings were put in place. Sharing information directly with the researchers from all study sites using Zoom as a tool for communication enables a sense of community and partnership with the global team [17]. The meeting agenda consists of updates regarding pediatric data, automation, adult data, and ancillary studies. Following these, a session takes place during which the participants get the opportunity to inquire about any unclear points. Content of the meeting sessions is summarized in Table 3.

Table 3. Project management significance: structure of the weekly meeting sessions.

Step/update	Specifics of the VIRUS[a] Registry weekly follow-up	Significance to project management
Automation update	• Data automation unit contributes with advancements in automated data collection. • They also offer assistance and collaboration opportunities to participating institutions that are interested in automation.	• Although there have been substantial developments in electronic health record–based automated data collection, there is still room for improvement. Not all sites may have the means for developing a feasible automation system all by themselves. Thus, the collaboration of sites under the supervision of a professional team helps individual sites to establish a system that would make reliable automated data collection possible.
Adult VIRUS registry data update	• The project principal investigator brings the researchers up to date about the collected data. • The study coordinating team provides comprehensive information about the updates in the standard operating procedure. • They also indicate inconsistent and missing data points and share tips about improving the quality of collected data.	• Acknowledgment of achievements, both for the whole project and for each site, increases the motivation and collaboration of the centers. • Detailed explanations of standard operating procedure amendments and their purpose help in attracting the attention of the collaborators to the updates and facilitate their compliance with the new or changed procedures. • Providing guidance about how to improve data quality and offering partnership increase the efficiency of teamwork.
Ancillary study updates	• The project management team provides insight into the ancillary study proposal submission and the approval process. • They extend guidance regarding approved ancillary study proposals.	• Being transparent in the ancillary proposal evaluation process helps with the building of trust within the study team. • Offering guidance as necessary increases efficiency and strengthens teamwork.
Pediatric update	• The VIRUS: COVID-19 Registry Pediatric Team provides detailed information regarding the current status of pediatric data. • They highlight pediatrics-specific data points.	• Providing pediatric-specific data during the general meeting strengthens the harmony between adult and pediatric sites.
Question and answer sessions	• More than half of weekly meetings is reserved for the question and answer session. • The participants get an opportunity to inquire about any unclear points, and receive explanations directly from the study principal investigator. • Fruitful discussions take place between the study coordinating team and participating researchers.	• The chance to have a personal discussion with the primary study team allows any issues to be clarified promptly. Additionally, because all investigators are a part of the conversation, it helps them to address the same situations quickly as they encounter them. • During these conversations the primary study team has the opportunity to get direct feedback from other investigators, which leads to adjustments of the project according to the needs of the researchers.

[a]VIRUS: Viral Infection and Respiratory Illness Universal Study.

Social Media Management

During these times of rapid information exchange, proper utilization of social media could be a great asset in such a large-scale global pandemic registry. The early recruitment of like-minded and motivated key team members could be done via professional social media platforms such as LinkedIn and Twitter. To enroll newer sites, the word could be spread through regular posts and tweets. A handful of social media–savvy champions to repost, reshare, and retweet with tagging and adding pertinent stakeholders would be a game changer. In our case, an early twitter handle (@covid19registry) was established even before the first site enrollment. On Twitter, more than 425 tweets have been sent in the first 8 months of registry initiation, accumulating over 725 followers. A robust social media presence would be a great strength for a global pandemic registry.

Results

The registry was established in 15 days from the inception of idea. Within the first 2 weeks, data agreements were submitted for 250 sites (Figure 4) and approval for 45 sites was achieved. A total of 69, 41, and 11 sites were enrolled in the first 3 months, respectively. Enrollment in the pediatrics registry also started from June 2020.The trend of patient enrollment is shown in Figure 2. The robust data automation management saved significant time over abstraction or manual data import processes. A total of 25 sites participate in data automation.

Figure 4. Active sites for the VIRUS: COVID-19 registry with the executive regional Leads.

The extensive data that are being collected within the scope of this study would help answer some fundamental questions about this new disease. In the end it will not only be useful for determining the impact of treatment strategies, but also could provide more insight into the pathogenesis. For this purpose, researchers that are involved in the study were given the opportunity to share their research questions with the core team. After a careful review, ancillary study proposals were approved to be conducted on the collected data and they are ongoing. The large database allows for further studies on observational outcomes. In addition, we periodically gathered queries related to VIRUS Registry REDCap and suggestions to improve the registry. We also updated REDCap fields accordingly in a weekly manner after acquiring the approval of the core committee. This process will also allow reconciling large amounts of data into concise targeted information summaries and reports for statistical analysis.

Discussion

In times of crisis, agile project management helped health care organizations form a self-organized team to set up a global registry, which implemented clear key metrics, and fostered collaboration among sites as well as sharing of resources. It also boosted quality output and improvised changes. Project management for a global registry requires rigorous training and monitoring on a large scale. We adopted the “hub and spoke” model of team organization. This style allows for more effective communication among the teams and is more adaptable to change. Establishing scope of relevant activities and assembling governance that has the right background, knowledge, and skills assured prompt response to barriers. Stakeholder analysis helped in forming alliance with shared understanding of objectives. The user interface of REDCap allowed for a quick initiation and formulation of data variables. It also proved to be an ideal tool for data import as 25 sites were using data automation. Its robust analytics helped in finding missing data and outliers. However, user maintenance is difficult in collaborating with other organizations. Robust project management was needed for successful launch of registry, which includes but not limited to, project plan, project charter, virtual meeting for project progression, and social media engagement. Our 5 key lessons learned were (1) preparation and anticipation of site needs, (2) regular communication with participating sites, (3) proactive data quality, (4) development of contingency plans through SOP, and (5) expressions of appreciation to participating sites on virtual meetings. This study can inform the project management implementation of future complex global registries.

Many studies in the past have provided guidance regarding trial management [7,18]. However, there is a lack of project management guidelines for a global registry. To our knowledge, we are the first to apply the agile project management approach to describe the conduct of a multicenter global COVID-19 registry. These results may help inform the planning of realistic clinical registry activities and set milestones for participating sites for another multicenter observational registry.

In conclusion, project management from initiation to execution with rapid availability of results in a global pandemic registry project is a herculean task. Recognizing the importance of implementing the VIRUS Registry to share best medical practices pushes the boundaries of project management and

challenges traditional project management methodologies (agile in our case) during rapid implementation. However, as illustrated in this protocol, it is achievable. A well-organized database that is led by a database manager is imperative to storing patient data securely and without errors. Perhaps the most important lesson we have learned is that the success of a registry of this scope depends crucially on the willingness of health care community to contribute to a joint initiative for a common good. This project will facilitate national consensus on data standardization and subsequent automation for rapid critical care trials and national registries.

Conflicts of Interest

RK is receiving grants from the Betty and Moore Foundation and Jensen & Jensen, LLC for supporting this registry. Neither of them has any role in writing, reviewing, or influencing this manuscript. Other coauthors have no conflict of interest.

Multimedia Appendix 1
Global pandemic registry team organization template.
[PNG File , 449 KB - resprot_v11i3e27921_app1.png]

References

1. Hospitals and health systems face unprecedented financial pressures due to COVID-19. American Hospital Association. 2020 May. URL: https://www.aha.org/guidesreports/2020-05-05-hospitals-and-health-systems-face-unprecedented-financial-pressures-due [accessed 2020-11-18]
2. Brosnahan SB, Jonkman AH, Kugler MC, Munger JS, Kaufman DA. COVID-19 and Respiratory System Disorders. ATVB 2020 Nov;40(11):2586-2597. [doi: 10.1161/atvbaha.120.314515]
3. Anesi GL. COVID-19: Management of the intubated adult. In: Manaker S, Finlay G, Bloom A, editors. UpToDate. Waltham, MA: Wolters Kluwer Health; Dec 09, 2021.
4. Walkey AJ, Kumar VK, Harhay MO, Bolesta S, Bansal V, Gajic O, et al. The Viral Infection and Respiratory Illness Universal Study (VIRUS): An International Registry of Coronavirus 2019-Related Critical Illness. Crit Care Explor 2020 May 20;2(4):e0113 [FREE Full text] [doi: 10.1097/cce.0000000000000113] [Medline: 32426754]
5. Walkey AJ, Sheldrick RC, Kashyap R, Kumar VK, Boman K, Bolesta S, et al. Guiding Principles for the Conduct of Observational Critical Care Research for Coronavirus Disease 2019 Pandemics and Beyond: The Society of Critical Care Medicine Discovery Viral Infection and Respiratory Illness Universal Study Registry. Crit Care Med 2020 Sep 16;48(11):e1038-e1044 [FREE Full text] [doi: 10.1097/CCM.0000000000004572] [Medline: 32932348]
6. Payne JM, France KE, Henley N, D'Antoine HA, Bartu AE, Elliott EJ, et al. Researchers' experience with project management in health and medical research: results from a post-project review. BMC Public Health 2011 Jun 04;11(1):424 [FREE Full text] [doi: 10.1186/1471-2458-11-424] [Medline: 21635721]
7. McCaskell DS, Molloy AJ, Childerhose L, Costigan FA, Reid JC, McCaughan M, et al. Project management lessons learned from the multicentre CYCLE pilot randomized controlled trial. Trials 2019 Aug 29;20(1):532 [FREE Full text] [doi: 10.1186/s13063-019-3634-7] [Medline: 31455384]
8. Pisano GP. You need an innovation strategy. Harvard Business Review. Boston (MA): Harvard Business Publishing; 2015 Jun. URL: https://hbr.org/2015/06/you-need-an-innovation-strategy [accessed 2020-11-19]
9. Nyström ME, Karltun J, Keller C, Andersson Gäre B. Collaborative and partnership research for improvement of health and social services: researcher's experiences from 20 projects. Health Res Policy Sys 2018 May 31;16(1):46 [FREE Full text] [doi: 10.1186/s12961-018-0322-0] [Medline: 29843735]
10. 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 2008 Oct 22;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
11. Map Chart. URL: https://mapchart.net/ [accessed 2020-11-19]
12. Beukers MW. Project management of life-science research projects: project characteristics, challenges and training needs. Drug Discov Today 2010 Dec 08;16(3-4):93-98. [doi: 10.1016/j.drudis.2010.11.015] [Medline: 21134487]
13. Kayworth T, Leidner D. The global virtual manager: a prescription for success. Eur Manag J 2000 Apr;18(2):183-194 [FREE Full text] [doi: 10.1016/S0263-2373(99)00090-0]
14. Groysberg B, Connolly BK. The COVID two-step for leaders: protect and pivot. Harvard Business School Working Knowledge. Boston (MA): Harvard Business Publishing; 2020 Sep 10. URL: https://hbswk.hbs.edu/item/the-covid-two-step-for-leaders-protect-and-pivot [accessed 2020-11-19]
15. Arts DG, De Keizer NF, Scheffer GJ. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. J Am Med Inform Assoc 2002 Oct 19;9(6):600-611 [FREE Full text] [doi: 10.1197/jamia.m1087] [Medline: 12386111]
16. The Society of Critical Care Medicine, Lyntek Medical Technologies. VIRUS COVID-19 Registry. Mount Prospect, IL: The Society of Critical Care Medicine; 2020 May 01. URL: https://sccmcovid19.org/ [accessed 2020-11-19]

17. Holtom B, Edmondson AC, Niu D. 5 tips for communicating with employees during a crisis. Crisis Communication. Boston (MA): Harvard Business Publishing; 2020 Jul 09. URL: https://hbr.org/2020/07/5-tips-for-communicating-with-employees-during-a-crisis [accessed 2020-11-19]
18. Arundel C, Gellatly J. Learning from OCTET - exploring the acceptability of clinical trials management methods. Trials 2018 Jul 15;19(1):378 [FREE Full text] [doi: 10.1186/s13063-018-2765-6] [Medline: 30005659]

Abbreviations

DAG: data access group
ISARIC: International Severe Acute Respiratory and emerging Infection Consortium
PEEP: Practical EHR (electronic health record) Export Pathways
REDCap: Research Electronic Data Capture
SCCM: Society of Critical Care Medicine
SOP: standard operating procedure
VIRUS: Viral Infection and Respiratory Illness Universal Study
WHO: World Health Organization

Edited by G Eysenbach; submitted 15.02.21; peer-reviewed by N Candula, Y Chu; comments to author 14.04.21; revised version received 29.06.21; accepted 20.07.21; published 15.03.22.

Please cite as:
Turek JR, Bansal V, Tekin A, Singh S, Deo N, Sharma M, Bogojevic M, Qamar S, Singh R, Kumar V, Kashyap R
Lessons From a Rapid Project Management Exercise in the Time of Pandemic: Methodology for a Global COVID-19 VIRUS Registry Database
JMIR Res Protoc 2022;11(3):e27921
URL: https://www.researchprotocols.org/2022/3/e27921
doi:10.2196/27921
PMID:34762062

Protocol

Development of a Digital Health Intervention for Rheumatoid Arthritis Symptom Management in a Biotechnology Industry Context: Protocol for the Application of a Human-Centered Design Framework

Lisa Nugent[1*], BFA, MFA; Robin Anthony Kouyate[1*], BA, MA, PhD; Shawna Jackson[1*], BS; Meredith Y Smith[1*], BA, MPA, PhD

Amgen, Thousand Oaks, CA, United States

[*]all authors contributed equally

Corresponding Author:
Robin Anthony Kouyate, BA, MA, PhD
Amgen
1224 Paseo Rancho Serrano
Thousand Oaks, CA, 91362
United States
Phone: 1 4102413997
Email: rakouyate@gmail.com

Abstract

Background: Involving chronically ill patients in the management of their health is widely recognized as a vital component of high-quality health care. However, to assume the role of informed participants, patients need both access to their health information and assistance in interpreting such data. Smartphone technology with SMS text messaging functionality offers a convenient and minimally demanding mechanism for providing such dual capabilities to patients. To date, a number of similar digital tools have been developed for use in various chronic and progressive disease conditions, including rheumatoid arthritis.

Objective: This paper aims to describe the development of a research protocol that applies a human-centered design (HCD) approach to develop a mobile health (mHealth) intervention to support symptom management and treatment adherence for rheumatoid arthritis.

Methods: To guide the development of the mHealth intervention for use within a commercial biotechnology context, we selected and applied an HCD framework consisting of three phases: understanding, ideation, and implementation.

Results: Leveraging the framework, we mapped the key objectives and research questions to each phase and identified the HCD techniques and methods most suitable for addressing them. In addition, we identified the need to include a fourth phase, one that referred to postimplementation assessment, which would enable evaluation of patient engagement and intervention impact on symptom self-management.

Conclusions: This paper presents a research protocol that applied an HCD framework to guide the development of an mHealth intervention within a commercial biotechnology context. This type of guidance is salient because commercial entities are becoming one of the leading producers of this type of intervention. However, the methodologies used and challenges faced from a research and development perspective are not well-represented in the published research literature to date. Our application of the HCD framework yielded important findings. Each phase of the HCD framework provided important guidance for increasing the likelihood that the final product would be understandable, acceptable, feasible, and engaging to use. Consistent with other researchers in the field of mHealth interventions, we identified the need to add a fourth phase to the HCD framework, one that focused on a postimplementation assessment to guide further improvements to support adoption in real-world settings.

International Registered Report Identifier (IRRID): RR1-10.2196/16430

(JMIR Res Protoc 2022;11(3):e16430) doi:10.2196/16430

KEYWORDS

human-centered design; patient-reported outcomes; rheumatoid arthritis; digital journal; patient diaries; data visualization; mobile phone

Introduction

Background

Patient-centeredness is increasingly recognized as a hallmark of high-quality health care delivery and drug development [1]. Empowering patients to manage their own health and disease conditions is a key aspect of patient-centeredness, which has been shown to improve treatment adherence and other patient outcomes [2]. However, to assume the role of empowered participants, patients need access to their health data and assistance in interpreting it. Smartphone technology with SMS text messaging functionality offers a convenient and minimally demanding mechanism for providing such dual capabilities to patients. To date, numerous types of digital tools have been developed for use in various chronic and progressive disease conditions, including for rheumatoid arthritis (RA).

RA is a chronic, progressive autoimmune disease. In the United States, an estimated 41 in 100,000 adults are diagnosed with this condition annually [3,4]. Typically, patients with this disease are older, White, and female [5]. Hallmarks of RA include functional disabilities that increase in severity over time and premature mortality [6,7]. Key disease markers include morning stiffness, pain, and fatigue. Treatments for RA feature a range of disease-modifying therapies, including biologics [6]. However, long-term adherence to therapy, especially biologic regimens has been shown to be poor [6,8]. Multiple factors contribute to nonadherence, including limited patient awareness and understanding of the disease, how it progresses, and how different treatments affect symptom expression over time [9,10].

Over the last two decades, patient-reported outcome measures (PROMs) have been developed to capture important aspects of RA disease and treatment [11]. Increasingly, static, one-time measurements have been replaced by a more dynamic approach in which PROMs are assessed at defined time points over a 24-hour period [11]. This type of periodic assessment is especially useful for evaluating disease progression and treatment impact and provides data that are informative for both health care professionals (HCPs) and patients alike [12]. To date, PROMs have been used predominantly in the clinical trial context. However, there is a growing recognition of their potential value in routine rheumatologic outpatient care as well [7]. Patient diaries have long been used to collect patient-reported outcomes (PROs), including RA-related symptoms [13]. Increasingly, digital diaries have supplanted paper-based versions largely because of the superior ease and timeliness of data capture [12,14]. Sharing the results of these assessments with patients visually in a graphical format has been shown to improve treatment adherence, increase patient trust in their physician, enhance patient-provider communication, boost patients' disease coping capability, and improve their understanding of the effects of disease activity and treatment [13].

Although the use of such electronic (*eHealth*) interventions has been growing, sustaining their use has remained elusive. A major contributing factor in this regard is that interventions have typically been developed with little to no involvement from the patient [13]. This lack of patient-centeredness has led to persistent usability problems and high attrition rates, resulting in mobile health (mHealth) interventions that are *high tech with a low impact* [15,16].

Currently, the peer-reviewed research literature describing patient-centered interventions to support RA symptom management and treatment adherence, digital or otherwise, is sparse [14]. Azevedo et al [17] used a patient-centered approach to develop a smartphone app for RA self-management. The study consisted of a cross-sectional patient survey to assess the usefulness of the app in supporting RA self-management, preferred features, and the degree to which patients would be willing to use and pay for it. However, no details were provided regarding whether and to what extent formative research and testing were conducted—a limitation common to many other published mHealth-based behavioral change interventions to date. To advance the science in this area, recent recommendations and guidelines call for detailed reporting of the types of methods used in each phase of the iterative design process [18-20].

Human-centered design (HCD) is an approach that can be applied to guide researchers in the reporting of these iterative design processes [20,21]. In particular, HCD engages participants in defining their unmet needs and designing solutions to address them. Within the context of health care and biopharmaceutical industries, HCD uses a patient-centered approach that emphasizes the human perspective, in addition to including criteria such as technological feasibility and economic viability when designing an intervention solution [22].

The application of HCD in the field of health and disease symptom management has been growing in recent years [21]. To date, HCD has been used to design interventions addressing a range of conditions and issues, including chronic obstructive pulmonary disease, diabetes, caregiver stress, and posttraumatic stress disorder and for a range of users (eg, patients, HCPs, and caregivers) [21]. Interventions based on an HCD approach have demonstrated greater satisfaction, usability, and effectiveness than traditional ones [21]. A defining feature of HCD is contextual inquiry—a method in which users are observed and questioned in their own environments to obtain rich information about practices, the social, technical, and physical environments, and user tools. This method can be particularly useful for understanding daily patient experiences and leveraging those insights to inform the design of tools to support patients' self-management.

Objectives

This paper describes the application of an HCD approach to guide the development of a research protocol to inform the design of a digital intervention for RA symptom management in the context of a biotechnology company. Specifically, we aim to describe the steps in the conceptualization process, the purpose of each step, and the corresponding methods and data sources used. We seek to contribute to the body of knowledge regarding the methods for designing mHealth tools to support RA symptom management and medication adherence in the real world.

Methods

We used an HCD approach to develop an intervention to assess RA-related symptoms and support treatment adherence. Signature features of HCD include the use of collaborative, multidisciplinary teams, an iterative design process involving rapid prototyping of solutions, and attention to the contexts in which the solution will be delivered [23]. The HCD approach is characterized by three main phases: (1) understanding, (2) ideation, and (3) implementation. Understanding involves exploring the dimensions, depth, and complexity of the opportunity or problem to be addressed. Ideation consists of generating, developing, and testing ideas or solutions for the identified problem. Finally, implementation involves rapid prototyping of ideas to produce solutions (eg, products and services), which are further refined via a series of subsequent iterations and feasibility assessments, including limited piloting or scaling-up efforts [23,24].

Ethical Considerations

No ethics board review was sought because the institution sponsoring this research (Amgen, Inc.) classified this research as market research. Amgen did not require ethics committee approval for healthcare market research undertaken by professional market researchers on behalf of pharmaceutical or medical device companies where such research is conducted by professional market researchers in accordance with the legal and ethical guidelines such as those issued by the British Healthcare Business Intelligence Association (BHBIA) except where otherwise required by law. Consistent with BHBIA ethics guidelines, the authors acquired informed consent of study participants.

Results

Overview

We applied the HCD framework by mapping the key research objectives to each phase. In addition, we identified the need to add a postimplementation phase as well. The purpose of this postimplementation phase was to inform future improvements to the digital intervention post launch. A summary of the steps of the framework, the purpose of each step, and the methods and data sources used for each phase are presented in Multimedia Appendix 1 and Table 1. Throughout the course of the project, a multidisciplinary team was leveraged to conduct various analyses to support the HCD process. The team consisted of experts in qualitative methods, health services research, design research, and digital health technology. In addition, as part of the iterative HCD process, patients and providers were integrated into the co-design and concept pretest phases.

Table 1. Data sources used to guide development of rheumatoid arthritis (RA) symptom management and treatment adherence intervention conceptualization.

Source	Objective	Methodology
Health care claims administrative data	To describe the size and characteristics of entire RA biologics nonadherent population	Conduct secondary data analysis of longitudinal patient data on adherence and persistence
Disease registry or medical chart data	To understand the rationale for nonadherence	Analyze aggregate RA registry data from people with RA who had initiated biologic treatments
Electronic health record	To characterize the different subtypes of patients based on rationale for drop, switch, or holiday and response rate	Analyze patient-level electronic health records for people with RA taking biologics
Patient social listening	To understand the underlying drivers of adherence to biologic treatment based on analysis of content of patient conversations with other patients	Scan social media for Patients with RA' conversations based on a list of keywords
HCP[a]-patient conversations and digital ethnography	To understand patient conversations with physicians and underlying drivers of adherence	Analysis of physician-patient with RA conversations (audio and transcripts) with redacted physician-client information
Call center	To gain insight into questions and concerns that patients have with treatment	Analysis of redacted Biologics: support call center conversations between nurses and patients
HCP ethnographic research	To gain insight into physician or office needs in helping to set RA treatment expectations and to support adherence	Conduct facility-based in-depth interviews with rheumatologists, including a simulated interaction with an actor-patient incorporating expectation-setting materials
Patient with RA ethnographic research	To gain insight into Patients with RA' experiences with using biologics and needs regarding support for adherence	Conduct interviews with patients with RA on biologics treatment, including at-home exercises, quantitative surveys and follow-up telephone in-depth interviews

[a]HCP: health care professional.

Phase I: Understanding

The understanding phase in the framework consists of a review and synthesis of a variety of different primary and secondary data sources to define the problem and to identify and frame the unmet needs to be addressed (Multimedia Appendix 1).

A range of primary and secondary data sources were identified and analyzed to characterize and explicate the rationale for medication nonadherence behaviors among patients diagnosed with moderate to severe RA who were being treated with a biologic product (Table 1).

We included four types of secondary data sources in the analysis: (1) administrative health care claims, (2) electronic health care records, (3) patient-level chart data, and (4) social media data. Primary data collection involved the use of ethnographic methods to obtain in-depth insights from both HCPs and people with RA.

Findings from this phase were instrumental in informing the problem definition for the intervention to address. Analysis and synthesis were conducted to identify distinct behavioral profiles related to adherence to RA biologic treatment, and criteria of actionability and measurability were used to select the profiles on which to intervene. Specifically, the profiles that were selected were ones in which there were defined behavioral objectives for both patients and providers, identifiable timing parameters for delivering an intervention, an understanding of the barriers to the desired behaviors (eg, *cloud of doubt* in patients' minds regarding whether the treatment was working or was continuing to work), and a defined critical *turning point* after which treatment adherence would be likely to decline (ie, 3-month mark posttreatment initiation).

The findings from this study enabled the research team to formulate a working hypothesis to guide the next steps in formative development.

Phase II: Ideation

The ideation phase consisted of three components: (1) a cocreation activity with patients and HCPs, (2) a patient journaling exercise, and (3) a literature review. The purpose of the cocreation activity was to inform the design team regarding patient and HCP needs and expectations during the patient-provider conversation to address barriers to medication adherence. Specific cocreation activities consisted of role plays with a small sample of rheumatologists, nurse practitioners, and patients in the clinic setting. In addition, patient journaling activities were used to gain a deeper understanding of patients' expectations regarding their treatment, how they managed their weekly routine, and the key differentiating factors between adherent and nonadherent patients. In addition, a literature review was conducted to identify previous e-diary interventions in patients with RA and validated PROMs to determine how (visual analog scales) and when the PROMs should be sent.

This phase resulted in a set of recommendations for an initial concept design. This concept was refined and tested in the implementation phase to address questions related to the understandability of the content and the optimal delivery of the intervention (ie, timing, frequency, and cadence).

Phase III: Implementation

The implementation phase consisted of (1) additional formative research to refine the intervention concept and (2) prototype-testing of the selected intervention to assess how to integrate it into patients' lives and to assess whether the content was understandable and acceptable.

Formative Research

The formative proof-of-concept research was conducted on a sample of patients with RA on treatment. The primary goals of the formative study were (1) to understand patient reactions to receiving multiple daily SMS text messages to assess the state of their RA symptoms and (2) to determine whether patients comprehended the content of the SMS text messages and found them to be useful. In total, 10 patients participated in the formative study for up to 4 weeks. The secondary goal was to identify a data structure that would enable the comparison of longitudinal symptom data. Determining the optimal time of day (if any) to prompt for reports of pain, fatigue, and length of morning stiffness would inform design decisions regarding how to capture and visualize the data in the next iteration.

Methods included 12 daily PROM surveys conducted via SMS text messages and weekly 30-minute patient interviews conducted via telephone. Weekly interviews were conducted to elucidate patient comprehension of the data, interpretation and utility of changes in the reported data over time, perceptions of the relationship between changes in reported data and current pain, fatigue, and morning stiffness, and feedback regarding the receipt of messages based on data. In addition, analyses were conducted to identify patterns in patients' responses to PRO text messages.

A protocol was developed to determine the optimal time for delivery of each PRO assessment [25]. The protocol probed for frequency of text messaging for pain assessment (randomly scheduled vs predetermined time points) and the type of pain scale to use (eg, 0-10 scale with 0=no fatigue to 10=totally exhausted). Similar questions were asked regarding the frequency and periodicity of the assessment and the preferred scale for measuring morning stiffness and fatigue.

The outcome of this phase included findings related to patients' (1) perceptions of the meaningfulness and usefulness of the data; (2) preferences for the timing, frequency, and cadence of the messages; and (3) the need for support in interpreting and responding to the PROs (pain, fatigue, and morning stiffness). The results of the formative phase yielded information regarding both aspects of the intervention prototype design and intervention impact. Table 2 presents examples of the types of findings at the completion of this phase.

On the basis of the learnings, the intervention was revised to reduce the number and timing of PRO assessments, to include the provision of a biweekly symptom report that visualized PRO data over time, and to send motivational and feedback messages to promote sustained patient engagement. The revised intervention then underwent concept testing in a new sample of patients with RA.

Table 2. Types of findings from the formative study.

Domain and specific constructs	Example of types of findings
Intervention delivery	
• Patients' preferences for the timing, frequency, and cadence of the messages	• Ideal number of messages • Preferences for a message schedule • Preferred amount of time to respond to PROM[a]-related messages
• Type of support to interpret and respond to the PROMs	• Understanding of different visual representations of their own data • Key elements to include in the data visualization • Comprehension and interpretation of how to respond to PROM messages (eg, whether it should be based on the last hour or the moment the message was received) • Feedback on motivational messages
Perceived impact	

• Patient perceptions of the meaningfulness and usefulness of the data	• Perceived potential impact of the intervention including
	  • Awareness of short- and long-term changes in symptom severity
	  • Usage of their data to have informed discussions with their rheumatologist regarding their symptoms
	  • Perceived usefulness to support medication adherence

[a]PROM: patient-reported outcome measure.

Prototype Testing

Following the formative phase, 2 working prototypes of the concept were pretested on patients with RA. The first prototype-testing was intended to capture patient feedback with regard to the modified intervention in a sample of biologics-naive patients with RA (prototype test 1). For example, patients provided feedback on the frequency of SMS text message requests to report symptoms and their willingness to engage in a 12-week program. The second prototype test sought to elicit feedback on the presentation of data in the symptoms journals when presented in different layouts in a subset of patients with RA who had completed the initial concept testing (prototype test 2).

Specific objectives of the prototype-testing 1 and 2 were to understand the use of a digital journal to help monitor or manage patients' disease conditions; learn if participants were able to understand and interpret the data presented (ie, the data visualizations) in 2 different presentation layouts—the original graph views and metaphoric landscape views that reflected a patient's data (eg, high levels of pain would create steep mountains vs low pain levels would create a green meadow); gauge participants' opinions on the various elements in the 2 new presentations—data visualizations and the surrounding templates; and appraise the participants' point of view on system components.

Feedback on the presentation of data in the journal was used to evaluate the best approach to visualize data in a graph format and to learn the merits of a graph versus a metaphoric view of data visualization.

Findings from this phase helped clarify the value proposition of the intervention for patients and the understandability and preferences for data visualizations. Specifically, findings were used to select the visual presentation of the data and guidance on data interpretation (eg, including question scaling) and to inform the final version of the intervention.

Phase IV: Postimplementation Assessment

A postimplementation assessment was conducted to assess patients' experiences using the intervention in a real-world setting. Specific objectives included understanding how the intervention was being used by both patients using a biologic product to treat RA symptoms and prescribers in clinical practice and how it could be changed to enhance its usefulness to patients. To address these objectives, an SMS text messaging–based survey was delivered to past intervention participants. In addition, interviews were conducted with 20 patients with RA and 10 rheumatologists who participated in the intervention. These surveys included conceptual stimuli to assist participants in *thinking aloud* and verbalizing their thoughts.

Results from a thematic analysis of the results from this phase of the research were intended to yield information regarding patient use patterns and descriptions of their experiences interacting with both the digital journaling tool and the weekly graphical output. Information was also obtained regarding both HCP and patient perceptions of the benefits of the intervention and their recommendations for its enhancement. Patients were asked to describe how the intervention affected their self-management of symptoms during the periods between HCP visits and how the intervention influenced their communication with their HCPs. These descriptions included the *emotional dimension* of their care experience (ie, their perceptions of their care and the feelings that the care experience evoked in them) while participating in the intervention.

Discussion

Principal Findings

The field of mHealth intervention research is growing rapidly. Mobile phones are both easily accessible and widely used, and they possess an ever-expanding array of features and technical capabilities [26,27]. Therefore, their use as a platform for health

interventions can only be expected to increase in the coming decades. Currently, research guidelines are calling for a systematic approach to documenting the iterative formative design processes to contribute to the evidence base for effective patient-centered mHealth interventions and to support the effective application of such interventions in real-world settings [18,19].

HCD offers a well-tested approach for addressing this gap and for helping develop an applied framework for use in the biopharmaceutical industry context. The adapted framework emerged from an inductive process derived from the development and pilot testing of a digital-based intervention to support symptom management among patients living with RA. The framework featured four separate phases: (1) understanding the dimensions and complexities of symptom management and treatment nonadherence among patients with RA, (2) intervention concept ideation, (3) iterative prototype development of the intervention and pretesting via piloting, and (4) postimplementation assessment.

Applying an HCD-based approach demands commitment to conducting in-depth formative research and an iterative approach to developing interventions. This is because an HCD approach emphasizes understanding of (1) the *context* of chronic disease management, which is a critical consideration for developing effective intervention strategies; (2) the *acceptability* of the proposed intervention to the intended recipients and those involved in its implementation; (3) *the demand* for and value of the intervention, as determined by piloting the use of selected intervention activities; and (4) the *implementation requirements*. In addition, using iterative design cycles, HCD guides ongoing intervention design in response to the circumstances and constraints encountered in real-world application [28].

Our goal in sharing this protocol is to increase the transparency of mHealth design efforts, thus aligning with recent legislative imperatives such as the 21st Century Cures Act [29]. A range of different frameworks are available to guide such efforts, including those that combine both human-centered and sociotechnical design considerations [30].

The importance of leveraging frameworks that consider the complexities of chronic disease management, the variety of stakeholders involved, and practical guidance for developing effective digital health solutions has been acknowledged [30]. Similar to our findings, van Gemert-Pijnen et al [30] emphasize the importance of systematic evaluation incorporating multiple stakeholders to ensure that solutions are user-informed, are fit for context, and add value.

Although such a comprehensive, holistic approach is recommended for use in future real-world applications of this type, the exact framework used is less important than the fact that a framework itself was applied to guide the development process. The application of a framework is critical for enabling a systematic approach to industry contributions to building, testing, and disseminating digital health interventions that help generate evidence regarding effective approaches in real-world settings.

Limitations

Arguably, a limitation of the HCD approach is that it excludes the postimplementation experience. To address this, we added a postimplementation assessment phase to our framework. However, our assessment was limited in scope, both in terms of outcomes evaluated and the duration of follow-up. Further work is needed to strengthen the postimplementation assessment phase so that the degree to which the intervention was adopted and sustained over time in the real-world context can be monitored and evaluated.

Another limitation concerns the use of real-world evidence and data analytics; there is a need for further guidance on a systematic approach to identify and evaluate the range of real-world data sources that might be appropriate for use.

Conclusions

The application of an HCD approach in a biotechnology industry setting helped inform the development of a research protocol for designing a digital health intervention for patients with moderate to severe RA. The application of this framework provided a structured road map for obtaining comprehensive, actionable insights regarding patients' daily experiences living with RA, the context of and barriers to symptom management, and treatment adherence from the perspectives of both the patient and HCPs. Collectively, such information helped directly inform the design of the intervention and increased the likelihood that it would prove acceptable, feasible, engaging, and impactful when implemented under real-world circumstances.

Acknowledgments

The authors acknowledge and thank the following individuals who contributed to the research and design of the digital health intervention referenced in this paper: Marisa Oge of *Amgen*; Tina Park and Miya Osaki of *Diagram*; and Eric Leven, Aditya Verma, Antoinette Schoenthaler, and Gillian McGovern of *Rip Road*.

Conflicts of Interest

At the time that this manuscript was developed, SJ, LN, RAK, and MYS were full-time employees of Amgen Inc and held shares in the company.

Multimedia Appendix 1

Framework overview of the development of a rheumatoid arthritis symptom management and treatment adherence digital intervention.
[DOCX File , 20 KB - resprot_v11i3e16430_app1.docx]

References

1. Topol E. The Patient Will See You Now: The Future of Medicine Is in Your Hands. New York: Basic Books; 2016.
2. Yadav UN, Hosseinzadeh H, Baral KP. Self-management and patient activation in COPD patients: an evidence summary of randomized controlled trials. Clin Epidemiol Global Health 2018 Sep 28;6(3):148-154 [FREE Full text] [doi: 10.1016/j.cegh.2017.10.004]
3. Myasoedova E, Crowson CS, Kremers HM, Therneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising?: results from Olmsted County, Minnesota, 1955-2007. Arthritis Rheum 2010 Jun 26;62(6):1576-1582 [FREE Full text] [doi: 10.1002/art.27425] [Medline: 20191579]
4. Ma VY, Chan L, Carruthers KJ. Incidence, prevalence, costs, and impact on disability of common conditions requiring rehabilitation in the United States: stroke, spinal cord injury, traumatic brain injury, multiple sclerosis, osteoarthritis, rheumatoid arthritis, limb loss, and back pain. Arch Phys Med Rehabil 2014 May;95(5):986-995 [FREE Full text] [doi: 10.1016/j.apmr.2013.10.032] [Medline: 24462839]
5. Senthelal S, Li J, Goyal A, Bansal P, Thomas M. Arthritis. Treasure Island (FL): StatPearls Publishing; 2021.
6. Thakur S, Riyaz B, Patil A, Kaur A, Kapoor B, Mishra V. Novel drug delivery systems for NSAIDs in management of rheumatoid arthritis: an overview. Biomed Pharmacother 2018 Oct;106:1011-1023 [FREE Full text] [doi: 10.1016/j.biopha.2018.07.027] [Medline: 30119166]
7. El Miedany Y. Adopting patient-centered care in standard practice: PROMs moving toward disease-specific era. Clin Exp Rheumatol 2014;32(5 Suppl 85):40-46. [Medline: 25365088]
8. Aletaha D, Smolen JS. Diagnosis and management of rheumatoid arthritis: a review. J Am Med Assoc 2018 Oct 02;320(13):1360-1372. [doi: 10.1001/jama.2018.13103] [Medline: 30285183]
9. Wong PK. Medication adherence in patients with rheumatoid arthritis: why do patients not take what we prescribe? Rheumatol Int 2016 Nov 24;36(11):1535-1542. [doi: 10.1007/s00296-016-3566-4] [Medline: 27665289]
10. Joplin S, van der Zwan R, Joshua F, Wong PK. Medication adherence in patients with rheumatoid arthritis: the effect of patient education, health literacy, and musculoskeletal ultrasound. Biomed Res Int 2015;2015:150658 [FREE Full text] [doi: 10.1155/2015/150658] [Medline: 26060812]
11. El Miedany Y. PROMs in inflammatory arthritis: moving from static to dynamic. Clin Rheumatol 2013 Jun 10;32(6):735-742. [doi: 10.1007/s10067-013-2228-0] [Medline: 23572036]
12. Stone AA, Broderick JE, Porter LS, Kaell AT. The experience of rheumatoid arthritis pain and fatigue: examining momentary reports and correlates over one week. Arthritis Care Res 1997 Jun;10(3):185-193. [doi: 10.1002/art.1790100306] [Medline: 9335630]
13. Black N. Patient reported outcome measures could help transform healthcare. Br Med J 2013 Jan 28;346:f167. [doi: 10.1136/bmj.f167] [Medline: 23358487]
14. El Miedany Y, El Gaafary M, Palmer D. Assessment of the utility of visual feedback in the treatment of early rheumatoid arthritis patients: a pilot study. Rheumatol Int 2012 Oct 11;32(10):3061-3068. [doi: 10.1007/s00296-011-2098-1] [Medline: 21909946]
15. Rudin RS, Fanta CH, Predmore Z, Kron K, Edelen MO, Landman AB, et al. Core components for a clinically integrated mHealth app for asthma symptom monitoring. Appl Clin Inform 2017 Oct;8(4):1031-1043 [FREE Full text] [doi: 10.4338/ACI-2017-06-RA-0096] [Medline: 29241243]
16. Wildenbos GA, Peute LW, Jaspers MW. Impact of patient-centered ehealth applications on patient outcomes: a review on the mediating influence of human factor issues. Yearb Med Inform 2016 Nov 10(1):113-119. [doi: 10.15265/IY-2016-031] [Medline: 27830238]
17. Azevedo R, Bernardes M, Fonseca J, Lima A. Smartphone application for rheumatoid arthritis self-management: cross-sectional study revealed the usefulness, willingness to use and patients' needs. Rheumatol Int 2015 Oct;35(10):1675-1685. [doi: 10.1007/s00296-015-3270-9] [Medline: 25903352]
18. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
19. Fjeldsoe BS, Miller YD, O'Brien JL, Marshall AL. Iterative development of MobileMums: a physical activity intervention for women with young children. Int J Behav Nutr Phys Act 2012;9:151 [FREE Full text] [doi: 10.1186/1479-5868-9-151] [Medline: 23256730]
20. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17 Suppl 1:11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
21. Altman M, Huang TT, Breland JY. Design thinking in health care. Prev Chronic Dis 2018 Sep 27;15:E117 [FREE Full text] [doi: 10.5888/pcd15.180128] [Medline: 30264690]
22. IDEO. URL: https://www.ideo.com/ [accessed 2021-10-19]

23. Berkowitz R, Vechakul J, Shrimali B, Anderson T. Human-centered design. In: Moving Life Course Theory Into Action: Making Change Happen. Washington, DC: American Public Health Association; 2018.
24. Brown T, Wyatt J. Design thinking for social innovation. Stanford Social Innovation Review. 2010. URL: https://ssir.org/articles/entry/design_thinking_for_social_innovation [accessed 2021-10-19]
25. Stone AA, Broderick JE, Schwartz JE. Validity of average, minimum, and maximum end-of-day recall assessments of pain and fatigue. Contemp Clin Trials 2010 Sep;31(5):483-490 [FREE Full text] [doi: 10.1016/j.cct.2010.06.004] [Medline: 20620239]
26. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
27. Wang P, Luo D, Lu F, Elias JS, Landman AB, Michaud KD, et al. A novel mobile app and population management system to manage rheumatoid arthritis flares: protocol for a randomized controlled trial. JMIR Res Protoc 2018 Apr 11;7(4):e84 [FREE Full text] [doi: 10.2196/resprot.8771] [Medline: 29643053]
28. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
29. Ratwani RM, Hodgkins M, Bates DW. Improving electronic health record usability and safety requires transparency. J Am Med Assoc 2018 Dec 25;320(24):2533-2534. [doi: 10.1001/jama.2018.14079] [Medline: 30489619]
30. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]

Abbreviations

HCD: human-centered design
HCP: health care professional
mHealth: mobile health
PRO: patient-reported outcome
PROM: patient-reported outcome measure
RA: rheumatoid arthritis

Edited by G Eysenbach; submitted 28.09.19; peer-reviewed by B Hesse, M Jaspers; comments to author 21.08.20; revised version received 15.01.21; accepted 19.01.21; published 22.03.22.

Please cite as:
Nugent L, Anthony Kouyate R, Jackson S, Smith MY
Development of a Digital Health Intervention for Rheumatoid Arthritis Symptom Management in a Biotechnology Industry Context: Protocol for the Application of a Human-Centered Design Framework
JMIR Res Protoc 2022;11(3):e16430
URL: https://www.researchprotocols.org/2022/3/e16430
doi:10.2196/16430
PMID:35315784

Protocol

Leveraging Large-Scale Electronic Health Records and Interpretable Machine Learning for Clinical Decision Making at the Emergency Department: Protocol for System Development and Validation

Nan Liu[1,2,3,4], PhD; Feng Xie[1], BSc; Fahad Javaid Siddiqui[1], MBBS, MSc; Andrew Fu Wah Ho[1,5], MBBS, MPH; Bibhas Chakraborty[1,6,7], PhD; Gayathri Devi Nadarajan[5], MBBS; Kenneth Boon Kiat Tan[5], MBBS; Marcus Eng Hock Ong[1,4,5], MBBS, MPH

[1]Programme in Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore
[2]Institute of Data Science, National University of Singapore, Singapore, Singapore
[3]SingHealth AI Health Program, Singapore Health Services, Singapore, Singapore
[4]Health Service Research Centre, Singapore Health Services, Singapore, Singapore
[5]Department of Emergency Medicine, Singapore General Hospital, Singapore, Singapore
[6]Department of Statistics and Data Science, National University of Singapore, Singapore, Singapore
[7]Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, United States

Corresponding Author:
Nan Liu, PhD
Programme in Health Services and Systems Research
Duke-NUS Medical School
8 College Road
Singapore, 169857
Singapore
Phone: 65 66016503
Email: liu.nan@duke-nus.edu.sg

Abstract

Background: There is a growing demand globally for emergency department (ED) services. An increase in ED visits has resulted in overcrowding and longer waiting times. The triage process plays a crucial role in assessing and stratifying patients' risks and ensuring that the critically ill promptly receive appropriate priority and emergency treatment. A substantial amount of research has been conducted on the use of machine learning tools to construct triage and risk prediction models; however, the black box nature of these models has limited their clinical application and interpretation.

Objective: In this study, we plan to develop an innovative, dynamic, and interpretable System for Emergency Risk Triage (SERT) for risk stratification in the ED by leveraging large-scale electronic health records (EHRs) and machine learning.

Methods: To achieve this objective, we will conduct a retrospective, single-center study based on a large, longitudinal data set obtained from the EHRs of the largest tertiary hospital in Singapore. Study outcomes include adverse events experienced by patients, such as the need for an intensive care unit and inpatient death. With preidentified candidate variables drawn from expert opinions and relevant literature, we will apply an interpretable machine learning–based AutoScore to develop 3 SERT scores. These 3 scores can be used at different times in the ED, that is, on arrival, during ED stay, and at admission. Furthermore, we will compare our novel SERT scores with established clinical scores and previously described black box machine learning models as baselines. Receiver operating characteristic analysis will be conducted on the testing cohorts for performance evaluation.

Results: The study is currently being conducted. The extracted data indicate approximately 1.8 million ED visits by over 810,000 unique patients. Modelling results are expected to be published in 2022.

Conclusions: The SERT scoring system proposed in this study will be unique and innovative because of its dynamic nature and modelling transparency. If successfully validated, our proposed solution will establish a standard for data processing and modelling by taking advantage of large-scale EHRs and interpretable machine learning tools.

International Registered Report Identifier (IRRID): DERR1-10.2196/34201

(JMIR Res Protoc 2022;11(3):e34201) doi:10.2196/34201

KEYWORDS

electronic health records; machine learning; clinical decision making; emergency department

Introduction

Background

Across the globe, there is increasing demand for emergency department (ED) services [1,2]. Increased ED visits have resulted in overcrowding and long waiting times [3-5]. Furthermore, adverse patient outcomes have been reported, such as mortality [6], poor patient satisfaction, and high costs [7,8]. As the first layer of emergency care in an ED, triage plays an essential role in assessing and stratifying patients' risks and ensuring that the critically ill receive appropriate emergency treatment promptly [9].

The triage process is commonly conducted by medical staff based on their own clinical experience, the patients' symptoms, and basic information obtained from patients during their presentation to the ED. To make this critical step more objective, triage systems have been introduced. Some examples of triage systems include the 5-level Emergency Severity Index [10] in the United States, the Australasian Triage Scale [11] in Australia, and the Patient Acuity Category Scale (PACS) [12] in Singapore. They are simple and easy to use but subjective and static. These scores are based on symptoms, but many critically ill patients may not have apparent symptoms when they arrive at the ED and their conditions deteriorate rapidly during their stay in the hospital. To address this limitation, more dynamic and accurate risk prediction tools are required for better patient monitoring throughout the ED journey [13].

In response to this gap of needs, researchers are interested in developing multivariable predictive models and clinical scores to identify patients in the ED at risk of adverse outcomes such as admission [14,15], death [16], cardiac arrests [17], and intensive care unit (ICU) admission [18]. Models such as these are primarily based on patient information, vital sign instability, changes in laboratory results, and administrative records. However, some parameters may appear similar between high-risk patients and other patients during an ED visit, making the prediction models less accurate.

Additional risk factors such as comorbidities, underlying chronic diseases, past hospitalization history, and other patient-related factors should be considered [19]. Furthermore, nonpatient factors are also integral components of patient care that can impact patient outcomes. Research has identified emergency boarding as a risk factor for mortality [6]. In addition, mortality rates were found to be higher for patients admitted during periods of high ED crowding regardless of their demographic characteristics, comorbidities, or primary diagnosis [20]. Changes in shift and high patient-to-nurse ratios have also been factors of concern [21].

In building predictive models, both traditional statistical methods and machine learning tools have been thoroughly investigated. Logistic regression is the most commonly used tool to construct multivariable prediction models [16,22,23]. In recent years, machine learning and artificial intelligence (AI) have gained popularity as tools for improving model performance. Fernandos et al [24] conducted an in-depth review of the current state of AI-based clinical decision support systems for triage. A recent study in the United States demonstrated the value of machine learning models for admission prediction in near real time [13].

While AI has proven successful in developing triage and prediction models, its solutions are often black box models, limiting model interpretation [25] and clinical adoption [26]. Consequently, efforts have been made to develop sparse predictive models by leveraging machine learning and conventional statistical analysis. Ustun and Rudin [27,28] proposed Supersparse Linear Integer Model–based methods for developing interpretable scoring systems. Xie et al [29] developed the interpretable machine learning–based AutoScore framework and used it to derive the score for emergency risk prediction to estimate the probability of mortality during an inpatient stay [30].

Objective

By leveraging large-scale electronic health records (EHRs) and machine learning, we intend to create an innovative, dynamic, and interpretable System for Emergency Risk Triage (SERT) for risk stratification in the ED. This protocol describes the detailed data collection procedures, data manipulation, and predictive modelling to accomplish our goals. In particular, we will employ the AutoScore framework to construct a dynamic SERT for risk assessment at multiple decision points in the ED. Our solution will also be compared with traditional clinical triage tools and black box machine learning algorithms.

Methods

Study Setting

This is a large-scale, retrospective, single-center study conducted in Singapore. As a city-state in Southeast Asia with an approximately 5.4 million population, Singapore provides affordable health care through partial subsidies and co-payments. The study site, Singapore General Hospital, is Singapore's largest and oldest tertiary referral hospital, with 1700 inpatient beds and over 30 clinical specialties. Each year, its ED sees more than 120,000 visits and admits 36,000 patients for inpatient care [16,31].

At public hospitals in Singapore, patients visiting EDs are triaged based on their symptoms according to the national PACS [32]. PACS-1 refers to patients who are seriously ill and require immediate medical care, PACS-2 refers to nonambulant patients who do not appear to be at risk of collapse, PACS-3 refers to ambulant patients, and PACS-4 refers to nonemergency cases. An initial triage is often recommended and used to identify patients who are more acutely ill and need immediate attention. As soon as resuscitation is required, the patient is taken directly to the resuscitation area. Otherwise, the patient will be directed

either to a critical care area or a waiting area, depending on the patient's condition.

Study Cohort and Design

The flowchart of the entire project is shown in Figure 1. In the extracted data set, there are 3 primary identifiers: "*ED Case No*," "*Admission Case No*," and "*Patient ID*," to represent the unique ED visit, the admission case, and the patient, respectively. Figure 2 illustrates how variables are constructed from and linked to these 3 identifiers. By consolidating the selected variables, a master data set will be created. Afterwards, the constructed master data set will be processed with outlier removal and missing value handling. The interpretable machine learning framework will then be implemented, and the models will be evaluated and compared with other baseline approaches, including traditional clinical scores, machine learning, and deep learning.

Figure 1. Flowchart of the study design. EHR: electronic health record.

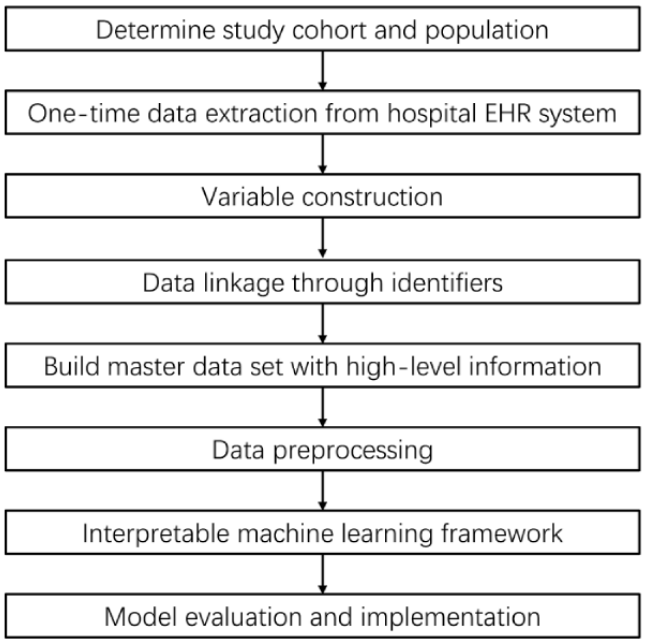
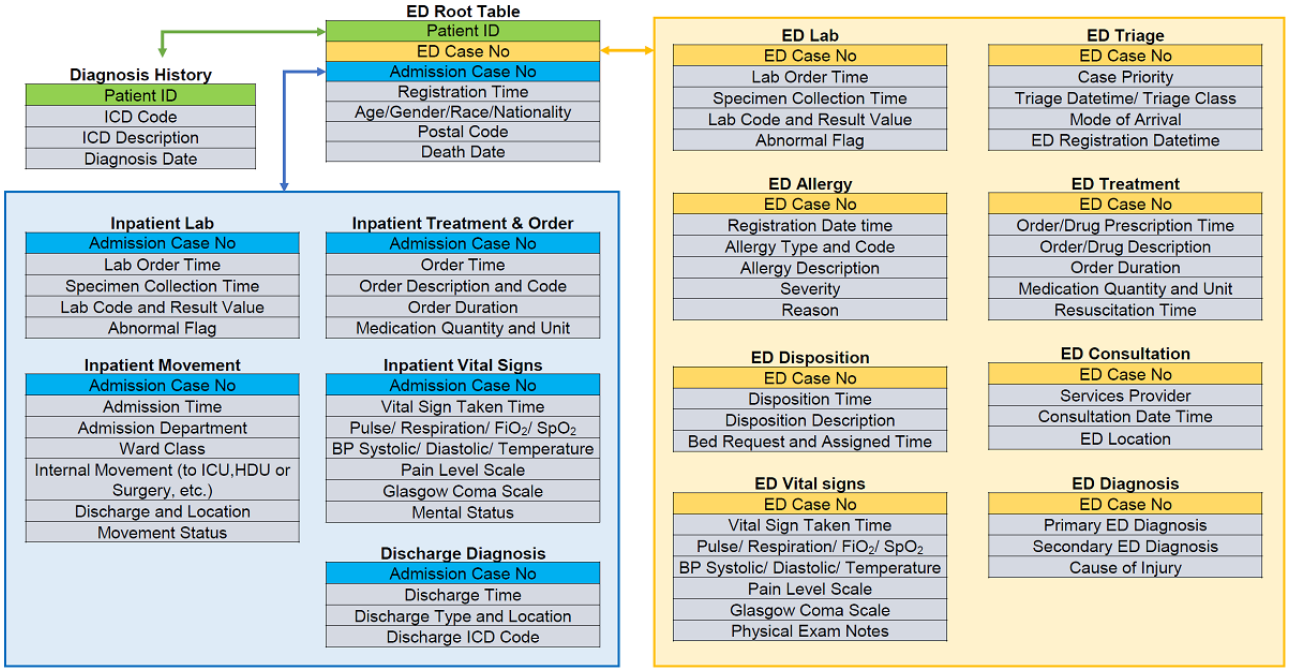


Figure 2. Illustration of the data linkage process of raw data tables through 3 primary identifiers. BP: blood pressure; ID: identification; ICD: International Classification of Diseases; ED: emergency department; ICU: intensive care unit; HDU: high dependency unit; SpO2: peripheral oxygen saturation; FiO2: fraction of inspired oxygen.



Singapore Health Services' Centralized Institutional Review Board approved this study (CIRB Ref: 2021/2122), and a waiver of consent was granted to collect and analyze EHRs.

Data Source and Extraction

Study subjects have been drawn from the hospital's EHRs using the SingHealth-IHiS Electronic Health Intelligence System,

which combines data from multiple clinical, operational, and finance data sources [33]. Before analysis, all data, including the 3 primary identifiers, have been de-identified to ensure that they are sufficiently anonymous. Records of deaths are obtained from the national death registry and are matched to specific patients in our database. Relevant variables are extracted from the beginning of the ED visits until the end of the patient's journey. Moreover, patients' medical histories are extracted and matched for each unique patient through "*Patient ID*." The extracted data were saved in multiple CSV files for subsequent processing and analysis.

Data Cleaning and Preprocessing

Data extracted from EHRs may contain many erroneous entries, as the EHRs are designed for clinical use and not explicitly modified for research purposes. This results in a lot of noise, missing values, outliers, and duplicate or incorrect records due to system problems or clerical errors. These issues will be addressed in several ways. First, wholly duplicated entries will be removed. Second, if the vital signs or laboratory test results are outside the normal range, they are considered outliers. All outliers are marked as missing values and are handled by appropriate imputation methods (eg, the mean or median value imputation based on the training data set). Third, a descriptive analysis will be conducted to determine whether the overall percentage and number are within a reasonable range.

Variable Construction

Candidate variables have been identified based on expert opinions as well as relevant literature [18,30,34-36]. Moreover, we have sought input from clinicians and informaticians familiar with the raw data to determine which features are feasible to extract and construct from the sources. The general rationale is to include all ED-relevant variables of high quality. Therefore, irrelevant, repeated, or largely missing variables will be excluded. For time-series data (such as laboratory test results and vital signs), the first, last, and average measurements are extracted and constructed for each ED episode. Past health care utilization will be derived per the patient's medical history.

Table 1 presents a list of high-level constructed variables. These variables are classified into 6 main categories depending on the time frame during which the variables could be collected: past medical history, ED triage, ED disposition, within the first 24 hours of inpatient stay, inpatient discharge, and after inpatient discharge. Variables of patient data include demographics, comorbidities, drug history, presenting vital signs, essential laboratory results, and treatments administered in the ED. There are also nonpatient variables such as ED waiting time from triage to consultation, ED boarding time (from consultation to ED disposition), patient load in the ED (number of other patients registered in the ED at that time), time of the day, and day of the week.

Outcomes

The clinical outcomes in this study include the following adverse events experienced by patients during their inpatient stay:

1. Admission: A hospital admission following an ED visit [37-39]. Each ED attendance is classified as admission or discharge according to the clinical decision made. As a result, patients who left before a decision could be made are excluded rather than considered discharged.
2. Inpatient death: A clinically certified death of a patient admitted to the hospital and who died during the hospitalization.
3. 2/7/30-day mortality: A clinically certified death of an admitted patient that occurred 2/7/30 days after the ED visit regardless of the place of death.
4. ICU transfer: Identified using the hospital's admission, transfer, and discharge database. Whenever a patient had more than one transfer from ward to ICU, only the data before the first transfer were included.
5. Cardiac arrest: Defined as the loss of a palpable pulse with attempted resuscitation in the ward.
6. Prolonged hospital length of stay: Defined as 21 days or more for the hospital stay.

Table 1. List of the high-level constructed variables in the master data set, along with their sources and categories.

Category	Subcategory	Source table	High-level variables extracted
Patient history			
	Health care utilization summary	Inpatient movement, ED[a] root table	Count of ED visits, emergency admissions, surgeries, ICU[b] or HDU[c] transfer in the patient's history (past 30/90/180/365 days)
	Comorbidities	Diagnosis history	Charlson Comorbidity Index (17 variables; chronic disease), Elixhauser Comorbidity Index (30 variables)
Information collected at triage station			
	Demographics	ED root table	Age, gender, race, nationality, postal code
	ED-prehospital	ED triage	Mode of arrival, high priority (chest pain/suspected stroke case), fever or not
	ED–triage information	ED triage	Triage waiting time, triage class (Patient Acuity Category Scale system), time of the day (midnight or not), day of the week (weekend or not)
	Triage vital signs	ED vital signs	Pulse, respiration, SpO_2[d], systolic BP[e], diastolic BP, temperature
Information collected at ED disposition			
	ED vital signs	ED vital signs	Vital measurement frequency and major ED vital readings: pulse, respiration, fraction of inspired oxygen, SpO_2, systolic BP, diastolic BP, temperature, pain level scale, Glasgow coma scale, alert (extracted from physical notes)
	ED laboratory	ED laboratory	Laboratory measurement frequency and major laboratory test results: potassium, creatinine, sodium, bicarbonate, albumin, creatine kinase-MB (mass), creatine kinase, prothrombin time, N-terminal pro–B-type natriuretic peptide, C-reactive protein
	ED consultation and treatment	ED consultation, ED treatment	Services provider, consultation waiting time, ED location, length of consultation, resuscitation, major emergency surgeries, pre-selected major ordering
	ED allergy	ED allergy	Major allergy types and reasons, severity
	ED disposition and diagnosis	ED disposition, ED diagnosis	Disposition type, major primary diagnosis, secondary diagnosis (eg, trauma)
	Outcomes	ED disposition, ED root table	Admissions, mortality within ED, direct transfer to ICU
Information collected within the first 24 hours of inpatient stay			
	Inpatient stay patient flow	Inpatient movement	ICU or HDU admission, ward class, duration of ICU or HDU stay, hospital departments, surgeries
	Inpatient vital	Inpatient vital signs	Pulse, respiration, SpO_2, systolic BP, diastolic BP, temperature, Glasgow coma scale, height, weight, BMI
	Inpatient laboratory	Inpatient laboratory	Laboratory measurement frequency and major laboratory test results: albumin, potassium, creatinine, sodium, bicarbonate, creatine kinase, creatine kinase_MB (mass), C-reactive protein, prothrombin time, procalcitonin, blood PH, glycated hemoglobin A1c, triglycerides, cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol
	Inpatient treatment	Inpatient treatment and order	Major medication prescription and order
Information collected at discharge			
	Health care utilization summary	Inpatient movement	Count of ED visits, ED admissions, surgeries, ICU or HDU admissions last year
	Discharge information	Discharge diagnosis	Primary discharge diagnosis, discharge location, length of stay
	Outcomes	Inpatient movement, discharge diagnosis	ICU transfer, inpatient mortality, cardiac arrest, prolonged hospital length of stay
Information collected after discharge			
	Outcomes	ED root table	2/7/30-day mortality, emergency readmission, ED revisit

[a]ED: emergency department.

[b]ICU: intensive care unit.

[c]HDU: high dependency unit.

[d]SpO_2: peripheral oxygen saturation.

[e]BP: blood pressure.

Predictive Modelling for Clinical Decision Making

In this study, we will develop and validate a novel interpretable triage system for risk stratification of patients in the ED. Our proposed solution will be compared with baseline risk prediction tools such as traditional clinical scores and black box machine learning models. The extracted data set will be split into training, validation, and testing sets to build and validate the predictive models. The ED visit episodes from January 1, 2008, to December 31, 2018, will be randomly divided into 2 non-overlapping cohorts: a training cohort (80%) and a validation cohort (20%). The ED visits dated in 2019 are assigned to one testing cohort, while those dated in 2020 are assigned to a second testing cohort covering the period of the COVID-19 pandemic [40,41]. Using this sequential testing design, we will be able to test whether the population shift and the COVID-19 pandemic would impact model performance [42]. Further details are presented below.

Proposed Method: Interpretable SERT

SERT consists of 3 scoring algorithms, each tailored to its application at different time points in the ED. On arrival at the triage station, SERT-1 is used to estimate patients' likelihood of admission (inpatient and ICU) and 2-day mortality. SERT-1 is intended to assess the patient's immediate urgency based on basic patient information, simple vital measurements, and medical histories readily available during triage. While in the ED, SERT-2 predicts patients' admission (inpatient and ICU) and 2/7-day mortality using a variety of variables, including laboratory test results, vital signs, ED treatment, diagnosis, and some administrative information. As an extension of the SERT-1 algorithm, SERT-2 incorporates additional variables obtained during ED stay to better predict outcomes. On admission, SERT-3 predicts the likelihood of 7/30-day mortality, ICU transfer, and prolonged length of stay using variables collected in the ED and during the first 24 hours of inpatient stay. In actual clinical implementation, in the case where a patient has incomplete information, SERT will use imputation methods to fill in the missing values before calculating the risk score. In summary, SERT allows for a comprehensive risk assessment and prediction in the ED in a dynamic manner.

The clinical risk-scoring models have been traditionally developed in 2 ways: through expert opinions or consensus and conventional cohort studies. However, both approaches are labor-intensive and are not easy to update over time. Recently, we developed an interpretable machine learning–based automatic clinical score generator, AutoScore, as a practical and universal solution for risk scoring [29]. Using the AutoScore framework, users could seamlessly generate parsimonious risk models (ie, point-based sparse risk scores), thereby supporting automated machine learning solutions in health care [43]. AutoScore comprises 6 modules. In module 1, random forest is used to rank variables in terms of their contribution to modelling. Module 2 categorizes continuous variables to address nonlinearity and facilitate the generation of point-based scores. Module 3 computes scores based on a subset of variables and logistic regression, while module 4 determines the optimal number of variables based on a parsimony plot. Module 5 enables fine-tuning of the cut-off values for categorizing continuous variables for preferable interpretation, and module 6 provides a final performance evaluation. AutoScore is used to develop the 3 SERT scoring algorithms with the candidate variables and the outcomes.

Baseline Methods: Traditional Clinical Scores

Several traditional clinical scores will be calculated for performance comparison with the SERT scores. They are the PACS triage system [32], Modified Early Warning Score [44], National Early Warning Score [45], Rapid Acute Physiology Score [46], Rapid Emergency Medicine Score [47], and Cardiac Arrest Risk Triage [48].

Baseline Methods: Black Box Machine Learning Models

Additionally, several machine learning techniques will be compared as baselines for predictive modelling. Of the many machine learning algorithms, we will apply the following popular ones as examples.

1. Random forest [49]: As the most commonly used tree-based prediction tool, its R package "RandomForest" will be used for model fitting. The parameters will be selected based on recommendations made in previous literature [50,51], where *ntree*= 100 and *mtry* is the principal square root of *m* (*ntree* number of trees grown; *mtry*: number of variables randomly sampled as candidates at each split).
2. Least absolute shrinkage and selection operator [52]: As a penalized regression technique, it is another popular method used in clinical modelling. It is a regression-based method that employs a regularization process for variable selection to increase the statistical model's predictive accuracy and interpretability. In our study, its regularization rate will be optimized through 10-fold cross-validation.
3. Deep learning [53]: As a branch of the machine learning field that uses deep neural networks, deep learning was initially widely adopted for computer vision and image understanding before being used for medical image analysis. More recently, researchers have begun to explore deep learning for EHR analysis [54,55]. We are particularly interested in applying deep learning algorithms for adverse event prediction, drawing on the rich sources of EHR data, as described earlier. Using the PyTorch library, we will construct a long short-term memory network [56]. In addition, a multilayer perceptron [57] will be used in conjunction with long short-term memory to learn nontemporal data.

Model Comparison and Performance Metrics

To evaluate the performance of all predictive models, receiver operating characteristic (ROC) analysis will be conducted on the 2 testing cohorts. An overall measure of predictive performance is represented by the area under the ROC curve. Moreover, we will calculate the measures of diagnostic accuracy, such as sensitivity, specificity, positive predictive value, and negative predictive value. These specific measures are determined by setting thresholds on each ROC curve. To achieve optimal balance between sensitivity and specificity, we will select the cut-off points closest to the plot area's upper-left corner. The 95% CIs for each model or score will also be reported and compared.

Statistical Analysis

We will perform data analysis using R version 4.0 (R Core Team). When summarizing descriptive results, frequency and percentages are reported for categorical variables, while means and SDs are reported for continuous variables. For categorical variables, the chi-square test or Fisher exact test will be used. For numeric variables, the *t* test will be applied. Further, univariable and multivariable logistic regressions will be used to identify common risk factors associated with the outcomes.

Results

The raw data have been extracted, and we are currently linking and cleaning the data. In the data extraction process, we included all patients who visited the ED at Singapore General Hospital between January 1, 2008, and December 31, 2020. Patients under the age of 21 years were excluded. If the patients were admitted through the corresponding ED visit, they would be followed throughout their inpatient stay. The data set contains more than 1.8 million ED visit episodes of over 810,000 unique patients. Approximately 650,000 of these ED visits resulted in subsequent hospitalizations. Our findings and modelling results are expected to be published by 2022.

Discussion

This paper presents a protocol designed to leverage large-scale EHRs and advanced machine learning techniques for risk stratification and triage in the ED. Among numerous ED triage and risk prediction scores and tools, our proposed SERT solution is unique and innovative because of its dynamic nature and modelling transparency. This project will build on the success of our previous research on risk modelling with EHRs for patients in the ED [14,16,30].

Significance

The identification of patients' risk at an early stage allows for better resource allocation. There is particular significance in this point because the instability of vital signs may occur later in the ward, leaving a limited time window for life-saving action or decision making, which can be especially difficult in a busy hospital. Patient groups at high risk should be identified earlier in the ED and, if possible, flagged for more stringent monitoring. Similarly, low-risk patients may require less intensive monitoring and treatment, thereby saving hospital resources. The SERT system that we propose has the potential to provide a feasible solution. This system allows medical personnel to assess patient risk at multiple decision points based on various clinical and nonclinical factors. In a dynamic way, SERT measures risk sequentially and in a manner that is perfectly suited to actual clinical needs.

Strengths

First, this study uses a large set of EHR data over a 13-year period, which contains comprehensive patient information. As Singapore's largest hospital, Singapore General Hospital provides medical care to a wide range of patients throughout the country; thus, its EHRs ensure good coverage for a large population. Additionally, the longitudinal data allow us to validate the SERT system using data before and after COVID-19. Thus, we will have the opportunity to evaluate the impact of the global pandemic on triage performance in the ED. The insights gained from system evaluation could be used to examine possible model adaptations in shifted clinical settings.

Second, the SERT triaging system we intend to develop will be transparent and easily understandable. All 3 SERT scores are parsimonious and point-based, as only the most significant variables are considered in their formulation. Their formats follow the same convention as widely used clinical scores such as the National Early Warning Score and Modified Early Warning Score, allowing for easy comprehension and quick adoption. In contrast, black box machine learning models are challenging to comprehend, making them inaccessible to clinicians [25]. Although there are techniques for post hoc model explanation, most machine learning models are not inherently interpretable [25].

Third, this project aims to develop a dynamic system capable of identifying risk strata at different decision points in the ED. During the initial triage process and the patient's stay in the ED, SERT predicts the likelihood of inpatient and ICU admissions. Whenever variables are altered, the scores can be updated, making the risk assessment dynamic and practical. In addition, SERT can make mortality predictions to assess the likelihood of the worst outcomes for patients who will be hospitalized.

Lastly, the simple form of the scores in SERT permits a variety of implementation schemes. As an example, the actual implementation can be as simple as a mobile app. Users may input relevant information into the app, which will return a risk score at the time of inquiry. The SERT scoring platform can also be easily integrated into existing information technology systems, which requires only simple calculations and therefore little computing power. The application can be designed and implemented in real time, similar to that seen in a recent study in the United States [13].

Limitations and Future Plan

Although the study site is the largest hospital in the country, the SERT system may not apply to international institutions where EDs operate differently. We intend to conduct cross-institutional validation of our system with both local and international partners. In the case that our SERT system is not feasible, the methods we use can easily be adapted to any

context because AutoScore is a generic, universal scoring tool that permits the creation of interpretable clinical scores. In addition, we anticipate a sparse data set with numerous missing values, particularly for comorbidities, medications, and time series records of vitals and laboratory test results. To address the issue of data sparsity, we will examine various data imputation strategies and feature representation techniques.

Our future efforts will include identifying opportunities to conduct a rigorously designed randomized trial to evaluate the system. In the long-term, we hope to expand the evaluation to a multicenter trial involving several countries.

Conclusions

Clinical decision making has widely benefited from the use of machine learning techniques. However, the black box models created by these methods prevent their use in actual clinical practice. Our study aims to address this issue by proposing an innovative SERT scoring system. An interpretable machine learning–based AutoScore framework will be used to create a series of 3 SERT scores that can be used in the medical setting at various decision points throughout the patient's journey. The SERT system is notable for its dynamic nature and transparency. If validated successfully, it will establish a standard for data processing and modelling by utilizing large-scale EHRs and interpretable machine learning. The proposed system may be well suited to bridge the gap between advanced computation and clinical applications.

Acknowledgments

This study is supported by Duke-NUS Medical School, Singapore. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

NL and MEHO are Scientific Advisors of TIIM Healthcare PTe Ltd, a startup with solutions in medical triaging. All other authors have no conflicts of interest to declare.

References

1. Lowthian JA, Curtis AJ, Cameron PA, Stoelwinder JU, Cooke MW, McNeil JJ. Systematic review of trends in emergency department attendances: an Australian perspective. Emerg Med J 2011 May;28(5):373-377. [doi: 10.1136/emj.2010.099226] [Medline: 20961936]
2. Derlet RW. Overcrowding in emergency departments: increased demand and decreased capacity. Ann Emerg Med 2002 Apr;39(4):430-432. [doi: 10.1067/mem.2002.122707] [Medline: 11919530]
3. Derlet RW, Richards JR. Overcrowding in the nation's emergency departments: complex causes and disturbing effects. Ann Emerg Med 2000 Jan;35(1):63-68. [doi: 10.1016/s0196-0644(00)70105-3] [Medline: 10613941]
4. Hoot NR, Aronsky D. Systematic review of emergency department crowding: causes, effects, and solutions. Ann Emerg Med 2008 Aug;52(2):126-136 [FREE Full text] [doi: 10.1016/j.annemergmed.2008.03.014] [Medline: 18433933]
5. Di Somma S, Paladino L, Vaughan L, Lalle I, Magrini L, Magnanti M. Overcrowding in emergency department: an international issue. Intern Emerg Med 2015 Mar;10(2):171-175. [doi: 10.1007/s11739-014-1154-8] [Medline: 25446540]
6. Singer AJ, Thode HC, Viccellio P, Pines JM. The association between length of emergency department boarding and mortality. Acad Emerg Med 2011 Dec;18(12):1324-1329 [FREE Full text] [doi: 10.1111/j.1553-2712.2011.01236.x] [Medline: 22168198]
7. Sabbatini AK, Kocher KE, Basu A, Hsia RY. In-hospital outcomes and costs among patients hospitalized during a return visit to the emergency department. JAMA 2016 Feb 16;315(7):663-671 [FREE Full text] [doi: 10.1001/jama.2016.0649] [Medline: 26881369]
8. Carter AJ, Chochinov AH. A systematic review of the impact of nurse practitioners on cost, quality of care, satisfaction and wait times in the emergency department. CJEM 2007 Jul 21;9(4):286-295. [doi: 10.1017/s1481803500015189] [Medline: 17626694]
9. FitzGerald G, Jelinek GA, Scott D, Gerdtz MF. Emergency department triage revisited. Emerg Med J 2010 Feb;27(2):86-92. [doi: 10.1136/emj.2009.077081] [Medline: 20156855]
10. Wuerz RC, Milne LW, Eitel DR, Travers D, Gilboy N. Reliability and validity of a new five-level triage instrument. Acad Emerg Med 2000 Mar;7(3):236-242 [FREE Full text] [doi: 10.1111/j.1553-2712.2000.tb01066.x] [Medline: 10730830]
11. Considine J, LeVasseur S, Villanueva E. The Australasian Triage Scale: examining emergency department nurses' performance using computer and paper scenarios. Ann Emerg Med 2004 Nov;44(5):516-523. [doi: 10.1016/j.annemergmed.2004.04.007] [Medline: 15520712]
12. Fong RY, Glen WSS, Mohamed Jamil AK, Tam WWS, Kowitlawakul Y. Comparison of the Emergency Severity Index versus the Patient Acuity Category Scale in an emergency setting. Int Emerg Nurs 2018 Nov;41:13-18. [doi: 10.1016/j.ienj.2018.05.001] [Medline: 29887281]

13. Fenn A, Davis C, Buckland D, Kapadia N, Nichols M, Gao M, et al. Development and validation of machine learning models to predict admission from emergency department to inpatient and intensive care units. Ann Emerg Med 2021 Aug;78(2):290-302. [doi: 10.1016/j.annemergmed.2021.02.029] [Medline: 33972128]
14. Parker CA, Liu N, Wu SX, Shen Y, Lam SSW, Ong MEH. Predicting hospital admission at the emergency department triage: A novel prediction model. Am J Emerg Med 2019 Aug;37(8):1498-1504. [doi: 10.1016/j.ajem.2018.10.060] [Medline: 30413365]
15. Sun Y, Heng BH, Tay SY, Seow E. Predicting hospital admissions at emergency department triage using routine administrative data. Acad Emerg Med 2011 Aug;18(8):844-850 [FREE Full text] [doi: 10.1111/j.1553-2712.2011.01125.x] [Medline: 21843220]
16. Xie F, Liu N, Wu SX, Ang Y, Low LL, Ho AFW, et al. Novel model for predicting inpatient mortality after emergency admission to hospital in Singapore: retrospective observational study. BMJ Open 2019 Sep 26;9(9):e031382 [FREE Full text] [doi: 10.1136/bmjopen-2019-031382] [Medline: 31558458]
17. Jang D, Kim J, Jo Y, Lee J, Hwang J, Park S, et al. Developing neural network models for early detection of cardiac arrest in emergency department. Am J Emerg Med 2020 Jan;38(1):43-49. [doi: 10.1016/j.ajem.2019.04.006] [Medline: 30982559]
18. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care 2019 Feb 22;23(1):64 [FREE Full text] [doi: 10.1186/s13054-019-2351-7] [Medline: 30795786]
19. Kirby SE, Dennis SM, Jayasinghe UW, Harris MF. Patient related factors in frequent readmissions: the influence of condition, access to services and patient choice. BMC Health Serv Res 2010 Jul 21;10:216 [FREE Full text] [doi: 10.1186/1472-6963-10-216] [Medline: 20663141]
20. Sun BC, Hsia RY, Weiss RE, Zingmond D, Liang L, Han W, et al. Effect of emergency department crowding on outcomes of admitted patients. Ann Emerg Med 2013 Jun;61(6):605-611.e6 [FREE Full text] [doi: 10.1016/j.annemergmed.2012.10.026] [Medline: 23218508]
21. Ball JE, Murrells T, Rafferty AM, Morrow E, Griffiths P. 'Care left undone' during nursing shifts: associations with workload and perceived quality of care. BMJ Qual Saf 2014 Feb;23(2):116-125 [FREE Full text] [doi: 10.1136/bmjqs-2012-001767] [Medline: 23898215]
22. Teubner DJ, Considine J, Hakendorf P, Kim S, Bersten AD. Model to predict inpatient mortality from information gathered at presentation to an emergency department: The Triage Information Mortality Model (TIMM). Emerg Med Australas 2015 Aug;27(4):300-306. [doi: 10.1111/1742-6723.12425] [Medline: 26147765]
23. Abraham WT, Fonarow GC, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, OPTIMIZE-HF InvestigatorsCoordinators. Predictors of in-hospital mortality in patients hospitalized for heart failure: insights from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). J Am Coll Cardiol 2008 Jul 29;52(5):347-356 [FREE Full text] [doi: 10.1016/j.jacc.2008.04.028] [Medline: 18652942]
24. Fernandes M, Vieira S, Leite F, Palos C, Finkelstein S, Sousa J. Clinical decision support systems for triage in the emergency department using intelligent systems: a review. Artif Intell Med 2020 Jan;102:101762. [doi: 10.1016/j.artmed.2019.101762] [Medline: 31980099]
25. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019 May 13;1(5):206-215. [doi: 10.1038/s42256-019-0048-x]
26. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24 [FREE Full text] [doi: 10.2196/medinform.8912] [Medline: 29669706]
27. Ustun B, Rudin C. Supersparse linear integer models for optimized medical scoring systems. Mach Learn 2015 Nov 5;102(3):349-391. [doi: 10.1007/s10994-015-5528-6]
28. Ustun B, Rudin C. Learning Optimized Risk Scores. 2017 Presented at: KDD '17: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 2017; Halifax, NS, Canada p. 1125-1134. [doi: 10.1145/3097983.3098161]
29. Xie F, Chakraborty B, Ong MEH, Goldstein BA, Liu N. AutoScore: a machine learning-based automatic clinical score generator and its application to mortality prediction using electronic health records. JMIR Med Inform 2020 Oct 21;8(10):e21798 [FREE Full text] [doi: 10.2196/21798] [Medline: 33084589]
30. Xie F, Ong MEH, Liew JNMH, Tan KBK, Ho AFW, Nadarajan GD, et al. Development and assessment of an interpretable machine learning triage tool for estimating mortality after emergency admissions. JAMA Netw Open 2021 Aug 02;4(8):e2118467 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.18467] [Medline: 34448870]
31. Shen Y, Tay YC, Teo EWK, Liu N, Lam SW, Ong MEH. Association between the elderly frequent attender to the emergency department and 30-day mortality: A retrospective study over 10 years. World J Emerg Med 2018;9(1):20-25. [doi: 10.5847/wjem.j.1920-8642.2018.01.003] [Medline: 29290891]
32. Fong RY, Glen WSS, Mohamed Jamil AK, Tam WWS, Kowitlawakul Y. Comparison of the Emergency Severity Index versus the Patient Acuity Category Scale in an emergency setting. Int Emerg Nurs 2018 Nov;41:13-18. [doi: 10.1016/j.ienj.2018.05.001] [Medline: 29887281]
33. Electronic Health Intelligence System. IHIS. URL: https://www.ihis.com.sg/Project_Showcase/Healthcare_Systems/Pages/eHINTS.aspx [accessed 2022-03-03]

34. Dickson SJ, Dewar C, Richardson A, Hunter A, Searle S, Hodgson LE. Agreement and validity of electronic patient self-triage (eTriage) with nurse triage in two UK emergency departments: a retrospective study. Eur J Emerg Med 2022 Feb 01;29(1):49-55. [doi: 10.1097/MEJ.0000000000000863] [Medline: 34545027]
35. Levin S, Toerper M, Hamrock E, Hinson JS, Barnes S, Gardner H, et al. Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the Emergency Severity Index. Ann Emerg Med 2018 May;71(5):565-574.e2. [doi: 10.1016/j.annemergmed.2017.08.005] [Medline: 28888332]
36. Dugas AF, Kirsch TD, Toerper M, Korley F, Yenokyan G, France D, et al. An electronic emergency triage system to improve patient distribution by critical outcomes. J Emerg Med 2016 Jun;50(6):910-918. [doi: 10.1016/j.jemermed.2016.02.026] [Medline: 27133736]
37. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emerg Med J 2015 Mar 13;32(3):174-179 [FREE Full text] [doi: 10.1136/emermed-2013-203200] [Medline: 24421344]
38. Kraaijvanger N, Rijpsma D, Roovers L, van Leeuwen H, Kaasjager K, van den Brand L, et al. Development and validation of an admission prediction tool for emergency departments in the Netherlands. Emerg Med J 2018 Aug;35(8):464-470. [doi: 10.1136/emermed-2017-206673] [Medline: 29627769]
39. Mowbray F, Zargoush M, Jones A, de Wit K, Costa A. Predicting hospital admission for older emergency department patients: insights from machine learning. Int J Med Inform 2020 Aug;140:104163. [doi: 10.1016/j.ijmedinf.2020.104163] [Medline: 32474393]
40. Jeffery MM, D'Onofrio G, Paek H, Platts-Mills TF, Soares WE, Hoppe JA, et al. Trends in emergency department visits and hospital admissions in health care systems in 5 states in the first months of the COVID-19 pandemic in the US. JAMA Intern Med 2020 Oct 01;180(10):1328-1333 [FREE Full text] [doi: 10.1001/jamainternmed.2020.3288] [Medline: 32744612]
41. Liu N, Chee M, Niu C, Pek P, Siddiqui F, Ansah J, et al. Coronavirus disease 2019 (COVID-19): an evidence map of medical literature. BMC Med Res Methodol 2020 Jul 02;20(1):177 [FREE Full text] [doi: 10.1186/s12874-020-01059-y] [Medline: 32615936]
42. Chee ML, Ong MEH, Siddiqui FJ, Zhang Z, Lim SL, Ho AFW, et al. Artificial intelligence applications for COVID-19 in intensive care and emergency settings: a systematic review. Int J Environ Res Public Health 2021 Apr 29;18(9):4749 [FREE Full text] [doi: 10.3390/ijerph18094749] [Medline: 33947006]
43. Waring J, Lindvall C, Umeton R. Automated machine learning: review of the state-of-the-art and opportunities for healthcare. Artif Intell Med 2020 Apr;104:101822 [FREE Full text] [doi: 10.1016/j.artmed.2020.101822] [Medline: 32499001]
44. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. QJM 2001 Oct;94(10):521-526. [doi: 10.1093/qjmed/94.10.521] [Medline: 11588210]
45. Smith GB, Redfern OC, Pimentel MA, Gerry S, Collins GS, Malycha J, et al. The National Early Warning Score 2 (NEWS2). Clin Med (Lond) 2019 May;19(3):260 [FREE Full text] [doi: 10.7861/clinmedicine.19-3-260] [Medline: 31092526]
46. Rhee KJ, Fisher CJ, Willitis NH. The Rapid Acute Physiology Score. Am J Emerg Med 1987 Jul;5(4):278-282. [doi: 10.1016/0735-6757(87)90350-0] [Medline: 3593492]
47. Olsson T, Terent A, Lind L. Rapid Emergency Medicine score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients. J Intern Med 2004 May;255(5):579-587 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01321.x] [Medline: 15078500]
48. Churpek MM, Yuen TC, Park SY, Meltzer DO, Hall JB, Edelson DP. Derivation of a cardiac arrest prediction model using ward vital signs*. Crit Care Med 2012 Jul;40(7):2102-2108 [FREE Full text] [doi: 10.1097/CCM.0b013e318250aa5a] [Medline: 22584764]
49. Breiman L. Random Forests. Machine Learning 2001;45(1):5-32. [doi: 10.1023/a:1010933404324]
50. Oshiro T, Perez P, Baranauskas J. How Many Trees in a Random Forest? In: Machine Learning and Data Mining in Pattern Recognition. Berlin, Heidelberg: Springer; 2012.
51. Probst P, Boulesteix AL. To tune or not to tune the number of trees in random forest. J Mach Learn Res 2017;18(1):6673-6690 [FREE Full text]
52. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. J R Stat Soc Series B Stat Methodol 2018 Dec 05;58(1):267-288. [doi: 10.1111/j.2517-6161.1996.tb02080.x]
53. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 May 28;521(7553):436-444. [doi: 10.1038/nature14539] [Medline: 26017442]
54. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. NPJ Digit Med 2018;1:18 [FREE Full text] [doi: 10.1038/s41746-018-0029-1] [Medline: 31304302]
55. Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. J Am Med Inform Assoc 2018 Oct 01;25(10):1419-1428 [FREE Full text] [doi: 10.1093/jamia/ocy068] [Medline: 29893864]
56. Sherstinsky A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. Physica D: Nonlinear Phenomena 2020 Mar;404:132306. [doi: 10.1016/j.physd.2019.132306]
57. Murtagh F. Multilayer perceptrons for classification and regression. Neurocomputing 1991 Jul;2(5-6):183-197. [doi: 10.1016/0925-2312(91)90023-5]

Abbreviations

AI: artificial intelligence
ED: emergency department
EHR: electronic health record
ICU: intensive care unit
PACS: Patient Acuity Category Scale
ROC: receiver operating characteristic
SERT: System for Emergency Risk Triage

Edited by G Eysenbach; submitted 11.10.21; peer-reviewed by Y Wang, T Ndabu; comments to author 23.11.21; revised version received 29.11.21; accepted 30.11.21; published 25.03.22.

Please cite as:
Liu N, Xie F, Siddiqui FJ, Ho AFW, Chakraborty B, Nadarajan GD, Tan KBK, Ong MEH
Leveraging Large-Scale Electronic Health Records and Interpretable Machine Learning for Clinical Decision Making at the Emergency Department: Protocol for System Development and Validation
JMIR Res Protoc 2022;11(3):e34201
URL: https://www.researchprotocols.org/2022/3/e34201
doi:10.2196/34201
PMID:35333179

Protocol

Economic Burden of Chronic Obstructive Pulmonary Disease and Lung Cancer Between 2000 and 2015 in Saskatchewan: Study Protocol

Erika Dianne Penz[1,2], SM, MD, MSc; Benjamin John Fenton[1,2]; Nianping Hu[3], MSc, PhD; Darcy Marciniuk[1,2], MD

[1]Division of Respirology, Critical Care and Sleep Medicine, University of Saskatchewan, Saskatoon, SK, Canada

[2]Respiratory Research Centre, University of Saskatchewan, Saskatoon, SK, Canada

[3]Saskatchewan Health Quality Council, Saskatoon, SK, Canada

Corresponding Author:
Erika Dianne Penz, SM, MD, MSc
Division of Respirology, Critical Care and Sleep Medicine
University of Saskatchewan
Room 537 Ellis Hall
103 Hospital Drive
Saskatoon, SK, S7N OW8
Canada
Phone: 1 306 844 1140
Fax: 1 306 844 1532
Email: erika.penz@usask.ca

Abstract

Background: Chronic obstructive pulmonary disease (COPD) and lung cancer are both detrimental diseases that present great burdens on society. Years of life lost (YLL), premature years of life lost (PYLL), working years lost (WYL), and productivity loss are all effective measures in identifying economic burden of disease.

Objective: We propose a population-based study to analyze comprehensive provincial cohorts of Saskatchewan residents with COPD, lung cancer, and combined COPD and lung cancer in order to identify the burden these diseases present.

Methods: Saskatchewan residents over the age of 35 years who had COPD, lung cancer, or both, between January 1, 2000, and December 31, 2015, will be identified and used in this study. Data for analysis including age, gender, and date of death, alongside Statistics Canada income estimates, will be used to estimate productivity loss and WYL. Statistics Canada life tables will be used to calculate YLL and PYLL by subtracting the patients' ages at death by their life expectancies, adjusted using sex and age at death. We will link the Saskatchewan cancer registry with Saskatchewan health administrative databases to create three cohorts: (1) COPD; (2) lung cancer; and (3) COPD and lung cancer. Individuals with lung cancer will be identified using ICDO-T (International Classification of Diseases for Oncology-Topography) codes, and those with COPD will be defined and identified as individuals who had at least 1 visit to a physician with a diagnosis of COPD or 1 hospital separation with a diagnosis of COPD. Those without a valid health care coverage for a consecutive 12 months prior to the first diagnostic code will be excluded from the study. Those with a combined diagnosis of COPD and lung cancer will be identified as individuals who were diagnosed with COPD in the 12 months following their lung cancer diagnosis or anytime preceding their lung cancer diagnosis.

Results: As of April 2021, we have had access to all relevant data for this study, have received funding (January 2020), and have begun the preliminary analysis of our data set.

Conclusions: It is well documented that COPD and lung cancer are both destructive diseases in terms of YLL, PYLL, WYL, and productivity loss; however, no studies have been conducted to analyze a cohort with combined COPD and lung cancer. Understanding the economic burden associated with each of our 3 cohorts is necessary in understanding and thus reducing the societal impact of COPD and lung cancer.

International Registered Report Identifier (IRRID): RR1-10.2196/31350

(JMIR Res Protoc 2022;11(3):e31350) doi:10.2196/31350

KEYWORDS

lung cancer; COPD; chronic obstructive pulmonary disease; productivity loss; years of life lost; premature years of life lost; working years lost; economic burden of disease; lung disease; health economics; Stats Canada; epidemiology; pulmonary disease; pulmonary health; disease burden

Introduction

Chronic obstructive pulmonary disease (COPD) and lung cancer both pose great burdens on society in terms of years of life lost (YLL), premature years of life lost (PYLL), absenteeism, and productivity loss. Lung cancer is the leading cause of cancer mortality in Canada [1], and COPD was the third leading cause of death worldwide in 2010 [2]. Both diseases are predominant in older age groups, with the majority of patients receiving diagnosis in their sixties [3,4].

Due to the impact that COPD and lung cancer have on the respiratory system and symptoms, individuals are often greatly affected in their ability to work, and many withdraw from the workforce early due to disability or death [5]. It has been found that lung cancer is a significant contributor to YLL, accounting for nearly as many YLL as colon, breast, and prostate cancers combined [6]. COPD has also been shown to account for 1.8% of all deaths in Poland between 1999 and 2014. In addition, the typical patient with COPD lost roughly 14.5 years of life due to the disease [7]. The authors concluded that since smoking remains the leading cause of COPD, public education on the risks of smoking could have a considerable effect on reducing the burden of COPD in Poland [7]. A study examining mortality trends in Ontarians with COPD from 1996 to 2012 found that PYLL due to COPD decreased over the period, while COPD prevalence increased. The authors posited that the decrease in PYLL due to COPD is likely driven by improvements in both cardiovascular prognosis and COPD management. It also demonstrated that the age-sex standardized mortality rates for lung cancer in the COPD population decreased between 1996 and 2009, although the absolute number of people with COPD dying of lung cancer increased over the same period [8]. Despite the positive trend in YLL over this time period, the authors reported that PYLL for individuals with COPD were nearly 5 times higher than that of the non-COPD population.

Although YLL is a measure of the impact of a disease on death in a population, productivity loss is another important measure of the impact of a disease, reflected in measures such as absenteeism and WYL due to disability or premature death [9-11]. It has been found that, of the total direct and indirect costs associated with COPD, productivity loss accounts for nearly one-third of total associated costs [11]. A study conducted in Spain in 2019 found that lung cancer was responsible for 60,846 WYL and €13.1 billion (US $14.8 billion) in productivity loss between the years 2008 and 2017 [9]. However, the impact of a combined diagnosis of COPD and lung cancer, a common clinical occurrence, on YLL or productivity loss has not been well described. Given factors such as symptom burden and potentially increased comorbidities from shared risk factors, one would intuitively expect the true burden to be even greater. It has been established that COPD is a risk factor for lung cancer, even after controlling for smoking [12]. Additionally, a high proportion of patients with lung cancer have concomitant COPD [13].

Saskatchewan is a Canadian prairie province with a population of 1.098 million, covering 588,000 square kilometers with a population density of 1.9 people per square kilometer [14]. In 2017, smoking prevalence in Saskatchewan was 17.8% compared with 16.9% in 2015 and was above the national average of 15.1% [15]. Among youth, Saskatchewan has roughly double the national smoking rates [15]. Due to a growing and aging population, the number of cases of COPD and lung cancer continues to rise, and as a consequence, both diseases are significant public health concerns. We believe the results from this work would be both interesting and informative to many beyond this geographical area. Identifying the extent of COPD and lung cancer's economic burden on Saskatchewan is crucial in understanding these diseases and what actions are appropriate in reducing their impact on society [16].

The objective of this study is to investigate the economic burden associated with COPD and lung cancer by estimating YLL, PYLL, WYL and productivity loss among a cohort of COPD and lung cancer patients in Saskatchewan diagnosed between the years 2000 and 2015.

Methods

Database and Data Linkage

We will link the Saskatchewan cancer registry with Saskatchewan health administrative databases to create three cohorts of patients: (1) COPD; (2) lung cancer; and (3) COPD and lung cancer between January 1, 2000, and December 31, 2015. The Saskatchewan cancer registry contains, among other variables, unique health services number, date of birth (year/month), date of diagnosis, date of death, primary cause of death, ICDO-T (International Classification of Diseases for Oncology-Topography) and ICDO-M (International Classification of Diseases for Oncology-Morphology) code of the primary tumor, and TNM overall stage. eHealth Saskatchewan is a government agency mandated to manage the information technology needs of the Saskatchewan Health Authority, which is responsible for and is the sole provider of government "single-payer" health care delivery in the province. It collects, combines, stores, and manages the electronic health records of Saskatchewan residents. Via eHealth Saskatchewan, the cancer registry data will be linked with the following Saskatchewan health administrative databases: Person Health Registry System, Discharge Abstracts Databases, and Physician Services Claims File. The Person Health Registry System contains unique health services number (encrypted), year of birth, sex, dates of health insurance coverage, and reason for termination. The Discharge Abstracts Databases contain health services number (encrypted), year and month of birth, date of admission, date of discharge, and discharge diagnosis (International Classification of Diseases 9th Revision [ICD-9]

or 10th Revision [ICD-10], all fields). The Physician Services Claims file contains health services number (encrypted), diagnostic code (ICD) associated with service, service code, and date of service.

Case Definitions

Lung Cancer

Using data from the Saskatchewan Cancer Agency registry, individuals 35 years of age or older diagnosed with small cell lung cancer (SCLC) or non–small cell lung cancer (NSCLC) between January 1, 2000, and December 31, 2015, will be included in the study. NSCLC and SCLC will be identified using the ICDO-T codes (Tables 1 and 2 in Multimedia Appendix 1). Individuals who were diagnosed at autopsy or death, who did not reside in Saskatchewan at the time of diagnosis, or who had a prior or subsequent cancer (except basal cell carcinoma or squamous cell carcinoma skin cancer) will be excluded.

Chronic Obstructive Pulmonary Disease

Individuals with COPD will be identified from Saskatchewan health administrative databases. COPD will be defined as an individual aged 35 years and older having at least 1 visit to a physician with a diagnosis of COPD in the first diagnostic field, or 1 hospital separation with a diagnosis of COPD in any diagnostic field, coded by the ICD-9-CM 491-492 (chronic bronchitis, emphysema), 496 (chronic airway obstruction, not otherwise specified), or ICD-10-CA J41-44 (chronic bronchitis) [17] until December 31, 2015. Individuals meeting these criteria prior to January 1, 2000, will be considered prevalent COPD cases while individuals meeting these criteria between January 1, 2000, and December 31, 2015, will be considered incident cases. Those without valid health care coverage for a consecutive 12 months prior to the first diagnostic code will be excluded from the study.

Chronic Obstructive Pulmonary Disease and Lung Cancer

All individuals identified with lung cancer between 2000 and 2015 who also meet the criteria for COPD in the 12 months following or anytime preceding their lung cancer diagnosis will be classified as a third cohort of individuals with both diagnoses.

Population Cohorts and Time Frame

The COPD cohort will be as follows: adults aged 35 years or older diagnosed with COPD as of January 1, 2000, plus all new COPD cases identified until December 31, 2015.

The lung cancer cohort will be as follows: Adults aged 35 years or older diagnosed with NSCLC or SCLC between January 1, 2000, and Dec 31, 2015.

The COPD and lung cancer cohort is as follows: Adults aged 35 years or older diagnosed with NSCLC or SCLC between January 1, 2000, and Dec 31, 2015, who also meet the criteria for COPD in the 12 months following or anytime preceding a lung cancer diagnosis.

All individuals will be followed to the end of study (December 31, 2018), end of coverage, or death, whichever occurs first.

Outcomes

Death

Date of death will be obtained from 2 sources, the cancer registry for the lung cancer and combined lung cancer and COPD cohorts, and the Saskatchewan personal health registry database for the COPD cohort. These data are updated quarterly from eHealth.

Years of Life Lost

YLL will be calculated for each cohort by subtracting the subject's age at death from their Saskatchewan age-specific and sex-specific life expectancy in their year of death. Life expectancy will be obtained from published life tables through Statistics Canada [18]. Average YLL will be calculated for all 3 cohorts by dividing YLL by the number of individuals who died in each cohort. YLL and average YLL will be stratified by gender, as well as type and stage of lung cancer.

Premature Life Years Lost

Premature life years lost (PYLL) will be reported for each cohort including only those individuals who died before their expected age-adjusted and sex-adjusted life expectancy at the time of death.

Average PYLL will be calculated for each cohort by dividing total PYLL by the number of individuals in each cohort who died prematurely. These results will be stratified by gender as well as type and stage of lung cancer.

Working Years Lost and Productivity Loss

Working years lost (WYL) will be calculated by subtracting the date of death from the birth date in the expected retirement year. If the value is positive, this is recorded as the WYL. If the value is negative, this is recorded as zero.

Productivity loss will be calculated by multiplying an individual's WYL, if positive, by their age-adjusted and sex-adjusted income using Saskatchewan-specific income statistics from Statistics Canada [19].

Base Case Model and Scenario Analysis

The base case model will use an expected retirement age of 65 years for both men and women for calculation of WYL. Median income will be used to calculate productivity loss. Scenario analyses will be conducted to explore difference in WYL and productivity loss according to different age at retirement as well as the use of mean income statistics from Statistics Canada. The scenarios to be analyzed are as follows: scenario 1—WYL based on retirement at age 63 years for men and 61 years for women [20]; scenario 2—WYL based on retirement at 55 years of age; and scenario 3—WYL based on retirement at 70 years of age.

Labor Force Participation

To adjust the estimated productivity loss for each cohort because not all individuals are employed in the labor force, we will incorporate age-specific and sex-specific labor force participation rates from Statistics Canada into our outcomes for each cohort (ie, expected WYL and expected productivity loss).

Inflation

We will report all productivity loss estimates in 2018 Canadian dollars. In the case of WYL prior to 2018, age-specific and sex-specific incomes will be obtained directly from Statistics Canada, which are provided in inflation-adjusted 2018 Canadian dollars. In the case of WYL after 2018, income will be inflated based on average age-specific and sex-specific growth rates calculated over the time period of 2000-2018.

Statistical Methods

Productivity Loss Model Structure

In order to identify lost productivity due to premature death, we must assess the affected individuals' lost incomes. Lost productivity will be measured through 2 specific metrics, which are productivity loss and WYL. We will report averages and totals, as well as the expected values for both metrics. The results will be stratified by gender and stage of lung cancer (for lung cancer and combined groups). In the first and last years of lost income, partial WYL will be calculated.

Formulas

We will calculate working years lost, year at expected retirement, total productivity loss, and expected total productivity loss. The total productivity loss formula will sum income adjusted using sex (s) and age (a) for years between year at death and year at expected retirement, while the expected total productivity loss formula will adjust for sex-specific and age-specific labor force participation rates as well. The formulas we will use are as follows:

Working years lost = expected retirement age - age at death;

YAR = working years lost + calendar year at death;

Where YAD is years between calendar year at death, YAR is calendar year at expected retirement, and LFPR is labor force participation rates.

Partial Years Correction

The time between date of death and December 31 in the year of death, and the time between January 1 and the individual's birthday in the expected year of retirement will be calculated as partial years in order to estimate productivity loss more precisely. For example, if someone died on December 1, 2015, the WYL and income attributed to that year would be equal to 30/365 for WYL and (30/365) multiplied by age-adjusted and sex-adjusted 2015 median income for productivity loss in 2015. All years between year at death and year at retirement will be accounted as full WYL.

Ethics Approval

Ethics exemption by the Research Ethics Board was received on May 31, 2017 (Bio 17-153).

Results

As of April 2021, we have gained access to all relevant data for this study, have received funding from the Lung Association of Saskatchewan (January 2020), and have begun preliminary analysis of our data set.

Discussion

Comparison With Prior Work

This study's aim is to estimate the economic burden, by way of YLL, PYLL, and productivity loss among COPD and lung cancer patients in Saskatchewan using comprehensive government health administrative and cancer registry data. Although prior work estimating indirect costs has been reported for COPD and lung cancer populations [2,9,11], there has not been a comparison between the 2 diseases over the same time period, from the identical general sample population for the extended time period we studied or describing indirect costs in a population of individuals with both diagnoses.

By understanding indirect costs associated with COPD and lung cancer through measures of PYLL, WYL, and productivity loss, these estimates can help inform policy makers, public health professionals, and the public to better understand the burden associated with lung disease experienced outside the health system. These costs are frequently missed or minimized even though they may be substantial to both the individual and the population. Moreover, from a governmental point of view, quantifying the magnitude of the economic impact of these diseases can assist policy makers when deciding on priorities related to health service delivery (eg, expanding access to diagnostic services to enable early detection) and efforts to improve prevention strategies (eg, expanding smoking cessation programs).

According to previous literature, lung cancer and COPD are comparable in terms of average YLL and average PYLL, so we expect to find similar results in those measures [6,7]. Due to an expected larger number of patients diagnosed with COPD compared with lung cancer, the total YLL, total PYLL, and total productivity loss will be highest in the COPD group. The benefit of reporting the aggregate productivity loss as well as YLL for each disease will be to estimate the overall magnitude of economic costs over the study period. If the studies performed by Kim et al [11] and Darba and Marsa [9] are indicators of what to expect from our data, productivity loss due to these diseases could exceed CAD $1 billion (US $788 million). It is unclear what the economic costs associated with the combined COPD and lung cancer group will be relative to either diagnosis alone; however, we hypothesize that mean YLL (as well as PYLL), WYL, and productivity loss may be greater than either diagnosis alone.

Limitations

Despite using sex-adjusted and age-adjusted average and median incomes for Saskatchewan residents from Statistics Canada for the productivity loss estimates in our study, total and average productivity loss will not be based on the individuals' actual income levels; this is a limitation with our analysis. Another

limitation is that productivity loss in our study does not take into account the time lost due to illness (as opposed to death), nor does it account for time away from work and the associated productivity loss attributed to caregivers. Lastly, life years and premature life years lost will be calculated using age-adjusted and sex-adjusted life tables from Statistics Canada in relation to death from any cause in each of the cohorts; therefore, YLL, and PYLL may not completely be attributed to COPD, lung cancer, or both diagnoses entirely.

Acknowledgments

Funding for this work was provided by a Lung Association of Saskatchewan research grant.

Availability of Data and Materials

The data sets analyzed during this study are not publicly available due to the terms and requirements of the data sharing agreements governing health administrative data access in Saskatchewan.

Disclaimer

This study is based in part on de-identified data provided by the Saskatchewan Ministry of Health, eHealth Saskatchewan, and the Saskatchewan Cancer Agency. The interpretation and conclusions contained herein do not necessarily represent those of the Government of Saskatchewan, Saskatchewan Ministry of Health, eHealth Saskatchewan, or the Saskatchewan Cancer Agency.

Authors' Contributions

EDP was a major contributor to the study design and in writing the protocol. DM contributed to study conception and the writing of the protocol. NH designed the statistical model and revised and edited the final draft of the protocol. BJF contributed to the study design and was a major contributor in writing the protocol.

Conflicts of Interest

EDP has received research funds paid to her institution from AstraZeneca and Saskatchewan Cancer Agency, unrelated to this work. She has received consulting fees and honoraria for participation on advisory boards, lecture series, educational events from AstraZeneca, GlaxoSmithKline, Sanofi, and Boehringer Ingelheim, unrelated to this work. She is cochair of the Canadian Thoracic Society COPD Assembly and Advisory Board member for the Institute of Cancer Research for the Canadian Institutes for Health Research, all unpaid work. DM has undertaken consulting with Alberta Health Services, Health Canada, Lung Association of Saskatchewan, Ontario Ministry of Health and Long-Term Care, Saskatchewan Health Authority, and Yukon Health and Social Services. He has provided research advisory and received research funding (held and managed by the University of Saskatchewan) from AstraZeneca, Boehringer Ingelheim, Canadian Institute of Health Research, GlaxoSmithKline, Grifols, Lung Association of Saskatchewan, Lung Health Institute of Canada, Novartis, Sanofi, Saskatchewan Health Research Foundation, and Schering-Plough. DM is an employee of the University of Saskatchewan and serves as Deputy Editor, CHEST Journal.

Multimedia Appendix 1
International Classification of Diseases for Oncology Topography and Morphology definitions.
[DOCX File , 20 KB - resprot_v11i3e31350_app1.docx]

References

1. Canadian cancer statistics 2011. Toronto, ON, Canada: Canadian Cancer Society; May 2011.
2. Quaderi SA, Hurst JR. The unmet global burden of COPD. Glob Health Epidemiol Genom 2018;3:e4 [FREE Full text] [doi: 10.1017/gheg.2018.1] [Medline: 29868229]
3. Shavelle RM, Paculdo DR, Kush SJ, Mannino DM, Strauss DJ. Life expectancy and years of life lost in chronic obstructive pulmonary disease: findings from the NHANES III Follow-up Study. Int J Chron Obstruct Pulmon Dis 2009;4:137-148 [FREE Full text] [doi: 10.2147/copd.s5237] [Medline: 19436692]
4. Hsu L, Chu N, Liu C, Tsai SYC, You D, Ko J, et al. Sex-associated differences in non-small cell lung cancer in the new era: is gender an independent prognostic factor? Lung Cancer 2009 Nov;66(2):262-267. [doi: 10.1016/j.lungcan.2009.01.020] [Medline: 19299032]
5. GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Respir Med 2017 Sep;5(9):691-706 [FREE Full text] [doi: 10.1016/S2213-2600(17)30293-X] [Medline: 28822787]
6. Brustugun OT, Møller B, Helland A. Years of life lost as a measure of cancer burden on a national level. Br J Cancer 2014 Aug 26;111(5):1014-1020 [FREE Full text] [doi: 10.1038/bjc.2014.364] [Medline: 24983370]

7. Dziankowska-Zaborszczyk E, Bryla M, Ciabiada-Bryla B, Maniecka-Bryla I. Standard expected years of life lost (SEYLL) due to chronic obstructive pulmonary disease (COPD) in Poland from 1999 to 2014. PLoS One 2019;14(3):e0213581 [FREE Full text] [doi: 10.1371/journal.pone.0213581] [Medline: 30861024]
8. Gershon A, Hwee J, Victor JC, Wilton A, Wu R, Day A, et al. Mortality trends in women and men with COPD in Ontario, Canada, 1996-2012. Thorax 2015 Feb;70(2):121-126. [doi: 10.1136/thoraxjnl-2014-205956] [Medline: 25428609]
9. Darbà J, Marsà A. The cost of lost productivity due to premature lung cancer-related mortality: results from Spain over a 10-year period. BMC Cancer 2019 Oct 23;19(1):992 [FREE Full text] [doi: 10.1186/s12885-019-6243-7] [Medline: 31646991]
10. Bradley CJ, Yabroff KR, Dahman B, Feuer EJ, Mariotto A, Brown ML. Productivity costs of cancer mortality in the United States: 2000-2020. J Natl Cancer Inst 2008 Dec 17;100(24):1763-1770 [FREE Full text] [doi: 10.1093/jnci/djn384] [Medline: 19066273]
11. Kim C, Kim Y, Yang DW, Rhee CK, Kim SK, Hwang YI, et al. Direct and Indirect Costs of Chronic Obstructive Pulmonary Disease in Korea. Tuberc Respir Dis (Seoul) 2019 Jan;82(1):27-34 [FREE Full text] [doi: 10.4046/trd.2018.0035] [Medline: 30302958]
12. Durham AL, Adcock IM. The relationship between COPD and lung cancer. Lung Cancer 2015 Nov;90(2):121-127 [FREE Full text] [doi: 10.1016/j.lungcan.2015.08.017] [Medline: 26363803]
13. Wang W, Dou S, Dong W, Xie M, Cui L, Zheng C, et al. Impact of COPD on prognosis of lung cancer: from a perspective on disease heterogeneity. Int J Chron Obstruct Pulmon Dis 2018;13:3767-3776 [FREE Full text] [doi: 10.2147/COPD.S168048] [Medline: 30538439]
14. Focus on geography series, 2016 census. Ottawa: Statistics Canada; 2017. URL: https://tinyurl.com/5859u9py [accessed 2020-07-06]
15. Reid J, Hammond D, Rynard V, Madill C, Burkhalter R. Tobacco use in Canada: patterns and trends. Waterloo, ON, Canada: University of Waterloo, Propel Centre for Population Health Impact; 2019.
16. Goodridge D, Lawson J, Duggleby W, Marciniuk D, Rennie D, Stang M. Health care utilization of patients with chronic obstructive pulmonary disease and lung cancer in the last 12 months of life. Respir Med 2008 Jun;102(6):885-891 [FREE Full text] [doi: 10.1016/j.rmed.2008.01.007] [Medline: 18313278]
17. Gershon AS, Wang C, Guan J, Vasilevska-Ristovska J, Cicutto L, To T. Identifying individuals with physcian diagnosed COPD in health administrative databases. COPD 2009 Oct;6(5):388-394. [doi: 10.1080/15412550903140865] [Medline: 19863368]
18. Archived - life expectancy and other elements of the life table, Canada and provinces. Statistics Canada. Ottawa; 2018 Jun 28. URL: https://tinyurl.com/457e5ctj [accessed 2020-07-19]
19. Archived - labour force characteristics by sex and detailed age group, annual, inactive. Statistics Canada. Ottawa; 2019 Jan 04. URL: https://tinyurl.com/mr2kaked [accessed 2020-07-19]
20. Retirement age by class of worker, annual. Statistics Canada. Ottawa; 2019. URL: https://tinyurl.com/2p8bfn7m [accessed 2019-07-27]

Abbreviations

COPD: chronic obstructive pulmonary disease
ICD: International Classification of Diseases
ICDO-M: International Classification of Diseases for Oncology-Morphology
ICDO-T: International Classification of Diseases for Oncology-Topography
NSCLC: non–small cell lung cancer
PYLL: premature years of life lost
SCLC: small cell lung cancer
WYL: working years lost
YLL: years of life lost

Edited by G Eysenbach; submitted 19.07.21; peer-reviewed by P Zarogoulidis, C Shah; comments to author 18.12.21; revised version received 28.12.21; accepted 28.12.21; published 04.03.22.

Please cite as:
Penz ED, Fenton BJ, Hu N, Marciniuk D
Economic Burden of Chronic Obstructive Pulmonary Disease and Lung Cancer Between 2000 and 2015 in Saskatchewan: Study Protocol
JMIR Res Protoc 2022;11(3):e31350
URL: https://www.researchprotocols.org/2022/3/e31350
doi:10.2196/31350
PMID:35254280

Protocol

Providing Accessible Recreation Outdoors—User-Driven Research on Standards (PARCOURS): Protocol for a Multiphase Study

Mike Prescott[1,2], PhD; François Routhier[1,3], PEng, PhD; Delphine Labbé[4,5,6], PhD; Marie Grandisson[1,3], PhD; Atiya Mahmood[5], PhD; Ernesto Morales[1,3], PhD; Krista L Best[1,3], PhD; Mir Abolfazl Mostafavi[1,2,7], PhD; Jaimie Borisoff[8,9], PhD; Stéphanie Gamache[1], OT, PhD; Bonita Sawatzky[10], PhD; William C Miller[6,9,11], PhD; Laura Yvonne Bulk[6], PhD; Julie M Robillard[12], PhD; Hailey-Thomas Jenkins[5], MSc; Kishore Seetharaman[5], MSc; W Ben Mortenson[6,9,11], BScOT, MSci, PhD

[1]Center for Interdisciplinary Research in Rehabilitation and Social Integration, Centre intégré universitaire de santé et de services sociaux de la Capitale-Nationale, Québec City, QC, Canada

[2]Centre for Research in Geospatial Data and Intelligence, Université Laval, Québec City, QC, Canada

[3]Department of Rehabilitation, Université Laval, Québec City, QC, Canada

[4]Disability and Human Development Department, University of Illinois, Chicago, IL, United States

[5]Department of Gerontology, Simon Fraser University, Vancouver, BC, Canada

[6]Department of Occupational Science & Occupational Therapy, University of British Columbia, Vancouver, BC, Canada

[7]Department of Geomatics Sciences, Université Laval, Québec City, QC, Canada

[8]Rehabilitation Engineering Design Lab, British Columbia Institute of Technology, Burnaby, BC, Canada

[9]International Collaboration on Repair Discoveries, Vancouver, BC, Canada

[10]Department of Orthopedics, University of British Columbia, Vancouver, BC, Canada

[11]GF Strong Rehabilitation Research Program, Vancouver, BC, Canada

[12]Department of Medicine, Division of Neurology, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
W Ben Mortenson, BScOT, MSci, PhD
International Collaboration on Repair Discoveries
818 W 10th Ave
Vancouver, BC, V5Z 1M9
Canada
Phone: 1 604 675 8870
Email: ben.mortenson@ubc.ca

Abstract

Background: Canada's national parks are world-renowned. However, despite recent attempts to improve access, many are not accessible to people with disabilities. With the advent of provincial and federal legislation, standards are being developed to assist with the design and management of parks.

Objective: The overarching objective of this study is to inform accessibility standards for federal parks that meet the needs of all park visitors, regardless of ability. The specific objectives of this study are to identify park accessibility standards that exist internationally, identify the accessibility challenges that people with disabilities face in park environments, and prioritize and recommend accessibility standards for national parks.

Methods: A 3-phase approach will be used to achieve the study objectives. In the first phase, a scoping review of the existing accessibility standards will be conducted. The second phase will include objective audits of trails and features in 6 parks, 3 in western Canada and 3 in eastern Canada, as well as mobile interviews with 24 diverse participants in each region regarding their experiences of and recommendations for improving the park's accessibility. In the final phase, a Delphi participatory consensus development process will be used, based on the data gathered in the first 2 phases, to prioritize recommendations for standards.

Results: We expect to find gaps in existing standards that do not account for the diverse range of accessibility requirements that people with disabilities have for visiting parks. We also expect to find that existing standards, on their own, may not be enough to ensure equitable access to all the experiences and amenities that parks have to offer. Development of subsequent

guidelines and best practices may be necessary to address complex scenarios for which standards may not be the best approach to ensuring accessibility.

Conclusions: The participatory and mixed methods approaches used in this study will provide rich insights for developing accessible park standards that consider the diverse needs of people with disabilities. The findings will also support the development or enhancement of park standards at all levels of government.

International Registered Report Identifier (IRRID): DERR1-10.2196/33611

(JMIR Res Protoc 2022;11(3):e33611) doi:10.2196/33611

KEYWORDS

parks; accessibility; standards; user-oriented research

Introduction

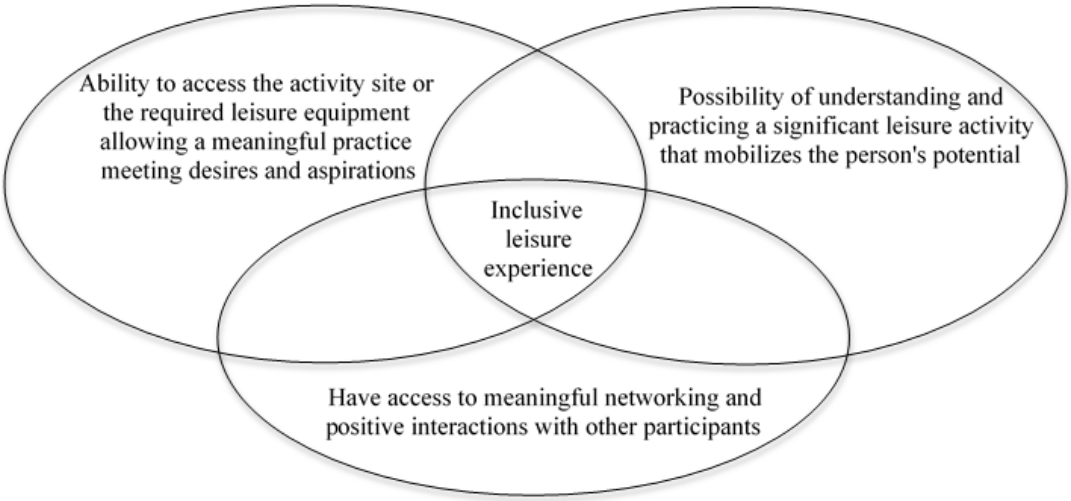
Background

Outdoor natural parks offer a variety of experiences that result in physical and psychological [1-3] as well as social and health benefits of access to green and blue spaces [3-10]. However, people with disabilities, who represent 22.3% of the population in Canada [11] and 25.7% in the United States [12], are often excluded from these spaces because of accessibility issues [13]. Limited access to outdoor spaces further contributes to the inequities that people with disabilities already face in employment, housing, and health care [14]. In response to challenges in built and natural environments, the Canadian federal government has enacted a legislation called the Accessible Canada Act to remove barriers to participation for people with disabilities [15]. This legislation includes a road map for developing accessibility standards that regulate organizations under federal responsibility, such as national parks. The intent is that these national standards will be adopted at all levels of government so that people with disabilities can expect the same level of service from every park they visit.

Historically, standards have focused on promoting access for people with physical disabilities, often neglecting the needs of people who experience cognitive, sensory, or multiple disabilities [16]. For example, wayfinding is emerging as a critical topic for different disabilities to identify accessible routes for planning purposes and to enable real-time navigation. Specific wayfinding standards also need to consider the dynamic nature of the environment to foster accessibility during trail construction, snowmelt, or massive rainstorms. In these instances, it is important that information about alternate routes consider accessibility requirements. The challenge of developing standards is compounded by the variety of assistive devices used by and capacities of people with diverse disabilities. Currently, some people with disabilities are excluded because the size of their mobility devices exceeds the space provided under existing building codes [17]. The use of hand cycles (3- or 4-wheeled cycles propelled by the arms rather than the legs) in parks will also require standards that mandate wider paths for turning [18]. Further complications arise when designs that meet the needs of one group conflict with the needs of others. For example, tactile paving at crosswalks, which warns those with vision loss that they are entering a street, can be hazardous for manual wheelchair users who may find that it is uncomfortable or causes pain [19] or might precipitate a fall [20]. Therefore, it is important to develop standards that consider and involve people with a wide variety of disabilities in the process.

Carbonneau et al [21] developed a conceptual framework (Figure 1) that promotes inclusive leisure experiences for people with disabilities in their communities. Being present and participating in some aspect or aspects of an activity does not guarantee the quality of the experience [21]. Not only do we need to consider the physical components of access, but we also have to take into consideration the significance of the activity for the participant and the necessity of positive interactions with other participants. Thus, if we wish to inform standards to make parks more inclusive, the experiences and preferences of people with disabilities need to be understood.

Figure 1. Components of an inclusive leisure experience [22].



Objectives

The overarching purpose of this participatory project is to inform the development of standards to make parks more accessible. The acronym for the project entitled Providing Accessible Recreation Outdoors: User-driven Research on Standards is PARCOURS. This is a French word for *trail*, that is, *un chemin pour aller d'un point à un autre*, which emphasizes our project's bilingual focus on developing standards to improve accessibility in parks across Canada. The specific objectives of this study are as follows: (1) to identify park accessibility standards that exist internationally; (2) to identify the accessibility challenges that people with disabilities face in park environments; and (3) to prioritize and recommend accessibility standards for national parks.

Methods

Overview

The research will be conducted in 3 phases over a 24-month period and in two provinces: British Columbia and Québec. The methods described in this proposal will be the same in both provinces. Advisory committees including individuals with a variety of disabilities have been created in both provinces (one in each province) to ensure the consideration of inputs or concerns of these individuals in the research project through a participatory research approach. These committees include individuals with mobility, visual, and hearing disabilities; intellectual disabilities; autism spectrum disorder; dementia; and Alzheimer disease. The committees will meet 2 to 3 times a year to provide feedback on the ongoing phases of the research project that are presented below. They will help us fine-tune the protocol.

A scoping review of existing standards will be conducted in the *first phase*. In the *second phase*, park audits and mobile interviews will be conducted with people who have a wide range of disabilities at 3 parks in each province in the summer and winter. The *third phase* will use the data collected in the first 2 phases to inform the selection and prioritization of park standards using a Delphi process. A final report will be presented to the granting agency (Accessibility Standards Canada) at the end of the study. The study has received approvals by research ethics boards in both provinces.

Phase 1: Scoping Review

The objective of the scoping review will be to compare and contrast existing international and national standards, along with novel research evidence to inform the development of revised standards. The scoping review will involve five steps: (1) identifying the research question, (2) identifying the key words, (3) identifying relevant standards and guidelines, (4) choosing standards, and (5) charting the data and reporting the results [22]. The question guiding the search for relevant studies will be, "What are the current accessibility standards in terms of outdoor spaces, including parks to allow people with disabilities to enjoy the natural environments?" The scoping review was conducted between June 2020 and February 2021 using Google search and governmental or official park websites. The search keywords will include accessibility terms (eg, access* standard*, disabilit* policy, regulation*, and guidelines), parks and nature (eg, parks, outdoors, natural, urban, trail*, path*, and national), mobility device, and disability types (eg, wheelchair*, scooter*, blind, partial sight, deaf, hard of hearing, cognitive, mental, and developmental). The search will cover international (eg, United States, World, World Health Organization, Europe, France, Switzerland, United Kingdom, England, Australia, and Spain) and Canadian national standards, including provincial guidelines. For feasibility purposes, we will exclude the guidelines and standards from the municipal level. The search will be conducted both in English and French, and some standards in Spanish will also be included, as these 3 languages are spoken by the research team. The data will be extracted and charted based on the features listed on the Parks Canada website and completed with the content of the other standards found. This list of features included paths and trails (eg, sidewalks, walkways, stairs, ramps, lighting, and obstacles), parking and drop-off areas and transit areas, amenities (eg, rest areas, visitor centers, outdoor shelters, point-of-sales, and washrooms), wayfinding and signage, park management (eg, policies, practices, and communications), and summer and winter activities (eg, access to activities, equipment, and installations). This research will provide an overview and

critique of the existing standards on outdoor spaces, along with the possible knowledge gaps on the subject.

Phase 2: Park Audits and Mobile Interviews

In phase 2, physical audits of accessibility will be conducted on-site, and people with disabilities will walk or wheel along a portion of these routes. The focus of the audits and participant interviews will be on park trails and features along the trails.

Park Audits

Approximately 10 km of park trails in 3 parks in each province will be audited for their accessibility, including the trails participants will be used in the mobile interviews. This will be used to provide context for the analysis of the conditions faced by participants in the mobile interview. Trail slope, cross slope, width, surface quality (ie, firm, level, and stable), and presence of obstacles and hazards will be measured and mapped using the High Efficiency Trail Assessment Process (HETAP; Beneficial Designs) cart equipped with automated GPS, distance and slope sensors, and a camera. Parks will be chosen to include a variety of settings (ie, mountains, coastlines, and forests) and features (eg, beaches, picnic areas, and camping) found in most national parks. The closest national or subnational parks (ie, provincial and regional) that have a diversity of features will be used to minimize travel for participants. This process will allow us to objectively assess the parks for their specific characteristics and provide a portrait of possible obstacles users can face while visiting parks.

Mobile Interviews

Overview

Mobile interviews will be conducted in three steps: a preinterview survey, a mobile interview, and study-specific interview activities. These mobile interviews were performed in parks with users living with various disabilities to gather more information on the lived experience of individuals and how, according to their abilities, their park experience is influenced. Mobile interviews allow the identification of unforeseen sociospatial interactions compared with traditional face-to-face interviews [23] and can help people with disabilities think about elements that they would not think of if they were not directly in the environment [24]. Participatory research is a broad method that can be carried out in many ways, depending on the objectives. The mobile interview method used is one way to accomplish our objective. As specified at the beginning of the Methods section, we have advisory committees that include individuals with lived experience and disability organization leaders to guide us through the process. The mobile interview process described in step 2 specifies how we are going to involve individuals in the process.

Step 1 (Preinterview Survey)

Each participant will complete a web-based questionnaire (Qualtrics) that asks about sociodemographic characteristics (eg, age and sex), disability and mobility status (eg, diagnoses and assistive aids used), wayfinding skills, preferences for park settings and activities, and transport mode to parks. A few days before the interview, participants will be contacted to remind them of the interview, review survey responses, and review the assigned park website to evaluate whether or not it provides the information they would need to feel confident about visiting the park.

Step 2 (Mobile Interview)

Interviews will take place in the park assigned to the participant [25-27]. The interviews will be administered by trained researchers. They will assist participants where necessary and ensure health precautions are followed. The mobile interview will take approximately 2 hours along 3 predetermined routes of 500-1300 m in length [26]. Participants will be encouraged to take breaks as needed.

Before starting a route, participants will be given a map of the intended route to help orient them and asked what their expectations are for the route (eg, how far do they think the route is, how hard will it be, and do they think they will enjoy it?). Specific adaptations will be made to the map for individuals with vision loss (ie, describe the map). The map will include the trail that will be used, features along the route, and landmarks that might assist their travel. Researchers will retrieve the map but make it available during the journey when requested. This will be repeated before each route. While traveling through the park, researchers will use a GoPro 8 (GoPro Inc; with GPS device) to film the participants and an additional audio recorder to capture the discussion along the journey. Researchers will follow participants as they travel along the route, redirecting them to the prescribed route if necessary. During this process, researchers will ask structured and semistructured questions about their experiences and take notes of any additional observations they see (eg, the participant appears to be struggling with the terrain and the entire width of the path was covered in mud). The semistructured component of the interview guide (Textbox 1) will focus on participants' experiences related to wayfinding and wayfaring.

Textbox 1. Semistructured interview questions about wayfinding and wayfaring experiences.

Wayfinding questions

- What direction do you think we should go? Why do you think this?
- Can you show on the map where we are and where we are going?
- What cues are you using to make that decision and why?
- What is drawing your eye or attention?

Wayfaring questions

- Routes
  - How does this path feel to you?
  - How safe do you feel along this route? Why?
  - How comfortable is this path for you? Why do you feel this way?
  - How do you feel about the overall experience of this path?
  - Are there any changes you would like to see made to this path?
- Features (eg, bench, picnic table, and washroom)
  - What are your thoughts about using this feature?
  - How would you use this feature?
  - Are there any changes you would make to this feature?

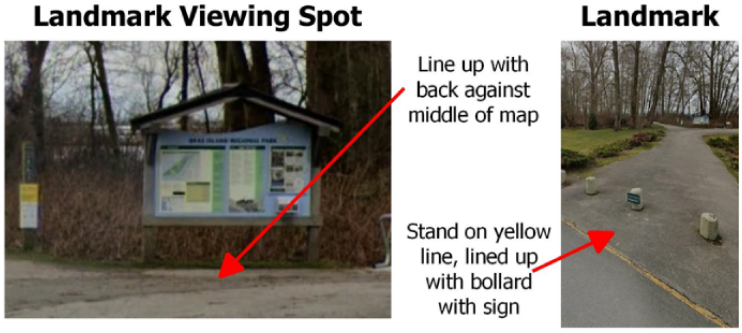
Structured questions about the presence or absence of features (eg, public toilet) and their characteristics (eg, accessibility and maintenance) will be asked as concise readable statements (eg, "The path leading to the public toilet is accessible"). This component draws on a park-specific adaptation of the Stakeholders Walkability/Wheelability Audit in Neighbourhood (SWAN) tool, a user-led microscale environmental audit tool that captures both objective and subjective data identifying features that hinder or support mobility and participation of people with mobility disabilities across five domains (ie, functionality, safety, land use features or destinations, maintenance or esthetics, and social aspects) [28]. Pilot testing of the original tool showed good interrater agreement across 90% of the items on the tool [29]. The park-specific adaptation is called the Stakeholders Walkability/Wheelability Audit in Nature (SWAN-Parks) tool. It includes blocks of questions associated with the aforementioned five domains that are tailored to different areas in the park, namely, amenities (eg, washrooms), paths (ie, that connect between different amenities), and trails. These blocks of questions were developed based on extant empirical evidence on accessibility issues in parks for different groups of people with disabilities, as well as pre-existing validated park audit tools, such as the Community Stakeholder Park Audit Tool [30] and Natural Environment Scoring Tool [31]. The semistructured and structured questions are sequenced into a cohesive interview guide and customized against the preidentified interview routes. Research assistants will orient participants (and personal assistants) to the different types of questions being asked during the mobile interview and provide prompts and clarification on an ongoing basis as they move through the mobile interview.

Sighted participants will also be asked to wear glasses that help track their gaze (Tobii Pro Glasses; 50-Hz eye-tracking glasses). These glasses will track their gaze during the journey to understand where they focus their attention during travel, including hazards, enjoyable features, and landmarks.

Step 3 (Postroute Interview Questions)

At the end of each route, participants will be asked to complete a series of tasks. First, participants will rate the route on a 7-point Likert scale, their perceived physical demand, mental demand, safety, enjoyment, and confidence to find their way independently. Follow-up questions for each scale would be regarding which changes they would recommend, including providing equipment or changes to the environment. Participants will also be asked to recall the route verbally or by drawing the route and all its features onto a route map [32]. The participant will be asked to describe the wayfaring and wayfinding experiences overall and provide additional feedback and recommendations. The final task will be to test objective spatial skills, including orientation and estimation skills necessary to the park environments [33]. Participants will be positioned at a predefined location and asked to point a compass in the direction of the origin of the route to measure orientation skills. At this same location, they will also be asked to estimate the distance and slope to a predefined landmark in the distance (Figure 2) [34,35]. Orientation and estimation skills are essential for reaching destinations and learning routes for future travel [36]. These skills may enhance confidence and encourage greater use of parks [37]. This information may also help parks to identify signage needs, including the type of information needed on those signs (eg, route slopes and distance markers). For winter evaluations, the data collection process will be repeated, changes in accessibility due to seasonality will be noted.

Figure 2. Test of objective spatial skills.



Recruitment

A purposive sample of 48 people (24 at each site) with a broad range of disabilities who use a variety of mobility devices will be eligible to participate in this study. To be included, participants will be at least aged 18 years, able to travel approximately 3 km with rests over a 2- to 3-hour period, and able to communicate directly with researchers (verbally) or indirectly through an assistant or attendant (Table 1 describes the sample distribution according to their individual characteristics).

Participants will be recruited through partners and participants from previous studies and selective advertising, if necessary. As there are 8 disability groups, 24 participants, and only 3 parks, participants will be assigned, where possible, to a park to ensure maximum variation in each park. Each participant will visit 1 of 3 park sites for a walking or wheeling interview during the summer months. Half of those participants (12/24, 50%) will visit the same park site in winter to account for the impacts of seasonal variation on potential accessibility standards. Participants in each province will complete 36 interviews, resulting in 72 separate interviews.

Table 1. Sample distribution according to participant characteristics (n=24).

Characteristics	Participants, n (%)
Scooter users	3 (13)
Power wheelchair users	3 (13)
Manual wheelchair users	3 (13)
Walker users	3 (13)
Cane or crutches users	3 (13)
People with visual impairments (including both white cane users and guide dog users)	3 (13)
People with hearing impairments	2 (8)
People with cognitive impairment (eg, dementia, autism, and intellectual disabilities)	4 (17)

Data Analysis

Participant Characteristics

Survey responses from step 1 and objective spatial skills tests from step 2 will be used to summarize the characteristics of the sample. Descriptive analysis will include counts of nominal data, means, and SDs for participant's socioeconomic status, mobility, assistive device use, and wayfinding skills. Results of the spatial skills tests will be reported as absolute error (pointing error in degrees, distance estimation in meters, and slope estimation in percent) and as relative error in percent.

Descriptive Analysis

The descriptive analysis will be carried out as follows:

- Concerning interview results, video and audio from the mobile interview will be transcribed to indicate what was being said or observed and by whom. Each quotation will be coded to reflect the feature or experience being explained (wayfinding or wayfaring) by the participant and any observation made by the researchers. The quotation and its code will be digitized in the geographic information system (GIS) at the location where it occurred. This will be linked to the participant survey responses through their ID as a separate file in the GIS (delimited file without spatial information). The results from the SWAN, the accessibility audit in phase 1, and sketch maps from the interview will also be uploaded into the GIS. They will be used to provide context about the physical environment and participants' familiarity and recommendations. These layers can then allow for filters to be used during the analysis based on any data collected during the survey. Wayfinding results include feedback from participants during the interview and researcher observations about wayfinding behaviors.
- A summary of the gaze position of participants will be completed by categorizing the amount of time looking at specific items and general gazing locations (eg, do manual wheelchair and walker users spend more time looking down at the ground compared with power wheelchair users or those who are able to see and walk?). Eye gaze data (analyzed with iMotions Eye Tracking Module software) will be added as a layer in the GIS that will be reviewed to

determine if patterns exist between layers (topographical, spatial transcript, sketch maps, video evidence, and researcher observations). These findings will be compared between mobility groups and between environments to determine if there are differences in how individuals with disabilities experience the environment and use landmarks to recall routes.

Global Analyses

The scoping review provides us with an additional layer of information that we can use to compare with the HETAP objective trail data and the subjective information from the mobile interviews. We can compare the subjective and objective layers using the existing standards from the scoping review.

Phase 3: Modified Delphi for the Prioritization and Identification of Standards

The Delphi panel process (participatory consensus developing process) [38,39] will be used to establish consensus and prioritize recommendations for accessibility standards using several national panels that focus on specific park areas (eg, trails and paths, information, and services). Drawing on what we will learn during the systematic review conducted in phase 1, and from the mobile interviews conducted in phase 2, we will then conduct additional individual interviews and focus groups of approximately 40 participants (including people with a wide spectrum of physical, cognitive, and sensory disabilities, as well as accessibility experts). Through a semistructured interview of approximately 1 hour, these experts will be asked to share their opinion on the hierarchy on which accessibility standards should be considered. The standards will be presented, and other participants' opinions will be shared, and the participants will have to comment.

During the second and third rounds of the Delphi, we will use a multimodal and iterative approach to encourage participants to reflect on existing standards and make recommendations for new standards. This process will include up to 100 participants (English- and French-speaking individuals from a range of ages, cultures, ethnicities, and income levels) that can participate with the assistance of a research assistant or independently on the web. Depending on the number of potential standards, these participants may be subdivided so they can review potential standards that are categorized into more manageable subgroups (eg, 3 or 4). From this process, we will create a list of recommended accessibility standards. The use of the Delphi process as it is presented considers the fact that we have 2 sites (French- and English-speaking) that we both want to consider rather than considering 2 separate sites. It also allows us to reach out to a greater number of experts in various fields.

Ethics

The protocol for this study was approved by the Research Ethics Boards at the University of British Columbia (H20-04036), the Centre intégré universitaire de santé et de services sociaux de la Capitale-Nationale (Project 2021-2120), and regional health authorities at each site. All study participants will provide informed consent. Evaluation in parks began in May 2021.

Results

Funding for this study was obtained from the Accessibility Standards, Canada. Results from the study will be reported for all phases including the systematic review, park audits, mobile interviews, and the Delphi process. A variety of spatial transcripts will be developed that show patterns in the data according to personal characteristics, age ranges, gender, or assistive device use or disability. A report of the final recommendations will conclude the results.

Discussion

Principal Findings

The purpose of this protocol is to describe the methodology for informing park accessibility standards. The combination of data gathered on parks and their use by individuals with various disabilities as well as the participatory approaches used to discuss existing standards should allow a better understanding of the conditions and dynamics required to propose positive inclusive leisure experiences in parks.

The data collected will inform future national park accessibility standards. Having a repository of existing standards (phase 1) and evidence supporting them will be beneficial to others who are undertaking the development of similar standards, to other levels of government in Canada, and internationally. It is anticipated that others will adopt this methodology to create user-driven accessibility standards that will have the following characteristics: (1) promote the widespread inclusion of people with disabilities in these spaces; (2) introduce broadly applicable standards (rather than siloed) and promote the widespread inclusion of people with disabilities in these spaces; (3) facilitate access by their families, parents with children, and older adults who may not self-identify as having a disability; (4) make park managers aware of the accessibility information that people with disabilities and their families need to plan trips to parks and enjoy their visits; (5) assist decision-makers in assessing to make improvements to parks and that people with disabilities and their families will use that information to make decisions about their visits; (6) continually be reviewed and assessed to ensure they meet the needs of people with disabilities as technologies and demand evolve; (7) create demand for similar mapping exercises to be undertaken in other parks; (8) identify modifications required for existing standards and new standards that should be developed; and (9) be adopted federally and that they will inform the development of similar standards provincially and municipally.

Limitations

The nature of the challenge addressed in this study introduced some limitations. The sample size was based on having adequate representation of all groups in different settings similar to federal parks but was fairly small. Given that the data will be collected outdoors, we were restricted by the nature of the exercise and requirements of the activity. Taking into consideration the grant funding and the fact that we wanted to offer a generous stipend to the participants, plus the cost of the equipment used, we were also limited in the number of participants we were able to reach.

Moreover, the COVID-19 pandemic forced us to increase researcher and participant safety measures by using parks that were not at the national level because they were closer to the participants' homes and required less travel. This may result in not assessing every feature or experience that might be expected in a national park. Limitations due to the size of our sample likely mean that not every potential perspective needed to inform standards could be accounted for in a population of such diversity of abilities, assistive device use, and park preferences. Finally, we may be limited in the causal conclusions that may be drawn because the study was not experimental in nature. This may be offset somewhat using multiple methods (eg, systematic review, mobile interview, and the Delphi method), the composition of the research team, and guidance from partners and the advisory committee.

Conclusions

This study will provide valuable insights from people with disabilities for recommending accessibility standards to be used in national parks and beyond. Although standards are not the only part of an effective inclusive park strategy, they are necessary for establishing a common language and set of expectations for accessibility in these spaces. They establish a baseline for park agencies to build on and ensure that the tremendous benefits they provide are available to all.

Acknowledgments

The authors would like to thank Accessibility Standards Canada for their generous funding and our partners and committee members for assisting in designing and recruiting participants for the study. FR is supported by a Quebec Health Research Funds Senior Salary Grant. KLB is supported by a Quebec Health Research Funds Junior 1 Salary Award. WBM was supported by a New Investigator Award from the Canadian Institutes of Health Research.

Conflicts of Interest

None declared.

References

1. Jakubec S, Carruthers Den Hoed D, Ray H, Krishnamurthy A. Mental well-being and quality-of-life benefits of inclusion in nature for adults with disabilities and their caregivers. Landscape Res 2016 Jul 07;41(6):616-627. [doi: 10.1080/01426397.2016.1197190]
2. Saitta M, Devan H, Boland P, Perry MA. Park-based physical activity interventions for persons with disabilities: a mixed-methods systematic review. Disabil Health J 2019 Jan;12(1):11-23. [doi: 10.1016/j.dhjo.2018.07.006] [Medline: 30100214]
3. Gascon M, Zijlema W, Vert C, White MP, Nieuwenhuijsen MJ. Outdoor blue spaces, human health and well-being: a systematic review of quantitative studies. Int J Hyg Environ Health 2017 Nov;220(8):1207-1221. [doi: 10.1016/j.ijheh.2017.08.004] [Medline: 28843736]
4. James L, Shing J, Mortenson WB, Mattie J, Borisoff J. Experiences with and perceptions of an adaptive hiking program. Disabil Rehabil 2018 Jun;40(13):1584-1590. [doi: 10.1080/09638288.2017.1302006] [Medline: 28325069]
5. Labbé D, Bahen M, Hanna C, Borisoff J, Mattie J, Mortenson WB. Setting the sails: stakeholders perceptions of an adapted sailing program. Leisure Sci 2019 Nov 11:1-15. [doi: 10.1080/01490400.2019.1686446]
6. Markevych I, Schoierer J, Hartig T, Chudnovsky A, Hystad P, Dzhambov AM, et al. Exploring pathways linking greenspace to health: theoretical and methodological guidance. Environ Res 2017 Oct;158:301-317. [doi: 10.1016/j.envres.2017.06.028] [Medline: 28672128]
7. Merrick D, Hillman K, Wilson A, Labbé D, Thompson A, Mortenson WB. All aboard: users' experiences of adapted paddling programs. Disabil Rehabil 2021 Oct;43(20):2945-2951. [doi: 10.1080/09638288.2020.1725153] [Medline: 32058823]
8. Rugel E. Green space and mental health: pathways, impacts, and gaps. National Collaborating Centre for Environmental Health. URL: https://ncceh.ca/sites/default/files/Full_Review-Greenspace_Mental_Health_Mar_2015.pdf [accessed 2022-02-02]
9. Rugel EJ, Carpiano RM, Henderson SB, Brauer M. Exposure to natural space, sense of community belonging, and adverse mental health outcomes across an urban region. Environ Res 2019 Apr;171:365-377. [doi: 10.1016/j.envres.2019.01.034] [Medline: 30716514]
10. Shanahan DF, Bush R, Gaston KJ, Lin BB, Dean J, Barber E, et al. Health benefits from nature experiences depend on dose. Sci Rep 2016 Jun 23;6:28551 [FREE Full text] [doi: 10.1038/srep28551] [Medline: 27334040]
11. Morris S, Fawcett G, Brisebois L, Hughes J. A demographic, employment and income profile of Canadians with disabilities aged 15 years and over, 2017. Canadian Survey on Disability Reports. URL: https://www150.statcan.gc.ca/n1/pub/89-654-x/89-654-x2018002-eng.htm [accessed 2022-02-02]
12. Okoro CA, Hollis ND, Cyrus AC, Griffin-Blake S. Prevalence of disabilities and health care access by disability status and type among adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Aug 17;67(32):882-887 [FREE Full text] [doi: 10.15585/mmwr.mm6732a3] [Medline: 30114005]

13. Burns N, Paterson K, Watson N. An inclusive outdoors? Disabled people's experiences of countryside leisure services. Leisure Studies 2009 Oct;28(4):403-417. [doi: 10.1080/02614360903071704]
14. Kavanagh AM, Krnjacki L, Aitken Z, LaMontagne AD, Beer A, Baker E, et al. Intersections between disability, type of impairment, gender and socio-economic disadvantage in a nationally representative sample of 33,101 working-aged Australians. Disabil Health J 2015 Apr;8(2):191-199. [doi: 10.1016/j.dhjo.2014.08.008] [Medline: 25278487]
15. Accessible Canada Act (S.C. 2019, c. 10). Justice Laws Website. URL: https://laws-lois.justice.gc.ca/eng/acts/A-0.6/ [accessed 2022-02-02]
16. Accessibility standards in outdoor environments. National Park Services, USA. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.626.4095&rep=rep1&type=pdf [accessed 2022-02-02]
17. Jang S, Mortenson W, Hurd L, Kirby R. Caught in-between: tensions experienced by community mobility scooter users. Disability Soc 2019 Dec 09;35(10):1577-1595. [doi: 10.1080/09687599.2019.1696749]
18. Hickman K. Disabled cyclists in England: imagery in policy and design. Proc Institution Civil Engineers Urban Design Planning 2016 Jun;169(3):129-137. [doi: 10.1680/udap.14.00048]
19. Ormerod M, Newton R, MacLennan H, Faruk M, Thies S, Kenney L, et al. Older people's experiences of using tactile paving. Proc Institution Civil Engineers Municipal Engineer 2015 Mar;168(1):3-10. [doi: 10.1680/muen.14.00016]
20. Lee H. The effects of truncated dome detectable warnings on travelers negotiating curb ramps in wheelchairs. J Visual Impairment Blindness 2019 Jan 07;105(5):276-286. [doi: 10.1177/0145482x1110500505]
21. Pour une expérience de loisir inclusive. Bulletin de l'Observatoire Québécois du Loisir. URL: https://oraprdnt.uqtr.uquebec.ca/pls/public/docs/FWG/GSC/Publication/170/377/1765/1/67752/5/F1367630963_Vol._12_No.11___Pour_une_exp_rience_de_loisir_inclusive.pdf [accessed 2022-02-02]
22. 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]
23. Carpiano RM. Come take a walk with me: the "go-along" interview as a novel method for studying the implications of place for health and well-being. Health Place 2009 Mar;15(1):263-272. [doi: 10.1016/j.healthplace.2008.05.003] [Medline: 18606557]
24. Hein J, Evans J, Jones P. Mobile methodologies: theory, technology and practice. Geography Compass 2008 Oct 24;2(5):1266-1285.
25. Parent L. The wheeling interview: mobile methods and disability. Mobilities 2016 Oct 31;11(4):521-532. [doi: 10.1080/17450101.2016.1211820]
26. Evans J, Jones P. The walking interview: methodology, mobility and place. Applied Geography 2011 Apr;31(2):849-858. [doi: 10.1016/j.apgeog.2010.09.005]
27. Wästerfors D. Required to be creative. Everyday ways for dealing with inaccessibility. Disability Soc 2020 Feb 04;36(2):265-285. [doi: 10.1080/09687599.2020.1720610]
28. Mahmood A, O'Dea E, Bigonnesse C, Labbe D, Mahal T, Qureshi M, et al. Stakeholders Walkability/Wheelability Audit in Neighbourhoods (SWAN): user-led audit and photographic documentation in Canada. Disability Soc 2019 Aug 26;35(6):902-925. [doi: 10.1080/09687599.2019.1649127]
29. Mahmood A, Chaudhury C, Oswald F, Konopik N. Seniors Walkability Audit in Neighbourhoods (SWAN): Development of a user-led observation tool to evaluate neighbourhood design features that fosters mobility in older adults. In: International Association People-Environment Studies. 2016 Presented at: IAPS 24; 2016; Lund, Sweden URL: http://www.iaps24.se/conference-information/scientific-program
30. Kaczynski AT, Stanis SA, Besenyi GM. Development and testing of a community stakeholder park audit tool. Am J Prev Med 2012 Mar;42(3):242-249. [doi: 10.1016/j.amepre.2011.10.018] [Medline: 22341161]
31. Gidlow C, van Kempen E, Smith G, Triguero-Mas M, Kruize H, Gražulevičienė R, et al. Development of the natural environment scoring tool (NEST). Urban Forestry Urban Greening 2018 Jan;29:322-333. [doi: 10.1016/j.ufug.2017.12.007]
32. Boschmann EE, Cubbon E. Sketch maps and qualitative GIS: using cartographies of individual spatial narratives in geographic research. Professional Geographer 2013 Apr 19;66(2):236-248. [doi: 10.1080/00330124.2013.781490]
33. Golledge R. Place recognition and wayfinding: Making sense of space. Geoforum 1992 May;23(2):199-214. [doi: 10.1016/0016-7185(92)90017-x]
34. Fajen B, Phillips F. Spatial perception and action. In: Handbook of Spatial Cognition. Washington, D.C., United States: American Psychological Association; 2013.
35. Hochmair H, Frank A. Influence of estimation errors on wayfinding-decisions in unknown street networks analyzing the least-angle strategy. Spat Cogn Comput Springer 2000;2(4):313 [FREE Full text]
36. Hund AM, Minarik JL. Getting from here to there: spatial anxiety, wayfinding strategies, direction type, and wayfinding efficiency. Spatial Cognit Comput 2006 Sep;6(3):179-201. [doi: 10.1207/s15427633scc0603_1]
37. Maus M, Lindeman D, Satariano W. Wayfinding, mobility, and technology for an aging society. In: Community Wayfinding: Pathways to Understanding. Cham: Springer International Publishing; 2016.

38. Mortenson WB, Miller WC, Boily J, Steele B, Crawford EM, Desharnais G. Overarching principles and salient findings for inclusion in guidelines for power mobility use within residential care facilities. J Rehabil Res Dev 2006;43(2):199-208 [FREE Full text] [doi: 10.1682/jrrd.2005.01.0031] [Medline: 16847786]
39. Rushton PW, Miller WC, Lee Kirby RL, Eng JJ, Yip J. Development and content validation of the Wheelchair Use Confidence Scale: a mixed-methods study. Disabil Rehabil Assist Technol 2011;6(1):57-66 [FREE Full text] [doi: 10.3109/17483107.2010.512970] [Medline: 20809874]

Abbreviations

GIS: geographic information system
HETAP: High Efficiency Trail Assessment Process
SWAN: Stakeholders Walkability/Wheelability Audit in Neighbourhood

Edited by G Eysenbach; submitted 07.10.21; peer-reviewed by H Carbonneau, É Robitaille; comments to author 08.11.21; revised version received 24.11.21; accepted 27.11.21; published 31.03.22.

Please cite as:
Prescott M, Routhier F, Labbé D, Grandisson M, Mahmood A, Morales E, Best KL, Mostafavi MA, Borisoff J, Gamache S, Sawatzky B, Miller WC, Bulk LY, Robillard JM, Jenkins HT, Seetharaman K, Mortenson WB
Providing Accessible Recreation Outdoors—User-Driven Research on Standards (PARCOURS): Protocol for a Multiphase Study
JMIR Res Protoc 2022;11(3):e33611
URL: https://www.researchprotocols.org/2022/3/e33611
doi:10.2196/33611
PMID:35357324

Corrigenda and Addenda

Correction: Assessing the Efficacy of an Individualized Psychological Flexibility Skills Training Intervention App for Medical Student Burnout and Well-being: Protocol for a Randomized Controlled Trial

Elizabeth Ditton[1,2,3], BA (Psych) Hons, MA Psych (Clin); Brendon Knott[4], BSc (Psych) Hons, MA Psych (Clin); Nicolette Hodyl[1,2], BSc Hons, GradCert Biostats, PhD; Graeme Horton[3], BS, MB, MEnvStud, PhD; Frederick Rohan Walker[1,2,3], BA (Psych) Hons, PhD; Michael Nilsson[1,2,3,5], MD, PhD

[1]Centre for Rehab Innovations, University of Newcastle, Callaghan, Australia

[2]Hunter Medical Research Institute, New Lambton Heights, Australia

[3]College of Health, Medicine and Wellbeing, University of Newcastle, Callaghan, Australia

[4]Contextual Interventions, Newcastle, Australia

[5]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

Corresponding Author:
Elizabeth Ditton, BA (Psych) Hons, MA Psych (Clin)
Centre for Rehab Innovations
University of Newcastle
University Drive
Callaghan, 2308
Australia
Phone: 61 2 404 ext 20738
Email: elizabeth.ditton@newcastle.edu.au

Related Article:

Correction of: https://www.researchprotocols.org/2022/2/e32992

(JMIR Res Protoc 2022;11(3):e37798) doi:10.2196/37798

In "Assessing the Efficacy of an Individualized Psychological Flexibility Skills Training Intervention App for Medical Student Burnout and Well-being: Protocol for a Randomized Controlled Trial" (JMIR Res Protoc 2022;11(2):e32992) the authors noted two errors.

First, in the originally published article a footnote appeared below Table 1 as follows:

> *[a]N/A: not applicable.*

This was deleted in the correction as it did not correspond to any values in Table 1.

Second, in the originally published article, the duration of the trial intervention was reported as 5 weeks. The actual duration of the intervention was 8 weeks.

The value appeared incorrectly in the following 7 instances:

1. Abstract; Methods

> *Participants in the individualized and nonindividualized intervention arms will have 5 weeks to access the app, which includes a PF concepts training session (stage 1) and access to short PF skill activities on demand (stage 2).*

This has been corrected as follows:

> *Participants in the individualized and nonindividualized intervention arms will have 8 weeks to access the app, which includes a PF concepts training session (stage 1) and access to short PF skill activities on demand (stage 2).*

2. Methods; Data Collection Tools and Procedures

> *Data will be collected at two time points: T1 (baseline) and T2 following the completion of the app-based intervention, commencing 5 weeks after baseline*

This has been corrected as follows:

> *Data will be collected at two time points: T1 (baseline) and T2 (following the completion of the app-based intervention, commencing 8 weeks after baseline).*

3. Methods; Intervention Stages

> *Participants who are allocated to the individualized and nonindividualized groups will have access to the 2-stage app for 5 weeks.*

This has been corrected as follows:

> *Participants who are allocated to the individualized and nonindividualized groups will have access to the 2-stage app for 8 weeks.*

4. Methods section; Intervention Stages

> *Participants may complete as many activities as they choose, but will be asked to complete at least four stage 2 skill activities during their 5-week period of access to the app.*

This has been corrected as follows:

> *Participants may complete as many activities as they choose, but will be asked to complete at least four stage 2 skill activities during their 8-week period of access to the app.*

5. Methods; Study Design; Figure 1

> *Access stage 2: individualized program (5-week access period)*

This has been corrected as follows:

> *Access stage 2: individualized program (8-week access period)*

6. Methods; Study Design; Figure 1

> *Access stage 2: nonindividualized program (5-week access period)*

This has been corrected as follows:

> *Access stage 2: nonindividualized program (8-week access period)*

7. Methods; Study Design; Figure 1

> *Complete t_2 assessments (t=5 weeks)*

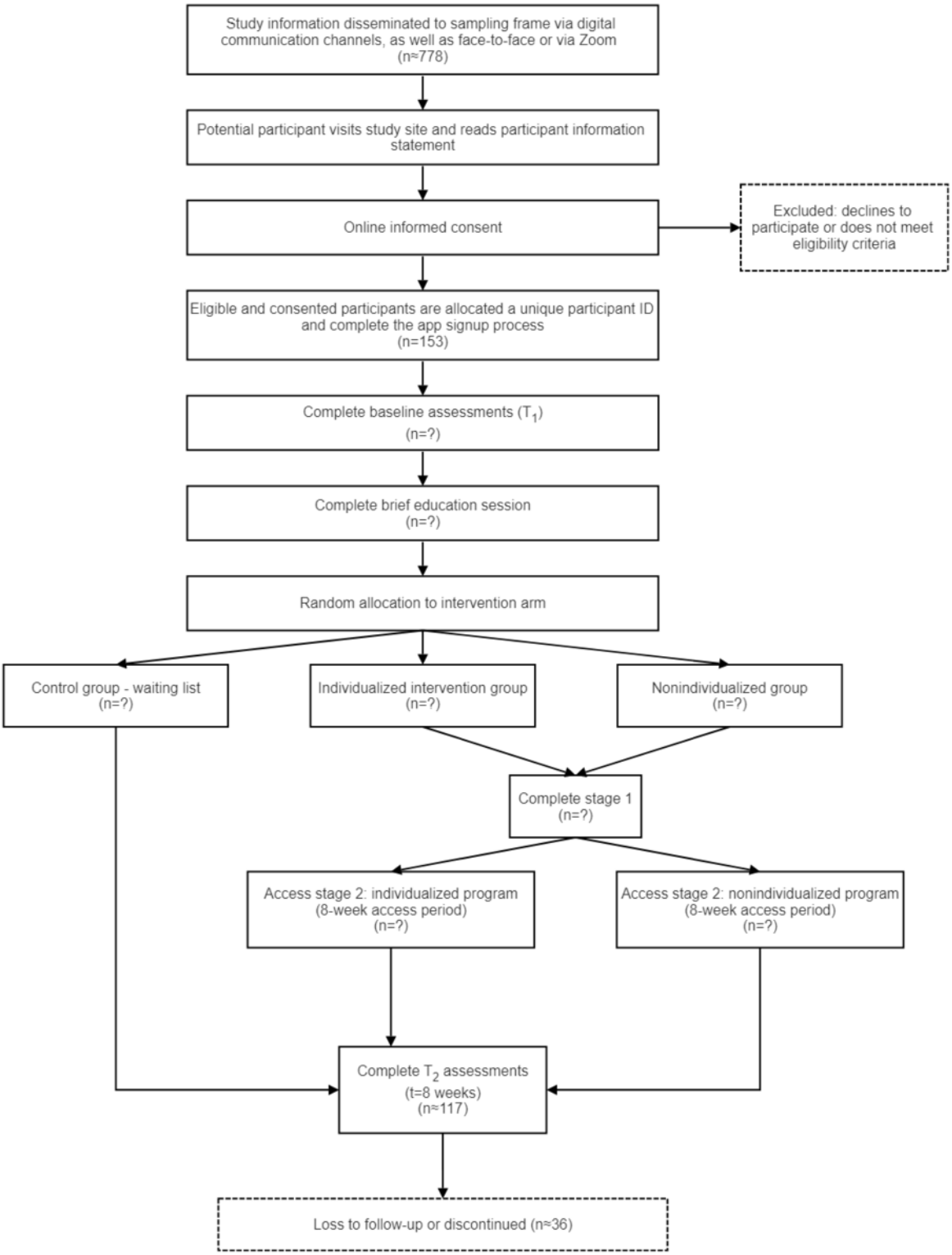
This has been corrected as follows:

> *Complete t_2 assessments (t=8 weeks)*

The corrected version of Figure 1 is included below. The originally published version of Figure 1 is attached as Multimedia Appendix 1.

The correction will appear in the online version of the paper on the JMIR Publications website on March 16, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Figure 1. Participant timeline.



Multimedia Appendix 1
Original Figure 1.
[PNG File , 78 KB - resprot_v11i3e37798_app1.png]

Submitted 07.03.22; this is a non–peer-reviewed article;accepted 08.03.22; published 16.03.22.

Please cite as:
Ditton E, Knott B, Hodyl N, Horton G, Walker FR, Nilsson M
Correction: Assessing the Efficacy of an Individualized Psychological Flexibility Skills Training Intervention App for Medical Student Burnout and Well-being: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2022;11(3):e37798
URL: https://www.researchprotocols.org/2022/3/e37798
doi:10.2196/37798
PMID:35294370

Corrigenda and Addenda

Correction: Rehabilitation Needs, Service Provision, and Costs in the First Year Following Traumatic Injuries: Protocol for a Prospective Cohort Study

Helene Lundgaard Soberg[1,2*], PhD; Håkon Øgreid Moksnes[1,3*], MD; Audny Anke[3,4,5], MD, PhD; Olav Røise[6,7], MD, PhD; Cecilie Røe[1,7], MD, PhD; Eline Aas[8], PhD; Unni Sveen[1,2], PhD; Christine Gaarder[7,9], MD, PhD; Pål Aksel Næss[7,9], MD, PhD; Eirik Helseth[7,10], MD, PhD; Hilde Margrete Dahl[7,11], MD, PhD; Frank Becker[7,12], MD, PhD; Marianne Løvstad[12,13], PhD; Kristian Bartnes[5,14], MD, PhD; Christoph Schäfer[1,4,5], MD; Mari S Rasmusssen[1,3], MSc; Paul Perrin[15], PhD; Juan Lu[16], MD, PhD; Torgeir Hellstrøm[1], MD, PhD; Nada Andelic[1,3*], MD, PhD

[1]Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Oslo, Norway

[2]Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway

[3]Institute of Health and Society, Research Centre for Habilitation and Rehabilitation Models & Services (CHARM), Faculty of Medicine, University of Oslo, Oslo, Norway

[4]Department of Rehabilitation, University Hospital of North Norway, Tromsø, Norway

[5]Faculty of Health Sciences, Department of Clinical Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway

[6]Division of Orthopaedic Surgery, Oslo University Hospital, Oslo, Norway

[7]Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

[8]Department of Health Management and Health Economics, Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway

[9]Department of Traumatology, Oslo University Hospital, Oslo, Norway

[10]Department of Neurosurgery, Oslo University Hospital, Oslo, Norway

[11]Department of Child Neurology, Oslo University Hospital, Oslo, Norway

[12]Sunnaas Rehabilitation Hospital, Nesoddtangen, Norway

[13]Department of Psychology, Faculty of Social Sciences, University of Oslo, Oslo, Norway

[14]Department of Cardiothoracic and Vascular Surgery, University Hospital of North Norway, Tromsø, Norway

[15]Departments of Psychology and Physical Medicine and Rehabilitation, Virginia Commonwealth University, Richmond, VA, United States

[16]Division of Epidemiology, Department of Family Medicine and Population Health, Virginia Commonwealth University, Richmond, VA, United States

[*]these authors contributed equally

Corresponding Author:
Helene Lundgaard Soberg, PhD
Department of Physical Medicine and Rehabilitation
Oslo University Hospital
PO Box 4956 Nydalen
Oslo, N-0424
Norway
Phone: 47 951009637
Email: h.l.soberg@medisin.uio.no

Related Article:

Correction of: https://www.researchprotocols.org/2021/4/e25980/

(JMIR Res Protoc 2022;11(3):e37723) doi:10.2196/37723

In “Rehabilitation Needs, Service Provision, and Costs in the First Year Following Traumatic Injuries: Protocol for a Prospective Cohort Study” (JMIR Res Protoc 2021;10(4):e25980), several errors were noted in author affiliations.

1. In the originally published paper, author affiliations 1 ("Department of Physical Medicine and Rehabilitation, Oslo University Hospital HF, Oslo, Norway") and 3 (“Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Oslo, Norway”) were the same. Therefore, “Oslo University Hospital HF” has been corrected to “Oslo University Hospital” in this affiliation for the following authors: Helene

Lundgaard Soberg, Mari S Rasmusssen, Torgeir Hellstrøm, Nada Andelic. The remaining affiliations have been renumbered as necessary.

2. For Audny Anke, the originally published author affiliation was:

> *Department of Cardiothoracic and Vascular Surgery, University Hospital of North Norway, Tromsø, Norway*

This has been corrected to the following:

> *Faculty of Health Sciences, Department of Clinical Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway*

The following new affiliation has also been added for this author:

> *Institute of Health and Society, Research Centre for Habilitation and Rehabilitation Models & Services (CHARM), Faculty of Medicine, University*

3. For Eline Aas, the originally published author affiliation was:

> *Department of Health Management and Health Economics, Institute for Health and Society, University of Oslo, Oslo, Norway*

This has been corrected to:

> *Department of Health Management and Health Economics, Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway*

4. For Marianne Løvstad, the originally published author affiliation was:

> *Department of Psychology, University of Oslo, Oslo, Norway*

This has been corrected to the following:

> *Department of Psychology, Faculty of Social Sciences, University of Oslo, Oslo, Norway*

5. For Kristian Bartnes, the originally published author affiliation was:

> *Institute of Clinical Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway*

This has been corrected to the following:

> *Faculty of Health Sciences, Department of Clinical Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway*

6. For Christoph Schäfer, the originally published author affiliation was:

> *Department of Cardiothoracic and Vascular Surgery, University Hospital of North Norway, Tromsø, Norway*

This has been corrected to the following:

> *Faculty of Health Sciences, Department of Clinical Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway*

The following new affiliation has also been added for this author:

> *Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Oslo, Norway*

In the originally published paper, the complete listing of authors and affiliations was as follows:

Helene Lundgaard Soberg [1,2], PhD; Håkon Øgreid Moksnes [3,4*], MD; Audny Anke [5,6], MD, PhD; Olav Røise [7,8], MD, PhD; Cecilie Røe [3,7], MD, PhD; Eline Aas [9], PhD; Unni Sveen [2,3], PhD; Christine Gaarder [7,10], MD, PhD; Pål Aksel Næss [7,10], MD, PhD; Eirik Helseth [7,11], MD, PhD; Hilde Margrete Dahl [7,12], MD, PhD; Frank Becker [7,13], MD, PhD; Marianne Løvstad [13,14], PhD; Kristian Bartnes [6,15], MD, PhD; Christoph Schäfer [5,6], MD; Mari S Rasmussen [1,4], MSc; Paul Perrin [16], PhD; Juan Lu [17], MD, PhD; Torgeir Hellstrøm [1], MD, PhD; Nada Andelic [1,4*], MD, PhD*

[1]*Department of Physical Medicine and Rehabilitation, Oslo University Hospital HF, Oslo, Norway*

[2]*Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway*

[3]*Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Oslo, Norway*

[4]*Institute of Health and Society, Research Centre for Habilitation and Rehabilitation Models & Services (CHARM), Faculty of Medicine, University of Oslo, Oslo, Norway*

[5]*Department of Rehabilitation, University Hospital of North Norway, Tromsø, Norway*

[6]*Department of Cardiothoracic and Vascular Surgery, University Hospital of North Norway, Tromsø, Norway*

[7]*Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway*

[8]*Division of Orthopaedic Surgery, Oslo University Hospital, Oslo, Norway*

[9]*Department of Health Management and Health Economics, Institute for Health and Society, University of Oslo, Oslo, Norway*

[10]*Department of Traumatology, Oslo University Hospital, Oslo, Norway*

[11]*Department of Neurosurgery, Oslo University Hospital, Oslo, Norway*

[12]*Department of Child Neurology, Oslo University Hospital, Oslo, Norway*

[13]*Sunnaas Rehabilitation Hospital, Nesoddtangen, Norway*

[14]*Department of Psychology, University of Oslo, Oslo, Norway*

[15]*Institute of Clinical Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway*

[16]*Departments of Psychology and Physical Medicine and Rehabilitation, Virginia Commonwealth University, Richmond, VA, United States*

[17]*Division of Epidemiology, Department of Family Medicine and Population Health, Virginia Commonwealth University, Richmond, VA, United States*

**these authors contributed equally*

Corresponding Author:

Helene Lundgaard Soberg, PhD

Department of Physical Medicine and Rehabilitation

Oslo University Hospital HF

PO Box 4956 Nydalen

Oslo, N-0424

Norway

Phone: 47 951009637

Email: h.l.soberg@medisin.uio.no

The complete corrected list of authors and affiliations is as follows:

Helene Lundgaard Soberg [1,2], PhD; Håkon Øgreid Moksnes [1,3*], MD; Audny Anke [3,4,5], MD, PhD; Olav Røise [6,7], MD, PhD; Cecilie Røe [1,7], MD, PhD; Eline Aas [8], PhD; Unni Sveen [1,2], PhD; Christine Gaarder [7,9], MD, PhD; Pål Aksel Næss [7,9],MD, PhD; Eirik Helseth [7,10] MD, PhD; Hilde Margrete Dahl [7,11], MD, PhD; Frank Becker [7.12], MD, PhD; Marianne Løvstad [12,13], PhD; Kristian Bartnes [5,14], MD, PhD; Christoph Schäfer [1,4,5,] MD; Mari S Rasmussen [1,3], MSc; Paul Perrin [15], PhD; Juan Lu [16], MD, PhD; Torgeir Hellstrøm [1], MD, PhD; Nada Andelic [1,3*], MD, PhD*

[1]*Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Oslo, Norway*

[2]*Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway*

[3]*Institute of Health and Society, Research Centre for Habilitation and Rehabilitation Models & Services (CHARM), Faculty of Medicine, University of Oslo, Oslo, Norway*

[4]*Department of Rehabilitation, University Hospital of North Norway, Tromsø, Norway*

[5]*Faculty of Health Sciences, Department of Clinical Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway*

[6]*Division of Orthopaedic Surgery, Oslo University Hospital, Oslo, Norway*

[7]*Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway*

[8]*Department of Health Management and Health Economics, Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway*

[9]*Department of Traumatology, Oslo University Hospital, Oslo, Norway*

[10]*Department of Neurosurgery, Oslo University Hospital, Oslo, Norway*

[11]*Department of Child Neurology, Oslo University Hospital, Oslo, Norway*

[12]*Sunnaas Rehabilitation Hospital, Nesoddtangen, Norway*

[13]*Department of Psychology, Faculty of Social Sciences, University of Oslo, Oslo, Norway*

[14]*Department of Cardiothoracic and Vascular Surgery, University Hospital of North Norway, Tromsø, Norway*

[15]*Departments of Psychology and Physical Medicine and Rehabilitation, Virginia Commonwealth University, Richmond, VA, United States*

[16]*Division of Epidemiology, Department of Family Medicine and Population Health, Virginia Commonwealth University, Richmond, VA, United States*

**these authors contributed equally*

Corresponding Author:

Helene Lundgaard Soberg, PhD

Department of Physical Medicine and Rehabilitation

Oslo University Hospital

PO Box 4956 Nydalen

Oslo, N-0424

Norway

Phone: 47 951009637

Email: h.l.soberg@medisin.uio.no

The correction will appear in the online version of the paper on the JMIR Publications website on March 23, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 04.03.22; this is a non–peer-reviewed article;accepted 04.03.22; published 23.03.22.

Please cite as:
Soberg HL, Moksnes HØ, Anke A, Røise O, Røe C, Aas E, Sveen U, Gaarder C, Næss PA, Helseth E, Dahl HM, Becker F, Løvstad M, Bartnes K, Schäfer C, Rasmusssen MS, Perrin P, Lu J, Hellstrøm T, Andelic N
Correction: Rehabilitation Needs, Service Provision, and Costs in the First Year Following Traumatic Injuries: Protocol for a Prospective Cohort Study
JMIR Res Protoc 2022;11(3):e37723
URL: https://www.researchprotocols.org/2022/3/e37723
doi:10.2196/37723
PMID:35320112

Protocol

Dashboards in Health Care Settings: Protocol for a Scoping Review

Danielle Helminski[1*], MPH; Jacob E Kurlander[1,2,3*], MD, MS; Anjana Deep Renji[4], MBA; Jeremy B Sussman[1,2,3], MD, MS; Paul N Pfeiffer[2,3,5], MD, MS; Marisa L Conte[4,6], MLIS; Oliver J Gadabu[4], MPH, PhD; Alex N Kokaly[7], MD, MHSA; Rebecca Goldberg[8], BS; Allison Ranusch[3], MA; Laura J Damschroder[3], MPH, MS; Zach Landis-Lewis[4], MLIS, PhD

[1]Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States
[2]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States
[3]Veterans Affairs Ann Arbor Center for Clinical Management Research, Ann Arbor, MI, United States
[4]Department of Learning Health Sciences, University of Michigan, Ann Arbor, MI, United States
[5]Department of Psychiatry, University of Michigan, Ann Arbor, MI, United States
[6]Taubman Health Sciences Library, University of Michigan, Ann Arbor, MI, United States
[7]Department of Medicine, UCLA Health, Los Angeles, CA, United States
[8]University of Michigan Medical School, Ann Arbor, MI, United States
[*]these authors contributed equally

Corresponding Author:
Danielle Helminski, MPH
Department of Internal Medicine
University of Michigan
NCRC Building 14
2800 Plymouth Road
Ann Arbor, MI, 48109
United States
Phone: 1 7346153952
Email: dhelmins@umich.edu

Abstract

Background: Health care organizations increasingly depend on business intelligence tools, including “dashboards,” to capture, analyze, and present data on performance metrics. Ideally, dashboards allow users to quickly visualize actionable data to inform and optimize clinical and organizational performance. In reality, dashboards are typically embedded in complex health care organizations with massive data streams and end users with distinct needs. Thus, designing effective dashboards is a challenging task and theoretical underpinnings of health care dashboards are poorly characterized; even the concept of the dashboard remains ill-defined. Researchers, informaticists, clinical managers, and health care administrators will benefit from a clearer understanding of how dashboards have been developed, implemented, and evaluated, and how the design, end user, and context influence their uptake and effectiveness.

Objective: This scoping review first aims to survey the vast published literature of “dashboards” to describe where, why, and for whom they are used in health care settings, as well as how they are developed, implemented, and evaluated. Further, we will examine how dashboard design and content is informed by intended purpose and end users.

Methods: In July 2020, we searched MEDLINE, Embase, Web of Science, and the Cochrane Library for peer-reviewed literature using a targeted strategy developed with a research librarian and retrieved 5188 results. Following deduplication, 3306 studies were screened in duplicate for title and abstract. Any abstracts mentioning a health care dashboard were retrieved in full text and are undergoing duplicate review for eligibility. Articles will be included for data extraction and analysis if they describe the development, implementation, or evaluation of a dashboard that was successfully used in routine workflow. Articles will be excluded if they were published before 2015, the full text is unavailable, they are in a non-English language, or they describe dashboards used for public health tracking, in settings where direct patient care is not provided, or in undergraduate medical education. Any discrepancies in eligibility determination will be adjudicated by a third reviewer. We chose to focus on articles

published after 2015 and those that describe dashboards that were successfully used in routine practice to identify the most recent and relevant literature to support future dashboard development in the rapidly evolving field of health care informatics.

Results: All articles have undergone dual review for title and abstract, with a total of 2019 articles mentioning use of a health care dashboard retrieved in full text for further review. We are currently reviewing all full-text articles in duplicate. We aim to publish findings by mid-2022. Findings will be reported following guidance from the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.

Conclusions: This scoping review will provide stakeholders with an overview of existing dashboard tools, highlighting the ways in which dashboards have been developed, implemented, and evaluated in different settings and for different end user groups, and identify potential research gaps. Findings will guide efforts to design and use dashboards in the health care sector more effectively.

International Registered Report Identifier (IRRID): DERR1-10.2196/34894

(JMIR Res Protoc 2022;11(3):e34894) doi:10.2196/34894

KEYWORDS

dashboard; mHealth; medical informatics; quality improvement; scoping review; health care; Cochrane library; Cochrane; stakeholder; health care sector; digital health; design; end user; development; implementation; evaluation; user need

Introduction

Background

Effectively measuring, monitoring, and responding to metrics about health-related decisions, practices, and outcomes has become an essential business function for modern health care organizations. Nimble health care organizations employ data for all manners of daily operational decision-making, ranging from supply chain management and staff scheduling to individual treatment planning and population health management [1]. For certain key performance metrics, payers have linked reimbursement to value-based payment programs [2] and accrediting bodies have required monitoring and disclosure of performance for accreditation or certification [3], incentivizing organizations to effectively monitor and track their performance against established benchmarks [4]. With the rapid proliferation of electronic health records, there is an abundance of patient- and provider-level data to use for assessing performance [5-7]. At the same time, vast data alone are of little use without systems to derive timely and actionable insights.

Health systems have increasingly adopted business intelligence software to track performance metrics in an automated way [8]. These applications have been defined by Loewen and Roudsari [9] as "specialized tools to collect, analyze, and present organizational data to operational leaders in user-friendly format(s) to support organizational objectives." One such tool that has seen considerable expansion in health care settings is the "dashboard," a business intelligence tool that uses data visualization to provide actionable feedback to improve performance, adherence to evidence-based practices, workflow management, and resource utilization [10,11]. Dashboards often display performance trends, peer comparisons, benchmarks, or goals, and use visual elements such as graphs and color-coding to improve interpretability [12].

To create an effective dashboard, developers must make multiple complex decisions. End users' information needs are highly contextual and depend on the clinical setting, professional roles, and the patient population, which impact selection of appropriate data elements, visualizations, and interactivity [13-15]. Although health care executives may prefer to see graphic performance trends over weeks or months, clinicians working with vulnerable patient groups may require real-time, patient-level health data so they can intervene quickly if needed. Indeed, numerous techniques for developing dashboards and selecting key metrics have been described, including focus groups, iterative usability testing, and a Delphi method [16,17]. More sophisticated dashboards also incorporate forecasting and decision support, which carry their own challenges [18,19]. The range of and most common strategies used to address these essential steps in dashboard development are unknown.

Developing effective dashboards tailored to the needs of the intended end user is only the first step in the health care performance improvement cycle. Developers and organizational leadership must also employ implementation strategies to promote uptake and use of the dashboard, such as the identification and involvement of "champions," ongoing training of end users, and changes in policy that mandate or incentivize dashboard use [20]. As development and maintenance of data-rich business intelligence tools, like dashboards, can be time- and resource-intensive, it is essential that these tools both function effectively and result in measurable improvements. Iterative evaluation of dashboard performance throughout development and implementation and beyond are critical to identify user- and system-level barriers to use as well as potential errors that may only be identified after extended use.

In this scoping review, we will survey peer-reviewed literature to describe the contexts in which dashboards have been used in health care settings, as well as how they were developed, implemented, and evaluated.

Aims and Comparison With Prior Work

This scoping review will provide a narrative overview of design elements and characteristics of health care dashboards, including where they exist geographically, the intended end users, information presented, whether/how the end user and setting impact dashboard design, and the processes used for development, implementation, and evaluation. Although previous reviews on health care dashboards have focused on identifying important design features and effectiveness of

dashboards in improving patient outcomes and clinician satisfaction [11,14,15,21,22], an updated review of *how* dashboard tools are used, and *by whom* will provide meaningful insight into how intended end user and setting impact the design, development and implementation of the dashboard (ie, the relationship of form and function). This information is essential to provide insights into (1) *how* and *why* dashboards work in different settings for different users, to allow relevant stakeholders to make more informed decisions about where to implement, and (2) how to effectively design dashboards based on their intended purpose and target audience. Given the rapidly evolving field of health informatics, the scoping review will also provide insight into the latest trends in dashboards, from initial conception and development through implementation and evaluation. Previous reviews of dashboards have included articles published only as recently as 2017 [11,14,21-23].

Methods

Study Design

The aims of this study can be best accomplished through a scoping review, which differs from a systematic review in that scoping reviews generally have a broader scope and are exploratory, not requiring critical quantitative appraisal of synthesized findings [24,25]. For this study, we will follow the framework for conducting scoping reviews developed by Arksey & O'Malley [26] and further refined by Levac et al [27]. A description of each step is provided below.

Step 1. Identifying the Research Questions

The key research questions, which were established through a process of team discussions and preliminary searches of the literature on health care dashboards, are as follows:

1. What design features are most frequently incorporated in health care dashboards?
2. For what purposes are dashboards developed in health care settings?
3. Where, and by whom, are dashboards used?
4. What processes and/or frameworks are used for development, implementation, and evaluation of dashboards?

Step 2. Identifying Relevant Studies

We searched MEDLINE, Embase, Web of Science, and the Cochrane Library databases in July 2020 for relevant articles using comprehensive search strategies for each database that were developed in collaboration with a research librarian (MLC) and are available in Multimedia Appendix 1. These databases were selected since they represent a broad sample of literature relevant to the health sciences. Search terms included a variety of keywords and medical subject headings (MeSH) related to clinical health care and information technology. Search strategies were developed around the following key terms: "dashboard," "information technology," "healthcare," "electronic health record," "electronic medical record," "quality," "safety," "key performance indicators," "decision making," "decision support," "benchmark," and "informatics." Boolean operators "AND," "OR," and "NOT" were used to construct each search, with "NOT" operators used to reduce the number of results related to automotive and learning analytics dashboards. No date, language, or other restrictions were imposed in the database searches. Grey literature sources were not searched.

Step 3. Study Selection

All articles retrieved by the search were imported into and initially reviewed using Covidence [28], a screening and data extraction tool adopted by Cochrane in 2015 as the standard platform for producing Cochrane Reviews. In addition, 2 of 4 authors (DH, ADR, MLC, OJG) independently screened all titles and abstracts to identify potentially eligible studies. All articles that mentioned use of a "dashboard" in a health care setting were reviewed in full text and are currently undergoing duplicate review by 2 of 7 authors (DH, ADR, MLC, OJG, ANK, RG, AR) to determine eligibility, applying the inclusion and exclusion criteria listed in Textbox 1. We excluded articles published prior to 2015 and those describing dashboards that were not successfully used in routine workflow or were only used in a pretesting environment. We believe these exclusions are justified as rapid advancements in technology warrant a focus on newer research that is more likely to be reproducible. Additionally, limiting our analysis to dashboards that were successfully implemented or used outside of a pretesting environment provides a clearer view of existing barriers and facilitators to designing and implementing dashboards in real-world practice. Any disagreements that arise during full-text screening will be resolved through adjudication by a third author. For any studies that are reviewed in full text but not deemed eligible for inclusion in the scoping review, a reason for exclusion will be documented and provided with the results of the scoping review in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

Textbox 1. Eligibility criteria for full-text review.

Inclusion criteria

- Peer-reviewed articles that describe the development, implementation, and/or evaluation of a dashboard used in a health care setting outside of a pretesting environment. Health care settings include clinics, hospitals, health systems, or any other settings where medical care is provided. Both quantitative and qualitative evaluations of dashboards will be included.

Exclusion criteria

- Non–English language publication
- Articles published prior to 2015
- Articles that describe pretesting of pilot or prototype dashboards that were not successfully implemented or used outside of a testing environment
- Articles that describe public health dashboards used for geographic tracking of disease or comparing city- or country-level data not used for clinical or management-level decision-making in a health care setting where patient care is provided
- Articles that describe dashboards used in undergraduate medical education, or in educational contexts where there is no direct association with patients, patient care, or facility management
- Articles for which the full-text manuscript is unavailable

Step 4. Charting the Data

A preliminary list of data elements for charting is presented in Textbox 2. However, in accordance with recommendations from Levac et al [27], an iterative process will be used to identify additional elements for data extraction and analysis as the study progresses. Using an iterative process improves the quality of the review by allowing reviewers to gain familiarity with included studies and add or revise data extraction elements accordingly. A standardized data extraction form will be developed and reviewed by all authors. The form will be pilot tested by two authors who will independently complete data extraction for a subset of articles to ensure consistency among extractors. Once a high level of agreement is achieved between extractors, the pilot extraction form will be approved, and two authors will independently extract data from each included study. Any disagreements in data extraction will be resolved by discussion between the two authors; if the reviewers are unable to reach consensus, a third author will serve as arbiter.

Textbox 2. Preliminary list of data extraction elements.

Article information

- Title
- Author
- Publication year
- Journal
- Study type

Contextual factors

- Geographic location of the described dashboard
- Health care setting
- Intended end user(s)

Primary purpose or goal of the dashboard

- Reason stated for development or use of dashboard

Development

- Software used
- Framework(s) used to guide development or pretesting
- Usability testing conducted
- Involvement of users in development process

Implementation

- Adjunct strategies used in conjunction with dashboard (such as academic detailing, audit and feedback, or financial incentivization)
- Identification of potential barriers and facilitators to use of dashboard prior to implementation
- Identification and involvement of champions
- Training of stakeholders or distribution of educational materials on how to use the dashboard
- Protocol or policy changes that mandate use of the dashboard

Evaluation

- Type of evaluation (qualitative or quantitative)

Design features

- Format (including delivery channel and timing)
  - Frequency of data updates
  - Use of visual elements
  - Delivery channel (eg, website, email, wall display)
- Information content
  - Descriptions of performance summary data (including indicators, time intervals, comparators, and their performance levels)
  - Patient lists (typically patients who have actionable data, such as guideline-discordant care; “yes” or “no”)
  - Patient-level data (“yes” or “no”)
  - Recommended actions (“yes” or “no”)
  - Metrics or evaluation based on benchmarks established by accrediting bodies, health care payer organizations, or national guidelines (“yes” or “no”)
- Functionality (“yes” or “no”)
  - Multilevel presentation of data
  - Interface customizability

- Goal setting/action planning
- Task performance (ie, ordering, flagging, prescribing)

Step 5. Collating, Summarizing, and Reporting the Results

Data extraction will be performed using Microsoft Excel (Microsoft Corp). The data elements for each dashboard identified will be displayed and coded in a spreadsheet, which will be used for analysis, mainly counts. This scoping review will follow the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [29] for reporting of methods and outcomes.

Results

In July 2020, electronic database searches were completed using the search strategies outlined in Multimedia Appendix 1, and 5188 results were retrieved and imported into Zotero reference management software (version 5.0.96.2; Corporation for Digital Scholarship) for management of records and retrieval of full-text articles prior to upload into Covidence online screening software [28]. After removal of duplicate results in Covidence [28], there were 3306 articles identified for title and abstract screening. A total of 2019 articles were retrieved for full-text review and will be reviewed in duplicate for eligibility. We aim to finish the review and draft the final report by mid-2022. Findings will be summarized in a narrative fashion while employing use of tables and graphs to illustrate key characteristics of dashboards in health care and will be submitted for publication along with the completed PRISMA-ScR reporting checklist.

Discussion

Future Planned Work

Currently, available literature lacks standard, consensus hierarchical descriptions of the different types of health care dashboards in use and their distinct design and implementation processes [15,30]. As the use of dashboards continues to increase, it is important for stakeholders to be able to communicate effectively with the designers and users of these tools. The authors intend to use the findings of this scoping review to inform the development of a taxonomy of the various types of dashboards a health care organization may choose to employ. This taxonomy will identify the relevant design elements that each type of dashboard includes to inform evidence about health care dashboard usability and purpose of use, and stakeholders, including end users. Finally, the review will provide evidence of the extent to which rigorous practices are used in the development, evaluation, and implementation of health care dashboards, each of which ultimately contributes to a dashboard's success.

The findings of this scoping review will additionally inform the design of a future meta-analysis and meta-synthesis of dashboard evaluations, if possible, in consideration of the heterogeneity of the studies identified in this scoping review.

Limitations

This scoping review methodology has several limitations. First, the search strategy does not include grey literature or conference abstracts since these are expected to provide insufficient detail for the data elements we plan to extract. This may cause some dashboards described in government and committee reports, dissertations, and conference proceedings to be overlooked. However, since the data extracted will mainly be summarized, and since we are not evaluating any causal effects or performing quantitative analyses, which would be more susceptible to publication bias, this will not be a major limitation. Second, because of the inclusion criteria, our findings will be most applicable to dashboards used in settings that provide direct health care; they will be less informative about public health tracking dashboards, including those used to monitor the COVID-19 pandemic and to perform contact tracing [31,32].

Conclusion

Health information technology continues to rapidly change the way health care organizations operate, and dashboards are an increasingly common tool. It is essential that key stakeholders have a clear understanding of what dashboards are, and which features are essential to specific end users for dashboard development. This scoping review will advance the field of health informatics by providing organizational leaders, clinical staff, dashboard developers, and quality improvement researchers with a clear and concise overview of the literature in this field, and by highlighting research gaps.

Acknowledgments

This scoping review was funded by the National Institute of Diabetes and Digestive and Kidney Diseases through a K23 award (K23DK118179) to JEK, and the US Department of Veterans Affairs (1 I50 HX003251-01) Maintaining Implementation Through Dynamic Adaptations (MIDAS; QUE 20-025). The funders played no role in the study design, decision to publish, or drafting of the manuscript.

Authors' Contributions

DH contributed to the study concept and design, pretesting and refinement of search strategy, citation and database management, screening of identified articles, and drafting of manuscript. JEK contributed to study concept and design, pretesting and refinement of search strategy, and drafting of manuscript. ADR contributed to study concept and design, screening of identified articles, and

critical revision of the manuscript. JS, LJD, ZLL, and PNP contributed to study concept and design and critical revision of the manuscript. MLC contributed to pretesting and refinement of search strategy, screening of identified articles, and critical revision of the manuscript. OJG contributed to study concept and design and screening of identified articles. ANK, RG, and AR screened identified articles. All authors read and approved the paper for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Database search strategies.
[DOCX File , 14 KB - resprot_v11i3e34894_app1.docx]

References

1. Unlocking Business Intelligence in Healthcare. Tableau. URL: https://www.tableau.com/learn/articles/business-intelligence/healthcare [accessed 2020-11-30]
2. What Is Pay for Performance in Healthcare? NEJM Catalyst, Innovations in Care Delivery. 2018 Mar 01. URL: https://catalyst.nejm.org/doi/abs/10.1056/CAT.18.0245 [accessed 2020-09-18]
3. 2021 ORYX Performance Measure Reporting Requirements. The Joint Commission. 2020 Oct. URL: https://www.jointcommission.org/-/media/tjc/documents/measurement/oryx/cy2021-oryx-reporting-requirements-oct2020.pdf [accessed 2021-03-25]
4. Kyerematenteng K, Robidoux R, D'Egidio G, Fernando SM, Neilipovitz D. An Analysis of Pay-for-Performance Schemes and Their Potential Impacts on Health Systems and Outcomes for Patients. Crit Care Res Pract 2019;2019:8943972 [FREE Full text] [doi: 10.1155/2019/8943972] [Medline: 31321097]
5. Office-based Physician Electronic Health Record Adoption. Office of the National Coordinator for Health Information Technology. 2019 Jan. URL: https://www.healthit.gov/data/quickstats/office-based-physician-electronic-health-record-adoption [accessed 2020-12-01]
6. Non-federal Acute Care Hospital Electronic Health Record Adoption. Office of the National Coordinator for Health Information Technology. 2017 Sep. URL: https://www.healthit.gov/data/quickstats/non-federal-acute-care-hospital-electronic-health-record-adoption [accessed 2020-12-01]
7. Atasoy H, Greenwood BN, McCullough JS. The Digitization of Patient Care: A Review of the Effects of Electronic Health Records on Health Care Quality and Utilization. Annu Rev Public Health 2019 Apr 01;40:487-500. [doi: 10.1146/annurev-publhealth-040218-044206] [Medline: 30566385]
8. Bonney W. Applicability of Business Intelligence in Electronic Health Record. Procedia - Social and Behavioral Sciences 2013 Feb;73:257-262. [doi: 10.1016/j.sbspro.2013.02.050]
9. Loewen L, Roudsari A. Evidence for Busines Intelligence in Health Care: A Literature Review. Stud Health Technol Inform 2017;235:579-583. [doi: 10.3233/978-1-61499-753-5-579] [Medline: 28423859]
10. Rivas C, Tkacz D, Antao L, Mentzakis E, Gordon M, Anstee S, et al. Automated analysis of free-text comments and dashboard representations in patient experience surveys: a multimethod co-design study. Health Serv Deliv Res 2019 Jul;7(23):1-160. [doi: 10.3310/hsdr07230] [Medline: 31287638]
11. Buttigieg SC, Pace A, Rathert C. Hospital performance dashboards: a literature review. J of Health Org and Mgt 2017 May 15;31(3):385-406. [doi: 10.1108/jhom-04-2017-0088] [Medline: 28686130]
12. Making Healthcare Safer III. A Critical Analysis of Existing and Emerging Patient Safety Practices. Agency for Healthcare Research and Quality. 2020 Mar. URL: https://www.ahrq.gov/sites/default/files/wysiwyg/research/findings/making-healthcare-safer/mhs3/making-healthcare-safer-III.pdf [accessed 2020-09-03]
13. Panicker V, Lee D, Wetmore M, Rampton J, Smith R, Moniz M, et al. Designing Tailored Displays for Clinical Practice Feedback: Developing Requirements with User Stories. Stud Health Technol Inform 2019 Aug 21;264:1308-1312 [FREE Full text] [doi: 10.3233/SHTI190438] [Medline: 31438137]
14. Khairat SS, Dukkipati A, Lauria HA, Bice T, Travers D, Carson SS. The Impact of Visualization Dashboards on Quality of Care and Clinician Satisfaction: Integrative Literature Review. JMIR Hum Factors 2018 May 31;5(2):e22. [doi: 10.2196/humanfactors.9328] [Medline: 29853440]
15. Vazquez-Ingelmo A, Garcia-Penalvo FJ, Theron R. Information Dashboards and Tailoring Capabilities - A Systematic Literature Review. IEEE Access 2019;7:109673-109688. [doi: 10.1109/access.2019.2933472]
16. Laurent G, Moussa MD, Cirenei C, Tavernier B, Marcilly R, Lamer A. Development, implementation and preliminary evaluation of clinical dashboards in a department of anesthesia. J Clin Monit Comput 2020 May 16;35(3):617-626. [doi: 10.1007/s10877-020-00522-x] [Medline: 32418147]
17. Bunch K, Allin B, Jolly M, Hardie T, Knight M. Developing a set of consensus indicators to support maternity service quality improvement: using Core Outcome Set methodology including a Delphi process. BJOG: Int J Obstet Gy 2018 Jun 14;125(12):1612-1618. [doi: 10.1111/1471-0528.15282] [Medline: 29770557]

18. Mould DR, Upton RN, Wojciechowski J, Phan BL, Tse S, Dubinsky MC. Dashboards for Therapeutic Monoclonal Antibodies: Learning and Confirming. AAPS J 2018 Jun 14;20(4):76. [doi: 10.1208/s12248-018-0237-2] [Medline: 29904814]
19. Eser A, Primas C, Reinisch S, Vogelsang H, Novacek G, Mould DR, et al. Prediction of Individual Serum Infliximab Concentrations in Inflammatory Bowel Disease by a Bayesian Dashboard System. The Journal of Clinical Pharmacology 2018 Jan 30;58(6):790-802. [doi: 10.1002/jcph.1069] [Medline: 29381220]
20. Waltz TJ, Powell BJ, Matthieu MM, Damschroder LJ, Chinman MJ, Smith JL, et al. Use of concept mapping to characterize relationships among implementation strategies and assess their feasibility and importance: results from the Expert Recommendations for Implementing Change (ERIC) study. Implementation Sci 2015 Aug 7;10(1):1-8. [doi: 10.1186/s13012-015-0295-0] [Medline: 26249843]
21. Dowding D, Randell R, Gardner P, Fitzpatrick G, Dykes P, Favela J, et al. Dashboards for improving patient care: Review of the literature. International Journal of Medical Informatics 2015 Feb;84(2):87-100. [doi: 10.1016/j.ijmedinf.2014.10.001] [Medline: 25453274]
22. Rivas C, Tkacz D, Laurence A, Mentzakis E, Gordon M, Anstee S, et al. Scoping review of clinical digital toolkit design. In: Automated analysis of free-text comments and dashboard representations in patient experience surveys: a multimethod co-design study. Southampton (UK): NIHR Journals Library; Jul 2019.
23. Wilbanks BA, Langford PA. A review of dashboards for data analytics in nursing. Comput Inform Nurs 2014 Nov;32(11):545-549. [doi: 10.1097/CIN.0000000000000106] [Medline: 25310224]
24. Aromatis E, Munn Z. JBI Manual for Evidence Synthesis. Joanna Briggs Institute. URL: https://jbi-global-wiki.refined.site/space/MANUAL [accessed 2020-08-18]
25. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Dec 19;18(1):143 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
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 KK. Scoping studies: advancing the methodology. Implement Sci 2010;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
28. Covidence. Melbourne, Australia: Veritas Health Information URL: https://www.covidence.org [accessed 2020-04-01]
29. 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 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
30. Barnum T, Vaez K, Cesarone D, Yingling CT. Your Data Looks Good on a Dashboard. Health Information and Management Systems Society (HIMSS). URL: https://www.himss.org/resources/your-data-looks-good-dashboard [accessed 2021-03-25]
31. Berry I, Soucy JR, Tuite A, Fisman D, COVID-19 Canada Open Data Working Group. Open access epidemiologic data and an interactive dashboard to monitor the COVID-19 outbreak in Canada. CMAJ 2020 Apr 14;192(15):E420 [FREE Full text] [doi: 10.1503/cmaj.75262] [Medline: 32392510]
32. Verhagen MD, Brazel DM, Dowd JB, Kashnitsky I, Mills MC. Forecasting spatial, socioeconomic and demographic variation in COVID-19 health care demand in England and Wales. BMC Med 2020 Jun 29;18(1):203 [FREE Full text] [doi: 10.1186/s12916-020-01646-2] [Medline: 32594909]

Abbreviations

MeSH: medical subject headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by G Eysenbach; submitted 11.11.21; peer-reviewed by T Ndabu, F Velayati; comments to author 28.12.21; revised version received 11.01.22; accepted 11.01.22; published 02.03.22.

Please cite as:
Helminski D, Kurlander JE, Renji AD, Sussman JB, Pfeiffer PN, Conte ML, Gadabu OJ, Kokaly AN, Goldberg R, Ranusch A, Damschroder LJ, Landis-Lewis Z
Dashboards in Health Care Settings: Protocol for a Scoping Review
JMIR Res Protoc 2022;11(3):e34894
URL: https://www.researchprotocols.org/2022/3/e34894
doi:10.2196/34894
PMID:35234650

Protocol

Evaluation of Digital Interventions for Physical Activity Promotion: Protocol for a Scoping Review

Karina Karolina De Santis[1], PhD; Tina Jahnel[1,2], PhD; Lea Mergenthal[1]; Hajo Zeeb[1], MD, PhD, Prof Dr; Katja Matthias[3], MD, Prof Dr

[1]Department of Prevention and Evaluation, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

[2]Faculty 11 Human and Health Sciences, University of Bremen, Bremen, Germany

[3]Faculty of Electrical Engineering and Computer Science, University of Applied Science Stralsund, Stralsund, Germany

Corresponding Author:
Karina Karolina De Santis, PhD
Department of Prevention and Evaluation
Leibniz Institute for Prevention Research and Epidemiology – BIPS
Achterstraße 30
Bremen, 28359
Germany
Phone: 49 421 218 56 908
Fax: 49 421 218 56 941
Email: desantis@leibniz-bips.de

Abstract

Background: Digital interventions (DIs) could support physical activity (PA) promotion, according to recent reviews. However, it remains unclear if and how DIs for PA promotion are evaluated; thus, it is unclear if they support behavior change in real-world settings. A mapping of evidence from published reviews is required to focus on the evaluation of DIs for PA promotion.

Objective: The aim of our study is to investigate evaluation strategies for any outcome in the context of DIs for PA promotion by conducting a scoping review of published reviews.

Methods: Our scoping review adheres to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. The information sources include bibliographic databases (MEDLINE, PsycINFO, and CINAHL) and the bibliographies of the selected studies. The electronic search strategy was developed and conducted in collaboration with an experienced database specialist. The electronic search was conducted in English with no limits up to March 19, 2021, for sources with the terms *digital intervention AND evaluation AND physical activity* in titles or abstracts. After deduplication, 300 reviews selected from 4912 search results were assessed for eligibility by 2 authors working independently. The inclusion criteria were (1) healthy or clinical samples (population), (2) DIs for PA promotion (intervention), (3) comparisons to any other intervention or no intervention (comparison), (4) evaluation strategies (methods, results, or frameworks) for any outcome in the context of DIs for PA promotion (outcome), and (5) any published review (study type). According to the consensus reached during a discussion, 40 reviews met the inclusion criteria—36 from the electronic search and 4 from the manual search of the bibliographies of the 36 reviews. All reviews reported the evaluation strategies for any outcomes in the context of DIs for PA promotion in healthy or clinical samples. Data coding and the quality appraisal of systematic reviews are currently being performed independently by 2 authors.

Results: Our scoping review includes data from 40 published reviews (1 rapid review, 9 scoping reviews, and 30 systematic reviews). The focus of data coding is on evaluation strategies in the context of DIs for PA promotion and on the critical appraisal of the included systematic reviews. The final consensus regarding all data is expected in early 2022.

Conclusions: Interventions for PA promotion that are supported by digital technologies require evaluation to ensure their efficacy in real-world settings. Our scoping review is needed because it addresses novel objectives that focus on such evaluations and are not answered in the published reviews identified in our search. The evaluation strategies addressing DIs for PA promotion will be mapped to synthesize the results that have been reported in published reviews so far.

International Registered Report Identifier (IRRID): DERR1-10.2196/35332

(JMIR Res Protoc 2022;11(3):e35332) doi:10.2196/35332

KEYWORDS

evaluation; digital intervention; physical activity promotion; scoping review; digital health

Introduction

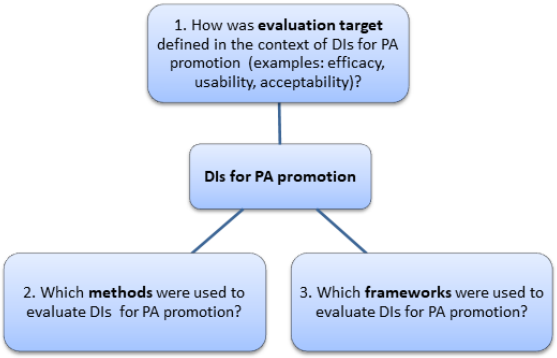
The field of digital public health is rapidly developing [1]. Digitization is likely to have a major impact on therapy in the future, and it already increasingly contributes to prevention and health promotion [2]. Interventions supported by modern technologies (ie, digital interventions [DIs]) are enormously popular in the context of healthy lifestyle and behavior change, including physical activity (PA) promotion [3,4]. A number of challenges exist in this rapidly developing field, including the gap between clinical or preventive interests and commercial interests [5], the ethics of data storage and usage [6], and development issues (theoretical and evidence-based foundations of new DIs) [7].

One key question in this new field of research is if DIs truly work in any health context. In light of the long history of evidence-based medicine with guidelines on how to assess the effectiveness of nondigital health interventions, comprehensive guidelines on the systematic evaluation of DIs are still scarce. Evaluation is also important for justifying and informing policy, program, and funding decisions. Although initial evaluation criteria and frameworks have been proposed for DIs, this preliminary work lacks guidance as to when and to which degree these criteria should be applied [8-10]. For example, as already pointed out in 2015 [11], it remains unclear if and how DIs for PA promotion are evaluated. Evaluation in this context is essential for understanding if DIs support behavior change in real-world settings, so that the sustainable, effective, and efficient use of DIs can be achieved [12,13]. However, real-world DIs are complex and difficult to evaluate. Among others, the challenges of evaluating DIs include contextual factors, such as settings, target populations, intervention functions, or intended outcomes; as well as organizational, political, or resourcing factors. Some of the practical challenges in conducting evaluations include using appropriate evaluation methods and tools, understanding what counts as evidence, and understanding how such evidence is applied and interpreted [14-16]. Therefore, a review of the literature is required to focus on evaluation methods in the context of DIs for PA promotion.

According to a PubMed search, 155 reviews with the terms *digital AND physical activity* have already been published up to November 17, 2021 (including 55 reviews in 2021 alone). Due to the high number of potentially relevant reviews that have already been published on this topic, the mapping of existing evidence is required to investigate if and how the evaluation of DIs for PA promotion was addressed in such publications. Mapping is important for identifying any evidence gaps that could be addressed in future reviews of primary studies.

The aim of this study is to investigate evaluation strategies for any outcome in the context of DIs for PA promotion by conducting a scoping review of published reviews. The three main objectives of this scoping review address the evaluation target, methods, and theoretical frameworks (Figure 1).

Figure 1. Objectives of this scoping review. DI: digital intervention; PA: physical activity.



Methods

Study Design

Our study uses a scoping review design. The study adheres to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [17]. The PRISMA-ScR checklist will be reported in an appendix once the scoping review is complete.

Protocol and Registration

The study was registered at the Center for Open Science [18] and was planned to include the following two parts: (1) an overview (review) of reviews and (2) a scoping review of primary studies. This protocol addresses only part 1 of the planned study. The need for part 2 will be established once the results of part 1 are available. In contrast to our registration [18], the study will be performed as a scoping review of reviews due to the availability of appropriate guidance (PRISMA-ScR [17]). Furthermore, our study has a broader focus on evaluation methods relative to that of overviews of reviews that typically address specific outcomes of interventions in the health context.

The electronic literature search was conducted on March 19, 2021, prior to study registration on May 3, 2021 [18]. The results were screened by 1 author for the presence of other overviews or scoping reviews addressing the same aims as those planned in our study. This step was necessary to prevent any research waste. According to meta-research [19-21], many reviews of interventions in the health context are redundant because they

either address the same aims as those addressed by other existing reviews or cite the same primary studies. This problem is so extensive that some reviews do not include any unique primary studies that are not cited in other reviews, and there are as many reviews as or even more reviews than there are primary studies in some fields [19-21]. The procedure of checking if a new review is required prior to study registration may be especially necessary in rapidly developing fields, such as digitally supported interventions, or when addressing commonly investigated outcomes, such as PA promotion.

Eligibility Criteria

The eligibility criteria for our scoping review are based on the Population, Intervention, Comparison, Outcome, and Study Type (PICOS) criteria (Textbox 1). Our review has a methods focus and thus targets any outcome.

Textbox 1. Eligibility criteria for the scoping review.

Inclusion criteria

1. Population: human samples of any age (children or adults) and health status (healthy or clinical)
2. Intervention: digital interventions for physical activity promotion as a primary outcome
3. Comparison: comparisons to any other intervention or no intervention
4. Outcome: evaluation of a digital intervention that is planned or performed by using any method for any target
5. Study type: any review (systematic, scoping, rapid, narrative, overview, or other)
6. Publication status: published in a peer-reviewed journal
7. Publication language: English or German
8. Access to the full texts of studies selected for data coding

Exclusion criteria

1. Nonhuman studies
2. Digital interventions for physical activity promotion are not applied or are not the primary intervention (included as a control to or part of another intervention)
3. Evaluation of digital interventions is not addressed (not planned or not performed)
4. Other study type: primary study, comment, correction, letter, editorial, or protocol
5. Other publication status: conference paper, unpublished report, thesis, or book
6. Language other than English or German
7. No access to the full texts of studies selected for data coding

Information Sources

The information sources for the scoping review include bibliographic databases (MEDLINE, PsycINFO, and CINAHL) as well as the bibliographies of the selected studies. These databases were chosen because they delivered the most relevant studies in our other searches for DIs in the context of public health.

Search

The electronic search strategy was developed iteratively by the team in consultation with a professional librarian. The search terms and corresponding MeSH (Medical Subject Headings) terms were derived to address the following three main search topics: (1) DIs (with a mobile app), (2) evaluation, and (3) PA. The full search strategy will be reported in an appendix once the scoping review is complete. A summary of the electronic search and its outcomes is shown in Table 1. In addition, the bibliographies of the included studies were manually screened for additional relevant sources.

Table 1. Summary of the electronic search strategy[a].

Databases (time frame)	Search strategy summary (search terms)	Studies (N=8272), n
MEDLINE through Ovid (from inception through to March 19, 2021)	Title OR abstract (*mobile application AND evaluation AND physical activity*)	4776
PsycINFO through Ovid (from inception through to March 19, 2021)	Title OR abstract (*mobile application AND evaluation AND physical activity*)	1157
CINAHL through EBSCO (from inception through to March 19, 2021)	Title OR abstract (*mobile application AND evaluation AND physical activity*)	2339

[a]The electronic search was conducted in English with no limits by a team assistant on March 19, 2021.

Selection of Sources of Evidence

The electronic search results (8272 studies) were stored in EndNote X9 (Clarivate). Following the removal of duplicates in EndNote, 4912 remaining studies were divided into 2 groups (reviews or nonreviews) by using the *smart groups* function in EndNote. All identified reviews (300/4912, 6.11%) were exported into a new EndNote library and divided into the following five groups, depending on review type, by using the *smart groups* function: (1) overview of reviews, (2) rapid review, (3) scoping review, (4) narrative review, and (5) systematic review.

Study selection was conducted in 3 steps. First, 2 authors independently screened all titles and abstracts in each smart group for inclusion and reached consensus during a discussion. Second, 2 authors independently screened the studies selected for the full-text inspection and reached consensus during a discussion. Third, once the final study selection from the electronic search was completed, 2 authors manually screened the bibliographies of the included studies for additional relevant sources and reached consensus during a discussion. The outcomes of the study selection are summarized in Table 2 and Figure 2. The complete list of included and excluded studies will be reported in an appendix once the scoping review is complete.

Table 2. Summary of study selection.

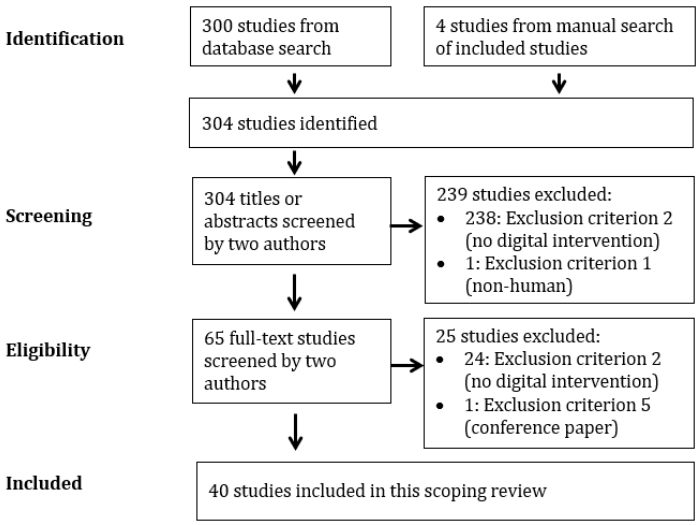
Review type[a]	Studies, n	Exclusion[b]
Reviews from electronic search (n=300)		
Overview	3	N/A[c]
Excluded based on title or abstract	3	Exclusion criterion 2
Excluded based on full text	0	N/A
Included	0	N/A
Rapid review	4	N/A
Excluded based on title or abstract	3	Exclusion criterion 2
Excluded based on full text	0	N/A
Included	1	N/A
Scoping review	21	N/A
Excluded based on title or abstract	1	Exclusion criterion 1
Excluded based on title or abstract	10	Exclusion criterion 2
Excluded based on full text	1	Exclusion criterion 2
Included	9	N/A
Narrative review	51	N/A
Excluded based on title or abstract	45	Exclusion criterion 2
Excluded based on full text	6	Exclusion criterion 2
Included	0	N/A
Systematic review	221	N/A
Excluded based on title or abstract	177	Exclusion criterion 2
Excluded based on full text	17	Exclusion criterion 2
Excluded based on full text	1	Exclusion criterion 5
Included	26	N/A
Reviews from manual search (n=4)		
Systematic review	4	N/A
Included from the manual search of the bibliographies of all included reviews	4	N/A

[a]Review type was established based on the information in titles or abstracts. A total of 40 reviews were included from the electronic and manual searches.

[b]Exclusion criteria are shown in Textbox 1.

[c]N/A: not applicable.

Figure 2. Study selection (PRISMA [Preferred Reporting Items for Systematic Reviews and Meta-Analyses] flowchart).



Data Charting

A form for coding and capturing all data was self-developed in Microsoft Excel and calibrated within the team. Two authors will code all data independently and reach consensus during a discussion. Data coding is currently in progress.

Data Items

Data items that will be coded in the scoping review are reported in Textbox 2. These items were chosen to address the objectives of the scoping review (Figure 1). All data will be reported in an appendix once the scoping review is complete.

In addition to the definitions provided in Textbox 2, the following operationalization definitions were used to improve the reliability and validity of study selection and data coding in the scoping review:

1. A *healthy* population was defined as samples without acute or chronic illnesses, but this could include samples at risk for various clinical illnesses.
2. *Digital intervention* was defined as an intervention delivered or supported by digital tools. *Digital tools* were defined as any digitally supported technologies for automated and continuous self-monitoring and feedback. This includes smartphone apps, activity trackers, and web-based software but excludes digital tools, such as pedometers and accelerometers, that do not offer tracked measures or feedback over time [22]. Pedometers or accelerometers were included as part of DIs or with other digital tools in the minority of primary studies in some reviews. Reviews were excluded if all or the majority of their primary studies used only pedometers or accelerometers.
3. *PA promotion* was defined as any primary outcome focusing on PA promotion. Reviews were excluded if PA promotion was assessed as part of a healthy lifestyle or as a secondary outcome for the management of weight, blood sugar, or sports injuries; for balance and mobility training following surgeries; or for such training in the management of neurological disorders.

Textbox 2. Data items in this scoping review.

Bibliographic information

- First author
- Year of publication
- Region of corresponding author (continent)
- Study title
- Study aim according to study authors
- Funding sources and conflicts of interest according to study authors

Population

- Population by health status (healthy individuals or individuals with clinical diagnoses)
- Population by diagnosis (none or diagnosis name)
- Population by diagnosis type (none, any, mental, neurological, or somatic)
- Population by age (any age; minors aged up to 18 years; or adults aged 18 years or older, including specific subgroups [eg, older adults])

Intervention

- Any digital intervention

Comparison

- Comparison (any, independent control group with another intervention, or baseline in pre-post studies without control groups)

Outcome

- Any outcome in the context of physical activity promotion
- Outcome focus (general fitness, mobility, or other)

Study (review) type

- Review type (rapid, scoping, or systematic)
- Primary studies in review (number)
- Published primary studies in review (number)
- Unique published studies that do not overlap with primary studies in other reviews (number)
- Primary study design in review (only randomized controlled trials [RCTs] or any designs, including RCTs and non-RCTs)

Evaluation

- Evaluation target (user outcomes [eg, efficacy, usability, acceptability, or tool performance or validation])
- Evaluation method (objective automated data from the tools, scales, or tests or the validation of tool data vs another method)
- Theory framework type (not reported or framework name)
- Theory framework description according to study authors (eg, frameworks used in tool development)
- Requirements for the efficacy of digital interventions according to study authors (eg, engagement with the tool)
- Evidence gaps according to study authors (recommendations for future research, limitations, and conclusions)

Critical Appraisal of Individual Sources of Evidence

The critical appraisal will be conducted only for the systematic reviews by using a tool that was specifically developed for such reviews (A Measurement Tool to Assess Systematic Reviews, Version 2 [AMSTAR2]) [23]. AMSTAR2 has acceptable psychometric properties and is an appropriate tool for appraising systematic reviews of interventions in the health context [23,24]. The tool includes 16 items that need to be rated to derive the overall confidence rating for the results of a systematic review (critically low, low, moderate, or high) [23]. All 16 items will be rated according to the AMSTAR2 scoring guidelines [23]. The overall confidence rating will be derived for each systematic review based on a combination of scores for 7 critical items and 9 noncritical items, in accordance with the AMSTAR2 guidelines [23]. In general, critically low ratings are assigned if at least 2 critical items are not fulfilled (rated as “no”) on AMSTAR2.

A form for appraising systematic reviews with AMSTAR2 was self-developed in Microsoft Excel. Two authors will appraise

all systematic reviews independently and reach consensus during a discussion. The appraisals will be performed in 2 phases by both authors.

During phase 1, the systematic reviews with critically low confidence ratings will be identified. This will be done by using 2 AMSTAR2 items (item 2: presence of a review protocol; item 7: presence of a list of excluded studies). Systematic reviews that do not fulfill both of these items will receive critically low confidence ratings. These two items address the most common limitations in systematic reviews of interventions in the health context [20], and they were selected according to a fast and frugal decision tree for the critical appraisal of systematic reviews [25].

During phase 2, the systematic reviews that fulfilled at least 1 of the 2 items in phase 1 (item 2 or 7) will be rated by using all 16 AMSTAR2 items. This will be done to identify further systematic reviews with critically low confidence ratings and other reviews with low, moderate, or high confidence ratings.

Once consensus is reached, the final overall confidence ratings for each systematic review derived from AMSTAR2 will be reported in an appendix. The appraisal procedure is currently in progress.

In addition to deriving the overall confidence ratings, the AMSTAR2 scores will be used for 2 meta-research studies. Both studies will be performed because the overall confidence ratings on AMSTAR2 alone poorly discriminate among systematic reviews of various interventions in the health context [20].

The first meta-research study will be performed in addition to our original registration [18]. This study will address the following two aims: (1) to identify common strengths and weaknesses in systematic reviews of DIs for PA promotion and (2) to assess the stability of the overall confidence ratings. To investigate both aims, all systematic reviews appraised with the two AMSTAR2 items in phase 1 will be appraised with all 16 AMSTAR2 items. Two authors will appraise all systematic reviews independently and reach consensus during a discussion. To address aim 1, the scores on the individual AMSTAR2 items for each of the 30 systematic reviews will be presented on a bar graph to visualize the strengths (fulfilled items rated as “yes” or “partial yes”) and the weaknesses (not fulfilled items rated as “no”) in each review. To address aim 2, the outcomes of appraisals involving 2 AMSTAR2 items and those involving 16 AMSTAR2 items will be compared descriptively according to the overall rating correctness and the total appraisal time. Furthermore, the “yes” or “yes + partial yes” ratings will be expressed as percentage scores out of all 16 ratings assigned to each systematic review in accordance with methods described elsewhere [26]. Such percentage scores for “yes” or “yes + partial yes” ratings will be compared between 2 groups of systematic reviews, which will be based on the reviews’ overall confidence ratings (critically low and low vs moderate and high). The comparisons will be computed in IBM-SPSS 24 (IBM Corporation) and reported as odds ratios with 95% CIs for the nominal variables or as mean difference scores with 95% CIs for the continuous variables. Finally, the overall confidence ratings derived from different combinations of critical items will be descriptively compared.

The second meta-research study will be performed in accordance with the plan in our original registration [18]. The aim of the second meta-research study is to compare the outcomes of AMSTAR2 appraisals by using the original scoring guidelines [23] and the revised scoring guidelines proposed by us. Two authors will appraise the same systematic reviews independently; one will use the original scoring guidelines, and one will use the revised scoring guidelines. The overall confidence ratings will be graphically summarized and descriptively compared. This will be done to identify the sources of similarities and discrepancies between the outcomes of both scoring methods and to test the usefulness of the revised scoring guidelines for appraising systematic reviews.

Synthesis of Results

Studies will be grouped according to their designs. The coded data will be synthesized either by using descriptive statistics (relative frequencies) or narratively within each group. If applicable, evidence maps [27] will be used to visualize the results according to the three objectives of the scoping review (Figure 1). The overall confidence ratings for all systematic reviews will be graphically synthesized by using a bar graph to visualize the outcomes of the critical appraisal.

Results

Included Studies

Our electronic search identified 6.11% (300/4912) studies designated as reviews of any type in titles or abstracts. Of the 300 reviews, 36 (12%) met the inclusion criteria (Textbox 1) for the scoping review. An additional 4 reviews were selected following a manual search of the bibliographies of the 36 included reviews. Thus, 40 reviews were included in the scoping review (Figure 2). Of the 40 reviews, 1 (2.5%) was a rapid review, 9 (22.5%) were scoping reviews, and 30 (75%) were systematic reviews (Table 2).

Further Results

Data coding and the critical appraisal of systematic reviews are currently in progress and are expected to be completed in early 2022.

Discussion

Principal Results

Our electronic search revealed that 300 reviews indexed in 3 bibliographic databases have already been published on interventions supported by digital technologies and designed for PA. Thus, to prevent research waste resulting from the contribution of another review of primary studies, our scoping review will synthesize the findings of such published reviews. Among the 300 reviews, we identified 40 reviews that specifically focused on evaluation strategies in the context of DIs for PA promotion (36 from the electronic search and 4 from the manual search of the bibliographies of the 36 included reviews). According to our preliminary findings, our scoping

review is needed because it addresses novel objectives that focus on evaluation strategies in the context of DIs for PA promotion.

Interest in DIs for PA Promotion

The large number of reviews on digitization and PA highlights 2 important issues so far. First, the academic field of DIs for the core aspects of public health (prevention and health promotion) is rapidly developing [1]. This development is probably related to the general technological progress [4] and economic interest in the digitization of health [5]. However, the digitization of health is also associated with various challenges, such as access to digital tools, digital health competence, and ethical issues related to data storage and usage [28]. Interestingly, although digital technologies for PA are enormously popular [3], it remains unclear if they work (ie, if they promote PA). Second, there is a need to carefully inspect the published literature on digitization and PA promotion before conducting a new review that may be redundant and may contribute to research waste. Our scoping review of other reviews will help to identify evidence gaps and possible research questions that have not been addressed in the academic literature so far. Such evidence gaps will be used to determine if a new scoping review or systematic review of primary studies is required in this rapidly developing field.

Limitations

Although the data coding and appraisal are still ongoing, 3 main limitations have already been identified in the scoping review. First, the development of the search strategy was difficult due to the heterogeneous terminology used in the field of DIs. Professional assistance with the development of the search syntax for bibliographic databases was an essential requirement for determining the validity of our search. Our search syntax was calibrated and extensively pretested, and the search was conducted under the supervision of an experienced librarian who specialized in bibliographic databases. Although a large number of relevant sources suggests that the search was valid, we acknowledge that additional search terms could have been used to identify further sources. Second, we searched only 3 electronic databases (MEDLINE, PsycINFO, and CINAHL). The pilot searches for reviews on DIs for PA promotion in the Cochrane Library and Scopus did not identify any additional relevant reviews. However, it cannot be ruled out that additional sources are available on other international databases. Third, the careful operationalization of definitions is required for study selection and data coding because this research field is young and is evolving. Thus, the definitions of some PICOS criteria (Textbox 1) had to be expanded to improve the reliability and validity of the study selection and data coding processes that were conducted after the pilot assessment of scoping reviews. Finally, the data sources for our scoping review (other published reviews) may not necessarily focus on the evaluation of DIs for PA promotion. Although this cannot be ruled out, our search strategy indeed included the term *evaluation*, meaning that all of the studies identified in the search included this term in their titles, abstracts, or keywords. Thus, the scoping review can systematically collate the information about the evaluation of DIs for PA promotion from other reviews to identify any evidence gaps that could be addressed in future reviews of primary studies. In general, our experiences so far highlight the need for high-quality documentation and the reporting of definitions in this relatively new and dynamically developing field.

Conclusions

Interventions for PA promotion supported by digital technologies require evaluation to ensure their efficacy in real-world settings. Our scoping review is needed because it addresses novel objectives that focus on such evaluations and are not answered in the published reviews identified in our search. The evaluation strategies addressing DIs for PA promotion will be mapped to synthesize the results that have been reported in published reviews so far.

Acknowledgments

We thank the database specialist—Lara Christianson, MLS—and the team assistant—Lisa Kühne, MSc—for their assistance in designing the search strategy and conducting the electronic search. The publication of this article was funded by the Open Access Fund of the University of Applied Science Stralsund. We gratefully acknowledge the support of the Leibniz-Science Campus Bremen Digital Public Health (lsc-diph.de), which is jointly funded by the Leibniz Association (grant W4/2018), the Federal State of Bremen, and the Leibniz Institute for Prevention Research and Epidemiology – BIPS.

Authors' Contributions

KKDS conceptualized the study, developed the methodology, selected the studies, wrote the first draft of the manuscript, and reviewed and edited the manuscript. TJ conceptualized the study and reviewed and edited the manuscript. LM selected the studies and reviewed and edited the manuscript. HZ conceptualized the study and reviewed and edited the manuscript. KM conceptualized the study, developed the methodology, selected the studies, wrote the first draft of the manuscript, and reviewed and edited the manuscript.

Conflicts of Interest

None declared.

References

1. Zeeb H, Pigeot I, Schüz B, Leibniz-WissenschaftsCampus Digital Public Health Bremen. [Digital public health-an overview]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Feb;63(2):137-144. [doi: 10.1007/s00103-019-03078-7] [Medline: 31919531]
2. De Santis KK, Jahnel T, Sina E, Wienert J, Zeeb H. Digitization and health in Germany: Cross-sectional nationwide survey. JMIR Public Health Surveill 2021 Nov 22;7(11):e32951 [FREE Full text] [doi: 10.2196/32951] [Medline: 34813493]
3. Cajita MI, Kline CE, Burke LE, Bigini EG, Imes CC. Feasible but not yet efficacious: A scoping review of wearable activity monitors in interventions targeting physical activity, sedentary behavior, and sleep. Curr Epidemiol Rep 2020 Mar;7(1):25-38 [FREE Full text] [doi: 10.1007/s40471-020-00229-2] [Medline: 33365227]
4. Taj F, Klein MCA, van Halteren A. Digital health behavior change technology: Bibliometric and scoping review of two decades of research. JMIR Mhealth Uhealth 2019 Dec 13;7(12):e13311 [FREE Full text] [doi: 10.2196/13311] [Medline: 31833836]
5. Eze ND, Mateus C, Hashiguchi TCO. Telemedicine in the OECD: An umbrella review of clinical and cost-effectiveness, patient experience and implementation. PLoS One 2020 Aug 13;15(8):e0237585. [doi: 10.1371/journal.pone.0237585] [Medline: 32790752]
6. Cordeiro JV. Digital technologies and data science as health enablers: An outline of appealing promises and compelling ethical, legal, and social challenges. Front Med (Lausanne) 2021 Jul 08;8:647897 [FREE Full text] [doi: 10.3389/fmed.2021.647897] [Medline: 34307394]
7. Sporrel K, Nibbeling N, Wang S, Ettema D, Simons M. Unraveling mobile health exercise interventions for adults: Scoping review on the implementations and designs of persuasive strategies. JMIR Mhealth Uhealth 2021 Jan 18;9(1):e16282 [FREE Full text] [doi: 10.2196/16282] [Medline: 33459598]
8. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. Lancet Digit Health 2019 Jun;1(2):e52-e54 [FREE Full text] [doi: 10.1016/S2589-7500(19)30013-5] [Medline: 33323229]
9. Kowatsch T, Otto L, Harperink S, Cotti A, Schlieter H. A design and evaluation framework for digital health interventions. it - Information Technology 2019 Nov 20;61(5-6):253-263 [FREE Full text] [doi: 10.1515/itit-2019-0019]
10. Fynn JF, Hardeman W, Milton K, Jones AP. A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. BMC Public Health 2020 Jun 26;20(1):1000 [FREE Full text] [doi: 10.1186/s12889-020-09062-0] [Medline: 32586301]
11. Tabak M, Dekker-van Weering M, van Dijk H, Vollenbroek-Hutten M. Promoting daily physical activity by means of mobile gaming: A review of the state of the art. Games Health J 2015 Dec;4(6):460-469. [doi: 10.1089/g4h.2015.0010] [Medline: 26397179]
12. Lynch C, Bird S, Lythgo N, Selva-Raj I. Changing the physical activity behavior of adults with fitness trackers: A systematic review and meta-analysis. Am J Health Promot 2020 May;34(4):418-430. [doi: 10.1177/0890117119895204] [Medline: 31858812]
13. Silva AG, Simões P, Queirós A, Rocha NP, Rodrigues M. Effectiveness of mobile applications running on smartphones to promote physical activity: A systematic review with meta-analysis. Int J Environ Res Public Health 2020 Mar 27;17(7):2251 [FREE Full text] [doi: 10.3390/ijerph17072251] [Medline: 32230769]
14. Lim MSC, Wright CJC, Carrotte ER, Pedrana AE. Reach, engagement, and effectiveness: a systematic review of evaluation methodologies used in health promotion via social networking sites. Health Promot J Austr 2016 Feb;27(3):187-197. [doi: 10.1071/HE16057] [Medline: 27719734]
15. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. NPJ Digit Med 2019 May 13;2:38 [FREE Full text] [doi: 10.1038/s41746-019-0111-3] [Medline: 31304384]
16. Moshi MR, Tooher R, Merlin T. Suitability of current evaluation frameworks for use in the health technology assessment of mobile medical applications: A systematic review. Int J Technol Assess Health Care 2018 Jan;34(5):464-475. [doi: 10.1017/S026646231800051X] [Medline: 30201060]
17. 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 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
18. De Santis KK, Jahnel T, Mergenthal L, Matthias K. Evaluation of digital interventions for physical activity promotion. Open Science Framework. 2021. URL: https://osf.io/v82d3 [accessed 2021-05-01]
19. De Santis KK, Kaplan I. Assessing the quality of systematic reviews in healthcare using AMSTAR and AMSTAR2. Z Psychol 2020 Mar 31;228(1):36-42. [doi: 10.1027/2151-2604/a000397]
20. De Santis KK, Lorenz RC, Lakeberg M, Matthias K. The application of AMSTAR2 in 32 overviews of systematic reviews of interventions for mental and behavioural disorders: A cross-sectional study. Res Synth Methods. Epub ahead of print 2021 Oct 19. [doi: 10.1002/jrsm.1532] [Medline: 34664766]
21. Kedzior KK, Kaplan I. Tai chi and Parkinson's disease (PD): A systematic overview of the scientific quality of the past systematic reviews. Complement Ther Med 2019 Oct;46:144-152. [doi: 10.1016/j.ctim.2019.08.008] [Medline: 31519271]
22. Laranjo L, Ding D, Heleno B, Kocaballi B, Quiroz JC, Tong HL, et al. Do smartphone applications and activity trackers increase physical activity in adults? Systematic review, meta-analysis and metaregression. Br J Sports Med 2021 Apr;55(8):422-432. [doi: 10.1136/bjsports-2020-102892] [Medline: 33355160]

23. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017 Sep 21;358:j4008 [FREE Full text] [doi: 10.1136/bmj.j4008] [Medline: 28935701]
24. Lorenz RC, Matthias K, Pieper D, Wegewitz U, Morche J, Nocon M, et al. A psychometric study found AMSTAR 2 to be a valid and moderately reliable appraisal tool. J Clin Epidemiol 2019 Oct;114:133-140. [doi: 10.1016/j.jclinepi.2019.05.028] [Medline: 31152864]
25. Lorenz R, Jenny M, Jacobs A, Matthias K. Fast and frugal decision tree for the critical appraisal of systematic reviews in situations with limited time periods. In: Ger Med Sci. 2021 Presented at: 22. Jahrestagung des Deutschen Netzwerks Evidenzbasierte Medizin e. V; February 24-26, 2021; Digital. [doi: 10.3205/21ebm039]
26. Matthias K, Rissling O, Pieper D, Morche J, Nocon M, Jacobs A, et al. The methodological quality of systematic reviews on the treatment of adult major depression needs improvement according to AMSTAR 2: A cross-sectional study. Heliyon 2020 Sep 01;6(9):e04776 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e04776] [Medline: 32939412]
27. Saran A, White H. Evidence and gap maps: a comparison of different approaches. Campbell Syst Rev 2018 Oct 12;14(1):1-38 [FREE Full text] [doi: 10.4073/cmdp.2018.2]
28. Dadaczynski K, Tolks D. Digitale public health: Chancen und herausforderungen internetbasierter technologien und anwendungen. Public Health Forum 2018 Sep 01;26(3):275-278. [doi: 10.1515/pubhef-2018-0059]

Abbreviations

AMSTAR2: A Measurement Tool to Assess Systematic Reviews, Version 2
DI: digital intervention
MeSH: Medical Subject Headings
PA: physical activity
PICOS: Population, Intervention, Comparison, Outcome, Study Type
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by T Leung; submitted 01.12.21; peer-reviewed by K Schneider; comments to author 22.12.21; revised version received 12.01.22; accepted 03.02.22; published 03.03.22.

Please cite as:
De Santis KK, Jahnel T, Mergenthal L, Zeeb H, Matthias K
Evaluation of Digital Interventions for Physical Activity Promotion: Protocol for a Scoping Review
JMIR Res Protoc 2022;11(3):e35332
URL: https://www.researchprotocols.org/2022/3/e35332
doi:10.2196/35332
PMID:35238321

Protocol

The Volume and Tone of Twitter Posts About Cannabis Use During Pregnancy: Protocol for a Scoping Review

Liam Cresswell[1,2], BHSc; Lisette Espin-Noboa[3], MSc; Malia S Q Murphy[1], PhD; Serine Ramlawi[1], MSc; Mark C Walker[1,4,5,6,7], MSc, MHM, MD; Márton Karsai[3], PhD, HDR; Daniel J Corsi[1,4,5,6,8], PhD

[1]Obstetrics & Maternal Newborn Investigations Research Group, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

[2]Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

[3]Department of Network and Data Science, Central European University, Vienna, Austria

[4]Department of Obstetrics and Gynecology, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

[5]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

[6]Better Outcomes Registry & Network Ontario, Ottawa, ON, Canada

[7]International and Global Health Office, University of Ottawa, Ottawa, ON, Canada

[8]CHEO Research Institute, Ottawa, ON, Canada

Corresponding Author:
Daniel J Corsi, PhD
CHEO Research Institute
501 Smyth Rd
Centre for Practice Changing Research, Room L1242, Box 241
Ottawa, ON, K1H 8L6
Canada
Phone: 1 613 737 8899
Email: dcorsi@ohri.ca

Abstract

Background: Cannabis use has increased in Canada since its legalization in 2018, including among pregnant women who may be motivated to use cannabis to reduce symptoms of nausea and vomiting. However, a growing body of research suggests that cannabis use during pregnancy may harm the developing fetus. As a result, patients increasingly seek medical advice from online sources, but these platforms may also spread anecdotal descriptions or misinformation. Given the possible disconnect between online messaging and evidence-based research about the effects of cannabis use during pregnancy, there is a potential for advice taken from social media to affect the health of mothers and their babies.

Objective: This study aims to quantify the volume and tone of English language posts related to cannabis use in pregnancy from January 2012 to December 2021.

Methods: Modeling published frameworks for scoping reviews, we will collect publicly available posts from Twitter that mention cannabis use during pregnancy and use the Twitter Application Programming Interface for Academic Research to extract data from tweets, including public metrics such as the number of likes, retweets, and quotes, as well as health effect mentions, sentiment, location, and users' interests. These data will be used to quantify how cannabis use during pregnancy is discussed on Twitter and to build a qualitative profile of supportive and opposing posters.

Results: The CHEO Research Ethics Board reviewed our project and granted an exemption in May 2021. As of December 2021, we have gained approval to use the Twitter Application Programming Interface for Academic Research and have developed a preliminary search strategy that returns over 3 million unique tweets posted between 2012 and 2021.

Conclusions: Understanding how Twitter is being used to discuss cannabis use during pregnancy will help public health agencies and health care providers assess the messaging patients may be receiving and develop communication strategies to counter misinformation, especially in geographical regions where legalization is recent or imminent. Most importantly, we foresee that our findings will assist expecting families in making informed choices about where they choose to access advice about using cannabis during pregnancy.

Trial Registration: Open Science Framework 10.17605/OSF.IO/BW8DA; www.osf.io/6fb2e

International Registered Report Identifier (IRRID): PRR1-10.2196/34421

(JMIR Res Protoc 2022;11(3):e34421) doi:10.2196/34421

KEYWORDS

cannabis; pregnancy; health information; social media; Twitter

Introduction

Recreational cannabis use has increased in Canada since its legalization in 2018, including among pregnant women [1]. Reductions in the perceived harms of cannabis use may occur around legalization, and as a result, pregnant women or individuals may find the activity to be low risk [2]. Cannabis and its derivative products are often marketed online as safe [3]. Certain groups and dispensaries may even promote the use of cannabis products during pregnancy for their antinausea and antiemetic effects [4,5]. Expecting mothers may also use the drug to stimulate appetite or treat depression, motivated by the perception that cannabis is natural and thus preferable to prescription medications [6]. However, a growing body of research suggests that cannabis and derivative products during pregnancy may harm the developing fetus. Cannabinoids readily cross the placenta and interfere with the endogenous cannabinoid system, a cell-signaling network that assists in neurodevelopment [7]. Consequently, maternal cannabis use has been associated with fetal growth restriction, higher rates of childhood affective disorders, and a greater incidence of learning disability and autism spectrum disorders among offspring [8-10].

Pregnant patients increasingly seek medical and health advice on online platforms, especially for emerging topics like cannabis use [11,12]. Although medical professionals and research groups may use these avenues to promote research findings, other Twitter users may use social media to promote commercial interests, share anecdotal stories, or spread misinformation [13-15]. For example, a 2019 study by Ishida et al [16] found that those who primarily rely on social media for their health information were 31% more likely than others to endorse the claim that cannabis use during pregnancy is safe and 56% more likely to endorse any form of misinformation about cannabis.

Given the possible disconnect between online messaging and evidence-based research about the effects of cannabis use during pregnancy, there is the possibility that advice taken from social media could have inaccuracies that may affect the health of mothers and their babies. Here, we propose a systematic search of Twitter to quantify the volume and tone of posts on the forum related to cannabis use in pregnancy. Twitter is a global platform, and our findings may have relevance in Canada, the United States, and other jurisdictions where access and availability to cannabis are increasing due to legalization. We will assess regional correlations in these data to determine if changes in the legalization of nonmedical cannabis affect online messaging of its use during pregnancy in Canada and states in the United States that have legalized recreational cannabis.

Methods

Overview

With reference to Arksey and O'Malley's [17] framework for scoping reviews, we will synthesize publicly available posts from Twitter to determine how cannabis use during pregnancy is being discussed on the platform [17]. The steps, as outlined by this framework and adapted for a Twitter-based analysis, will be:

- Identifying the research question
- Identifying relevant Twitter posts
- Selecting eligible Twitter posts
- Charting the data
- Collating, summarizing, and reporting the results

Past research from Cavazos-Rehg et al [18] has identified Twitter as a good source for analyzing online discussions about cannabis use because of its popularity and acceptance of substance use disclosure. We will use this to model a novel scoping review approach to explore Twitter posts about cannabis use during pregnancy. We will report our findings following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [19].

Step 1: Identifying the Research Question

How is cannabis use during pregnancy discussed on Twitter regarding the volume, tone, content, and authors/users?

Step 2: Identifying Relevant Twitter Posts

Our search strategy will follow an iterative approach according to our population, concept, and context of interest (Textbox 1). We will first use Twitter's native search function to conduct a preliminary scan of English language tweets about cannabis use during pregnancy and assemble a list of commonly used keywords and hashtags based on our findings. We will then refine our list to capture the breadth of online discussion while excluding mimicker terms (eg, non–drug-related uses of the word "high"). Finally, our search strategy will include a list of terms for pregnancy combined with terms for cannabis to search the Twitter Application Programming Interface (API; Textbox 2), for example, (pregnancy OR pregnant OR prenatal) AND (cannabis OR weed OR pot OR marijuana), with the final search strategy to be developed following preliminary findings. We will use the Twitter API for Academic Research for data collection. We will perform a full archive search of all English language tweets containing the keywords of interest posted from January 2012, when Colorado became the first English-speaking jurisdiction to legalize cannabis, to December 2021 [20].

Textbox 1. Population, Concept, Context framework.

Population
Twitter posts containing information relevant to pregnancy or pregnant individuals
Concept
Discussion or mention of cannabis use in relation to pregnancy or the developing fetus
Context
All English language Twitter posts (tweets) made from January 2012 to December 2021. Geographical analyses will be restricted to Canada and states in the United States where recreational cannabis use is legal.

Textbox 2. List of keywords related to cannabis use in pregnancy used to search the Twitter Application Programming Interface.

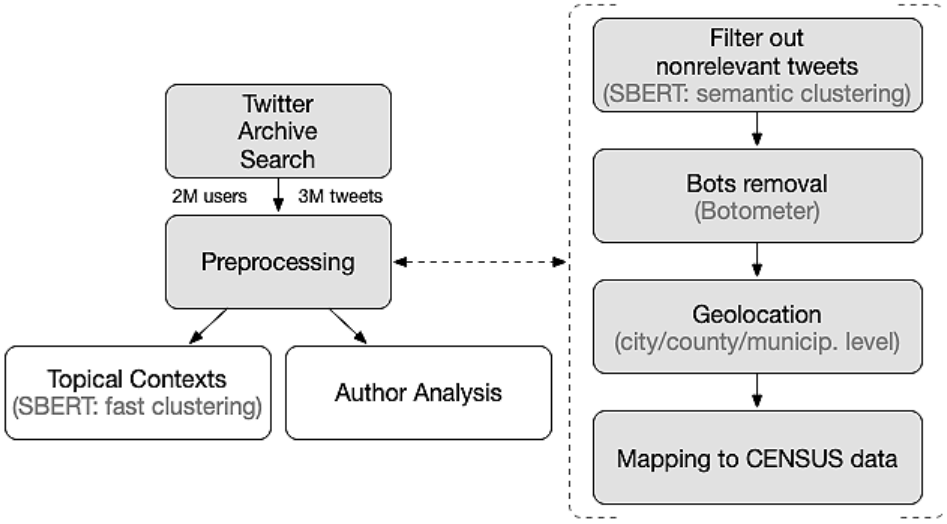
Pregnancy related
Pregnancy, pregnant, baby, fetus, fetal, prenatal, perinatal, womb, preggo, “pregnant life,” “baby bump,” “mom to be,” “mommy to be,” “baby on the way,” “preggers,” “pregnant af”
Cannabis related
cannabis, weed, pot, marijuana, marihuana, MJ, ganja, purp, bud, keef, kief, dope, “mary jane,” thc, cbd, cannamom, opiate, mdma, ecstasy, mmj, medical marijuana, blunt, bong, budder, hash, hemp, indica, kush, reefer, sativa

Step 3: Selecting Eligible Twitter Posts

Following the Twitter Archive search, we will preprocess the corpora to filter out content unrelated to *cannabis use during pregnancy*. Additionally, we will remove bot accounts [21], and tweets without geotags will be further analyzed to infer a location from their authors' profile [22,23] (Figure 1). We will filter out all tweets containing our keywords but that are unrelated to the consumption of cannabis during pregnancy via a *symmetric semantic search* using Sentence Bidirectional Encoder Representations from Transformers (BERT) [24]. This search assigns a score to each tweet for each given query (Textbox 3). The higher the score, the more semantically close the tweet is to the query. Tweets with a score lower than, for example, 0.6 for all queries are discarded since they are likely unrelated. The cutoff value of 0.6 was selected here for illustrative purposes. In the final analysis, we will tune this parameter and select the score that gives optimal classification results. We will also perform a topical context analysis to provide meaning and classify tweets by performing a *semantic community detection* using Sentence BERT [24]. We will use the "Fast clustering" algorithm together with "all-MiniLM-L6-v2" a pretrained sentence-transformer model for large-scale data sets [25]. In this model, we will set the minimum size of communities (or clusters) to 10 and a threshold similarity of 0.6. In other words, clusters will contain at least 10 tweets, and the similarity between tweets of the same cluster will be at least 60%. We will further classify related tweets into broad categories related to cannabis during pregnancy and medical cannabis or cannabis and youth, or legalization of cannabis. In addition, we will classify tweets related to cannabis during pregnancy into commercial, anecdotal/conversational/babble, misinformation, memes, and research studies.

Note, that most irrelevant tweets are pruned out by Sentence BERT in the preprocessing phase (Figure 1). We will evaluate the accuracy of this filtering by randomly sampling both types of tweets and label them as relevant or irrelevant by three independent reviewers and report precision and recall based on majority voting. Similarly, in the clustering phase, we will revise the inferred clusters and merge (if necessary) those that might be related to the same topical context.

Figure 1. Overview of the proposed data collection methodology, preprocessing, and analytical process for tweets about cannabis use during pregnancy. SBERT: Sentence Bidirectional Encoder Representations from Transformers.



Textbox 3. Queries passed to the symmetric semantic search of Sentence Bidirectional Encoder Representations from Transformers.

Queries

- Cannabis during pregnancy
- Kids, children, and youth smoking cannabis
- Smoking cannabis while pregnant
- Medical cannabis for people
- The effects of cannabis on pregnant women
- Legalization of cannabis
- Smoking or consuming drugs during pregnancy

Step 4: Charting the Data

Data charting will include an automated analysis of all tweets returned by our search. A manual analysis will then be conducted on the smaller subset of tweets included during the process outlined in Step 2.

Using the Twitter API for Academic Research [26], we will collect the timestamp of each returned tweet and analyze its text for sentiment (positive or negative) by integrating with the Natural Language Toolkit in Python and other techniques such as latent Dirichlet allocation [27], Sentence BERT [24], or recurrent neural networks [28]. We will also analyze the number and types of health effects mentioned in association with cannabis use in pregnancy and will extract location data when available from each tweet, either from geotagged tweets or from the location associated with the user's profile [29].

Three independent reviewers will manually review the smaller subset of randomly sampled tweets. We will verify the number of favorites and retweets each tweet has received against the automatic data collection via the API. We will use publicly available user lists to determine the category of organization or individual user that posted the tweet (government or public health agency, obstetrical society/network, university, hospital, news outlet, cannabis industry source, or other individual) [30], and we will manually (via majority voting from three reviewers) assign a category for organizations not appearing on the user lists. Finally, we will assess if the tweet mentions positive or adverse health effects on mothers or developing fetus/infants, and the specific health effects mentioned. For each tweet, data will be extracted by one reviewer and validated by a second reviewer. A third independent reviewer will resolve discrepancies if they arise.

Separately, we will also extract CENSUS or population-level data on birth rates and maternal and infant mortality rates across the study period in Canada and the United States. It has been shown that Twitter is a good proxy to infer health-related statistics, including teenage birth rates [31]. Thus, we want to verify whether certain geographical areas with certain CENSUS characteristics behave similarly with respect to their opinions on cannabis use during pregnancy. These vital statistics data will be sourced from Statistics Canada and the Centers for Disease Control and Prevention in the United States [32,33].

Step 5: Collating, Summarizing, and Reporting the Results

We will first report the total number of tweets returned over the search period and temporal trends in the number of tweets posted over the study period. Next, the number of tweets sampled in the automated and manual analyses will be reported. From the automated analysis, we will report the number and percentage of the returned posts that discuss cannabis use during pregnancy positively or negatively as determined by our sentiment analysis. Subsequently, we will calculate the standardized mean difference in the number of favorites and retweets received by positive and negative tweets, and to compute the odds (ratio) that positive posts originate from each category of organization or individual and mention health effects. We will further calculate the number of times each health effect was mentioned as a percentage of the total health effect mentions. These statistics will be presented in tabular form.

The location-based component of our analysis will be restricted to tweets that offer location data and originate from Canada and legal states within the United States, as these are the only English-speaking regions that have legalized the sale of nonmedical cannabis. If any regions (eg, New Zealand or the United Kingdom) legalize cannabis before our analysis is conducted, this restriction will be changed to include them. We will match location data from these jurisdictions to the timestamp for each tweet to calculate the proportion of tweets originating from our predefined geographical regions for each week of the search period. Next, we will visualize each region on a line graph that plots time versus the volume of posts with a marker to indicate when that region legalized cannabis. A line graph that plots time versus percentage of positive posts will be plotted using the same process. We will then use a repeated cross-sectional design to analyze the correlation of these data with population-level vital statistics data and determine if trends in cannabis messaging on Twitter correlate with birth rates and maternal and infant mortality rates.

In addition to these numerical analyses, we will develop qualitative profiles of influential accounts. These profiles will include elements such as the user's background (eg, political leaning, socioeconomic status, or education/interests); their Twitter following; whether Twitter has verified their account as "authentic, notable, and active" [34]; and how they contribute to the discussion about cannabis use during pregnancy on the platform. Comparisons and contrasts will be drawn between the typical supportive and opposing posters based on these elements.

Ethics and Dissemination

This study was exempted from ethics review on the basis that it will collect and synthesize publicly available data. Therefore, the research does not require ethical approval.

Results

Using our data collection method, combing the search_all_tweets function from Tweepy [35] together with the Twitter API for Academic Research, we collected 2,000,000 tweets and 1,000,000 retweets that are potentially related to cannabis use during pregnancy. These results cover all English language tweets posted from January 1, 2012, to December 31, 2021 (10 years), that include both *pregnancy*- and *cannabis*-related keywords. Of the 3,000,000 unique tweets, only 4.3% of them are geotagged (Figure 2). Note that these tweets are concentrated mainly in English-speaking cities or countries. This finding is expected since our search explicitly requested English tweets. Figure 3 shows the frequency distribution of all 3,000,000 tweets per day since 2012. Colorado was the first English-speaking jurisdiction to legalize cannabis in 2012, and Canada legalized cannabis in 2018. Our Twitter search includes 47 distinct keywords; we plotted the number of times each keyword appears in our corpora (Table 1).

Figure 2. Geographic distribution of geotagged tweets containing pregnancy and cannabis-related keywords posted between January 1, 2012, to December 31, 2021.

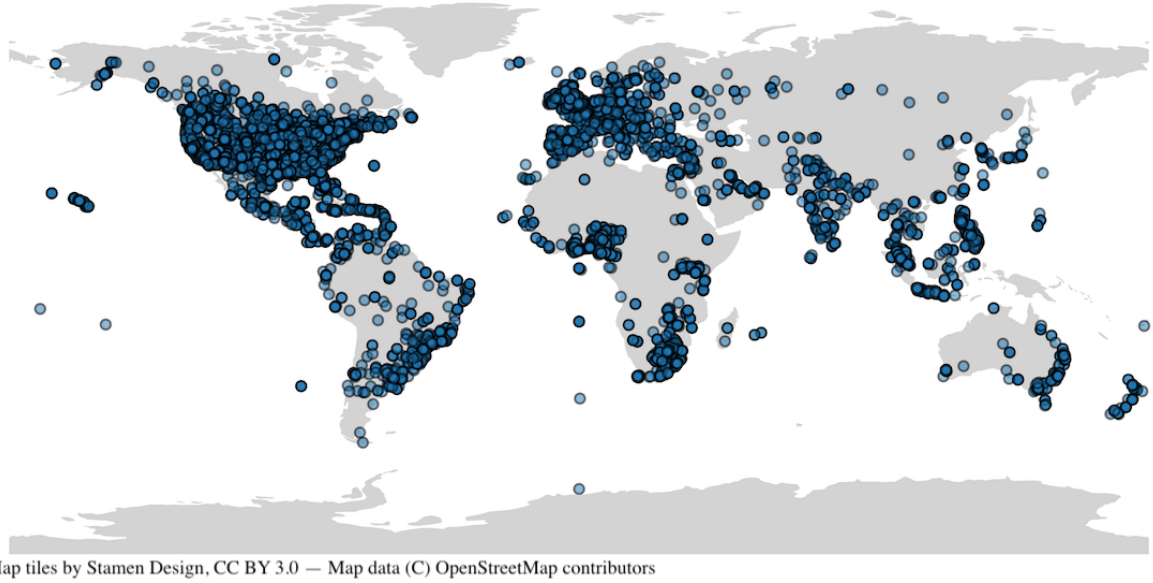


Figure 3. Number of tweets per day related to cannabis in pregnancy, January 1, 2012, to December 31, 2021.

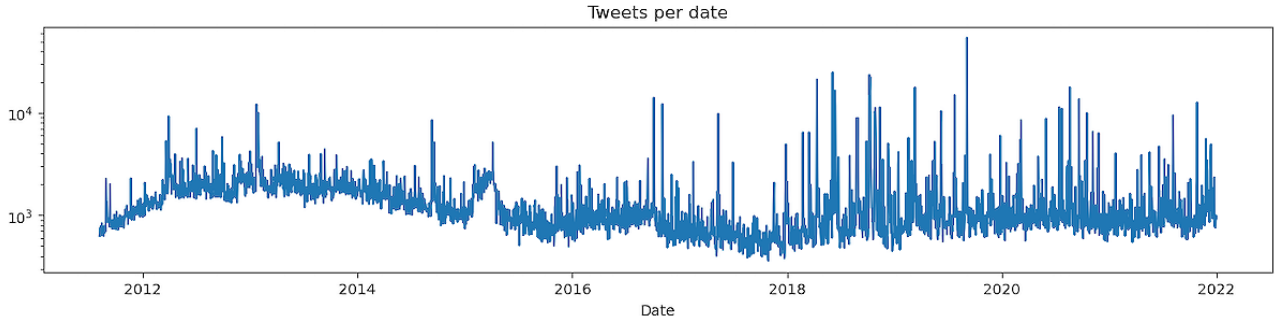


Table 1. Frequency of cannabis-related keywords identified in tweets posted between January 1, 2012, to December 31, 2021.

Keyword	Count
weed	1,047,115
dope	688,153
blunt	556,865
pot	399,444
keef	356,605
marijuana	183,409
bud	161,328
bong	116,876
kush	99,916
thc	44,970
hash	44,906
cbd	39,287
ecstasy	33,989
hemp	28,514
purp	25,353
ganja	24,641
indica	8447
reefer	6125
opiate	4102
kief	3092
mdma	2459
mmj	1386
budder	643
marihuana	637
cannamom	40
medicalmarijuana	13

The semantic community detection algorithm detected 220 clusters within the 3,000,000 tweets from our corpora. We manually inspected the top 5 and bottom 5 tweets of each cluster and assigned an appropriate label that best described the topical context of those tweets. For example, we found 9 topical clusters related to *cannabis use during pregnancy* (Figure 4). A sample of paraphrased tweets from one identified cluster, "Cannabis exposure on infants," is shown (Table 2).

We expect to conclude this study in December 2022.

Figure 4. Topical contexts (clusters) identified from tweets collected about cannabis use during pregnancy.

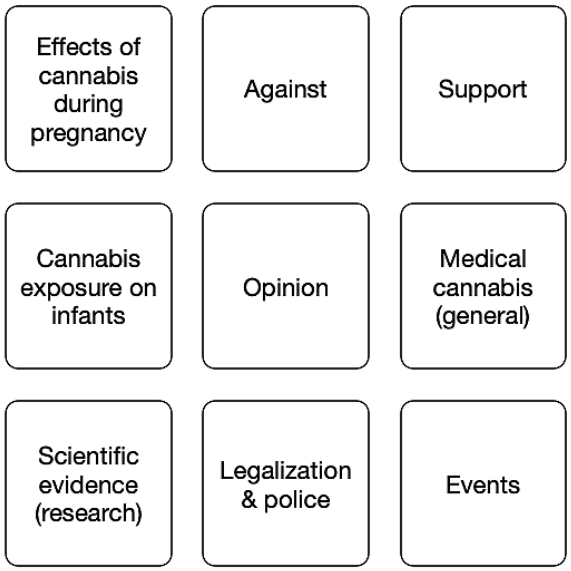


Table 2. Top 3 and bottom 3 tweets selected from the cluster "Cannabis exposure on infants."[a]

No.	Paraphrased tweet
1	random *thc* found in *baby* soap
2	newborns test positive *marijuana* from *baby* soap
3	*marijuana* in newborns from *baby* soap are false positives.
45	*Baby* you only do *thc*, you need help.
46	pediatric doctor advises passing *thc* via placenta and breast feeding (previously thought to damage *baby* brain)
47	expert on *thc* exposure during pregnancy is Dr. X at Clinic Y.

[a]Italicized words represent our set of query keywords.

Discussion

Preliminary Findings

This study will infer how cannabis use in pregnancy is portrayed on Twitter, the content and origin of supportive posts, and how legal status changes influence the volume and tone of posts related to cannabis in pregnancy. Our findings will help inform policy strategies to public health agencies, care providers, and other stakeholders. Moreover, they will suggest future avenues for research. Our preliminary findings suggest that this work is feasible and that we have identified a sufficiently robust corpus of tweets for more detailed analyses.

Limitations and Future Work

Twitter is an extensive online platform to share news and opinions [36]. However, it is not representative of the whole population [37]. A 2016 survey found that only 21% of Americans use Twitter [38]. Users are, on average, younger and better educated than nonusers, and they are more liberal and pay more attention to politics [37]. However, a recent study [39] has shown that young adults (25-44 years) that were active on an abortion debate on Twitter were well represented compared to the 2017 CENSUS representation in Chile. While this age range overlaps with the women's reproductive age (15-44 years), birth rates decreased for females aged 15 to 34 years, increased for females aged 35 to 44 years, and were unchanged for females aged 10 to 14 years and 45 to 49 years from 2018 to 2019 in the United States [40].

Besides Twitter, there are several online platforms used to share opinions, for instance, Facebook, Reddit, and Quora. To the best of our knowledge, only Facebook has been used to study people's opinions on cannabis [41] and during pregnancy [42]. However, in these studies, authors run surveys by targeting people via Facebook ads (ie, findings are based on answers to questionaries) and did not analyze free-text opinions. Here, we opt to use Twitter data since it has been shown that there is rich content to study health-related issues [20,43], including opinions on the use of cannabis during pregnancy [44-46]. Besides, Twitter is one of the largest social media platforms allowing discussions and debates with 187,000,000 daily users [47]. Future research may focus on other platforms to study how people discuss cannabis use during pregnancy and verify whether all these users combined can make a better representation of their offline population.

Conclusions

We will submit the final results of our review for publication in a peer-reviewed journal, present at academic conferences,

and share through publicly available streams such as the professional and institutional social media accounts and webpages associated with the research team. The results will provide insight into how frequently and in what context Twitter is being used to discuss cannabis use during pregnancy. We anticipate that this knowledge will help public health agencies and health care providers assess the messaging patients may be receiving on Twitter and develop communication strategies to counter misinformation, especially in geographical regions where legalization is recent or imminent. Most importantly, we foresee that our findings will assist expecting families in making informed choices about where they choose to access advice about using cannabis during pregnancy.

Acknowledgments

This work was supported by a Canadian Institutes of Health Research Team Grant awarded to DJC (funding reference CA3-170126). The authors would like to thank Roberto Ulloca for support with the Twitter Academic Application Programming Interface and Indira Sen for advice on the state-of-the-art natural language processing techniques.

Authors' Contributions

LC, LEN, MSQM, MK, and DJC conceptualized the study and designed the methodology. LC wrote the original draft. LC, LEN, MSQM, SR, MCW, MK, and DJC reviewed and edited the manuscript. DJC and MSQM supervised the study and acquired the funding. All authors read and approved the final version of this manuscript.

Conflicts of Interest

None declared.

References

1. Lowry DE, Corsi DJ. Trends and correlates of cannabis use in Canada: a repeated cross-sectional analysis of national surveys from 2004 to 2017. CMAJ Open 2020;8(3):E487-E495 [FREE Full text] [doi: 10.9778/cmajo.20190229] [Medline: 32737023]
2. Mark K, Gryczynski J, Axenfeld E, Schwartz RP, Terplan M. Pregnant women's current and intended cannabis use in relation to their views toward legalization and knowledge of potential harm. J Addict Med 2017;11(3):211-216. [doi: 10.1097/ADM.0000000000000299] [Medline: 28252456]
3. Fattore L, Fratta W. Beyond THC: the new generation of cannabinoid designer drugs. Front Behav Neurosci 2011;5:60. [doi: 10.3389/fnbeh.2011.00060] [Medline: 22007163]
4. Cavazos-Rehg PA, Krauss MJ, Cahn E, Lee KE, Ferguson E, Rajbhandari B, et al. Marijuana promotion online: an investigation of dispensary practices. Prev Sci 2019 Feb;20(2):280-290 [FREE Full text] [doi: 10.1007/s11121-018-0889-2] [Medline: 29629505]
5. Westfall RE, Janssen PA, Lucas P, Capler R. Survey of medicinal cannabis use among childbearing women: patterns of its use in pregnancy and retroactive self-assessment of its efficacy against 'morning sickness'. Complement Ther Clin Pract 2006 Feb;12(1):27-33. [doi: 10.1016/j.ctcp.2005.09.006] [Medline: 16401527]
6. Chang JC, Tarr JA, Holland CL, De Genna NM, Richardson GA, Rodriguez KL, et al. Beliefs and attitudes regarding prenatal marijuana use: perspectives of pregnant women who report use. Drug Alcohol Depend 2019 Mar 01;196:14-20 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.11.028] [Medline: 30658220]
7. Richardson KA, Hester AK, McLemore GL. Prenatal cannabis exposure - the "first hit" to the endocannabinoid system. Neurotoxicol Teratol 2016;58:5-14. [doi: 10.1016/j.ntt.2016.08.003] [Medline: 27567698]
8. Corsi DJ, Donelle J, Sucha E, Hawken S, Hsu H, El-Chaâr D, et al. Maternal cannabis use in pregnancy and child neurodevelopmental outcomes. Nat Med 2020 Oct;26(10):1536-1540. [doi: 10.1038/s41591-020-1002-5] [Medline: 32778828]
9. El Marroun H, Tiemeier H, Steegers EAP, Jaddoe VWV, Hofman A, Verhulst FC, et al. Intrauterine cannabis exposure affects fetal growth trajectories: the Generation R Study. J Am Acad Child Adolesc Psychiatry 2009 Dec;48(12):1173-1181. [doi: 10.1097/CHI.0b013e3181bfa8ee] [Medline: 19858757]
10. Gray KA, Day NL, Leech S, Richardson GA. Prenatal marijuana exposure: effect on child depressive symptoms at ten years of age. Neurotoxicol Teratol 2005;27(3):439-448. [doi: 10.1016/j.ntt.2005.03.010] [Medline: 15869861]
11. Althouse BM, Scarpino SV, Meyers LA, Ayers JW, Bargsten M, Baumbach J, et al. Enhancing disease surveillance with novel data streams: challenges and opportunities. EPJ Data Sci 2015;4:17 [FREE Full text] [doi: 10.1140/epjds/s13688-015-0054-0] [Medline: 27990325]
12. Meacham MC, Paul MJ, Ramo DE. Understanding emerging forms of cannabis use through an online cannabis community: an analysis of relative post volume and subjective highness ratings. Drug Alcohol Depend 2018 Jul 01;188:364-369 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.03.041] [Medline: 29883950]
13. Allem J, Escobedo P, Dharmapuri L. Cannabis surveillance with Twitter data: emerging topics and social bots. Am J Public Health 2020 Mar;110(3):357-362. [doi: 10.2105/AJPH.2019.305461] [Medline: 31855475]

14. Peiper NC, Baumgartner PM, Chew RF, Hsieh YP, Bieler GS, Bobashev GV, et al. Patterns of Twitter behavior among networks of cannabis dispensaries in California. J Med Internet Res 2017 Jul 04;19(7):e236 [FREE Full text] [doi: 10.2196/jmir.7137] [Medline: 28676471]
15. Tapper EB, Mirabella R, Walicki JJ, Banales JM. Optimizing the use of twitter for research dissemination: the "Three Facts and a Story" randomized-controlled trial. J Hepatol 2021 Aug;75(2):271-274. [doi: 10.1016/j.jhep.2021.05.020] [Medline: 34062178]
16. Ishida JH, Zhang AJ, Steigerwald S, Cohen BE, Vali M, Keyhani S. Sources of information and beliefs about the health effects of marijuana. J Gen Intern Med 2020 Jan;35(1):153-159 [FREE Full text] [doi: 10.1007/s11606-019-05335-6] [Medline: 31637640]
17. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
18. Cavazos-Rehg PA, Sowles SJ, Krauss MJ, Agbonavbare V, Grucza R, Bierut L. A content analysis of tweets about high-potency marijuana. Drug Alcohol Depend 2016 Sep 01;166:100-108 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.06.034] [Medline: 27402550]
19. 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 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
20. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. Am J Public Health 2017 Jan;107(1):e1-e8. [doi: 10.2105/AJPH.2016.303512] [Medline: 27854532]
21. Sayyadiharikandeh M, Varol O, Yang KC, Flammini A, Menczer F. Detection of novel social bots by ensembles of specialized classifiers. In: Proceedings of the 29th ACM International Conference on Information & Knowledge Management. 2020 Oct Presented at: CIKM '20; October 19-23, 2020; Virtual Event. [doi: 10.1145/3340531.3412698]
22. Ajao O, Hong J, Liu W. A survey of location inference techniques on Twitter. J Inf Sci 2015 Nov 20;41(6):855-864. [doi: 10.1177/0165551515602847]
23. Mahmud J, Nichols J, Drews C. Where is this tweet from? inferring home locations of twitter users. 2012 Presented at: The Sixth International AAAI Conference on Weblogs and Social Media; June 4-8, 2012; Trinity College in Dublin, Ireland.
24. Reimers N, Gurevych I. Sentence-BERT: sentence embeddings using siamese BERT-Networks. arXiv. Preprint posted online on August 27, 2019. [doi: 10.18653/v1/d19-1410]
25. Pretrained models. Sentence-Transformers. URL: https://www.sbert.net/docs/pretrained_models.html [accessed 2021-12-31]
26. Academic Research access. Twitter Developer Platform. URL: https://developer.twitter.com/en/products/twitter-api/academic-research [accessed 2021-08-09]
27. Hoffman M, Bach F, Blei D. Online learning for latent dirichlet allocation. Adv Neural Inf Processing Syst 2010;23:1-9.
28. Nistor SC, Moca M, Moldovan D, Oprean DB, Nistor RL. Building a Twitter sentiment analysis system with recurrent neural networks. Sensors (Basel) 2021 Mar 24;21(7):2266 [FREE Full text] [doi: 10.3390/s21072266] [Medline: 33804900]
29. Schlosser S, Toninelli D, Cameletti M. Comparing methods to collect and geolocate tweets in Great Britain. J Open Innovation Technol Market Complexity 2021 Jan 25;7(1):44. [doi: 10.3390/joitmc7010044]
30. van Draanen J, Krishna T, Tsang C, Liu S. Keeping up with the times: how national public health and governmental organizations communicate about cannabis on Twitter. Subst Abuse Treat Prev Policy 2019 Sep 12;14(1):38 [FREE Full text] [doi: 10.1186/s13011-019-0224-3] [Medline: 31511026]
31. Culotta A. Estimating county health statistics with twitter. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Apr Presented at: CHI '14; April 26-May 1, 2014; Toronto, ON p. 1335-1344. [doi: 10.1145/2556288.2557139]
32. National Vital Statistics System. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nvss/index.htm [accessed 2021-08-09]
33. Statistics Canada, Canadian Vital Statistics - Birth database (CVSB). Statistics Canada. URL: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3231 [accessed 2021-08-09]
34. About verified accounts. Twitter Help Center. URL: https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts [accessed 2021-08-09]
35. Roesslein J. tweepy / tweepy. GitHub. 2020. URL: https://github.com/tweepy/tweepy [accessed 2022-03-13]
36. Rosenstiel T, Sonderman J, Loker K, Ivancin M, Kjarval N. Twitter and the news: how people use the social network to learn about the world. American Press Institute. 2015. URL: https://www.americanpressinstitute.org/wp-content/uploads/2015/09/Twitter-and-News-How-people-use-Twitter-to-get-news-American-Press-Institute.pdf [accessed 2022-03-15]
37. Mellon J, Prosser C. Twitter and Facebook are not representative of the general population: political attitudes and demographics of British social media users. Res Polit 2017 Jul 13;4(3):205316801772000. [doi: 10.1177/2053168017720008]
38. Greenwood S, Perrin A, Duggan M. Social media update 2016. Downtown Dubuque, Iowa. 2016. URL: http://downtowndubuque.org/wp-content/uploads/2017/01/Social-Media-Update-2016.pdf [accessed 2022-03-15]
39. Graells-Garrido E, Baeza-Yates R, Lalmas M. How representative is an abortion debate on Twitter? In: Proceedings of the 10th ACM Conference on Web Science. 2019 Jun Presented at: WebSci '19; June 30-July 3, 2019; Boston, MA p. 133-134. [doi: 10.1145/3292522.3326057]

40. Martin JA, Hamilton BE, Osterman MJK, Driscoll AK. Births: final data for 2019. Natl Vital Stat Rep 2021 Apr;70(2):1-51. [Medline: 33814033]
41. Borodovsky JT, Marsch LA, Budney AJ. Studying cannabis use behaviors with Facebook and web surveys: methods and insights. JMIR Public Health Surveill 2018 May 02;4(2):e48 [FREE Full text] [doi: 10.2196/publichealth.9408] [Medline: 29720366]
42. Pike CK, Sofis MJ, Budney AJ. Correlates of continued cannabis use during pregnancy. Drug Alcohol Depend 2021 Oct 01;227:108939. [doi: 10.1016/j.drugalcdep.2021.108939] [Medline: 34358772]
43. Berry N, Lobban F, Belousov M, Emsley R, Nenadic G, Bucci S. #WhyWeTweetMH: understanding why people use Twitter to discuss mental health problems. J Med Internet Res 2017 Apr 05;19(4):e107 [FREE Full text] [doi: 10.2196/jmir.6173] [Medline: 28381392]
44. Dakkak H, Brown R, Twynstra J, Charbonneau K, Seabrook J. The perception of pre- and post-natal marijuana exposure on health outcomes: a content analysis of Twitter messages. J Neonatal Perinatal Med 2018;11(4):409-415. [doi: 10.3233/NPM-17133] [Medline: 29843262]
45. Pang RD, Dormanesh A, Hoang Y, Chu M, Allem J. Twitter posts about cannabis use during pregnancy and postpartum: a content analysis. Subst Use Misuse 2021;56(7):1074-1077. [doi: 10.1080/10826084.2021.1906277] [Medline: 33821757]
46. Pradyumn M. Systematic review of literature using Twitter as a tool. The ScholarShip at ECU. 2018 Jul. URL: https://thescholarship.ecu.edu/bitstream/handle/10342/6950/PRADYUMN-MASTERSTHESIS-2018.pdf?sequence=1 [accessed 2022-03-15]
47. Kastrenakes J. Reddit reveals daily active user count for the first time: 52 million. The Verge. URL: https://www.theverge.com/2020/12/1/21754984/reddit-dau-daily-users-revealed [accessed 2022-03-13]

Abbreviations

API: Application Programming Interface
BERT: Bidirectional Encoder Representations from Transformers
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by T Leung; submitted 22.10.21; peer-reviewed by T Ndabu, M Waring, JP Allem; comments to author 17.12.21; revised version received 26.01.22; accepted 08.02.22; published 29.03.22.

Please cite as:
Cresswell L, Espin-Noboa L, Murphy MSQ, Ramlawi S, Walker MC, Karsai M, Corsi DJ
The Volume and Tone of Twitter Posts About Cannabis Use During Pregnancy: Protocol for a Scoping Review
JMIR Res Protoc 2022;11(3):e34421
URL: https://www.researchprotocols.org/2022/3/e34421
doi:10.2196/34421
PMID:35348465

Protocol

Mapping Digital Public Health Interventions Among Existing Digital Technologies and Internet-Based Interventions to Maintain and Improve Population Health in Practice: Protocol for a Scoping Review

Laura Maaß[1,2], BA, MA; Chen-Chia Pan[2,3], BSc, MA; Merle Freye[2,4], First State Examination

[1]Department of Health, Long-Term Care and Pensions, Research Center on Inequality and Social Policy, Bremen, Germany

[2]Leibniz ScienceCampus Digital Public Health Bremen, Bremen, Germany

[3]Department for Epidemiological Methods and Etiological Research, Leibniz Institute for Prevention Research and Epidemiology, Bremen, Germany

[4]Institute for Information, Health and Medical Law, Bremen, Germany

Corresponding Author:
Laura Maaß, BA, MA
Department of Health, Long-Term Care and Pensions
Research Center on Inequality and Social Policy
Mary-Somerville-Straße 3
Bremen, 28359
Germany
Phone: 49 421 218 58610
Email: laura.maass@uni-bremen.de

Abstract

Background: Rapid developments and implementation of digital technologies in public health domains throughout the last decades have changed the landscape of health delivery and disease prevention globally. A growing number of countries are introducing interventions such as online consultations, electronic health records, or telemedicine to their health systems to improve their populations' health and improve access to health care. Despite multiple definitions for digital public health and the development of different digital interventions, no study has analyzed whether the utilized technologies fit the definition or the core characteristics of digital public health interventions. A scoping review is therefore needed to explore the extent of the literature on this topic.

Objective: The main aim of this scoping review is to outline real-world digital public health interventions on all levels of health care, prevention, and health. The second objective will be the mapping of reported intervention characteristics. These will include nontechnical elements and the technical features of an intervention.

Methods: We searched for relevant literature in the following databases: PubMed, Web of Science, CENTRAL (Cochrane Central Register of Controlled Trials), IEEE (Institute of Electrical and Electronics Engineers) Xplore, and the Association for Computing Machinery (ACM) Full-Text Collection. All original study types (observational studies, experimental trials, qualitative studies, and health-economic analyses), as well as governmental reports, books, book chapters, or peer-reviewed full-text conference papers were included when the evaluation and description of a digital health intervention was the primary intervention component. Two authors screened the articles independently in three stages (title, abstract, and full text). Two independent authors will also perform the data charting. We will report our results following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.

Results: An additional systematic search in IEEE Xplore and ACM, performed on December 1, 2021, identified another 491 titles. We identified a total of 13,869 papers after deduplication. As of March 2022, the abstract screening state is complete, and we are in the state of screening the 1417 selected full texts for final inclusion. We estimate completing the review in April 2022.

Conclusions: To our knowledge, this will be the first scoping review to fill the theoretical definitions of digital public health with concrete interventions and their characteristics. Our scoping review will display the landscape of worldwide existing digital public health interventions that use information and communication technologies. The results of this review will be published in a peer-reviewed journal in early 2022, which can serve as a blueprint for the development of future digital public health interventions.

International Registered Report Identifier (IRRID): DERR1-10.2196/33404

(JMIR Res Protoc 2022;11(3):e33404) doi:10.2196/33404

KEYWORDS

digital public health; telemedicine; electronic health records; ePrescription; eReferral; eConsultation; eSurveillance; eVaccination registries; scoping review; protocol

Introduction

Background

The potential of digital technology for improving the health of individuals, communities, and populations is unprecedented. Technological advancements empower individuals to engage in self-management and well-being [1]. There is also the unparalleled opportunity of digital technologies to transform the prevention [2], health promotion [3,4], health monitoring [5], health management [6,7], health equity [8], and surveillance for public health disasters [9-11].

Digital technologies for health are intrinsically interdisciplinary, including computer science, engineering, information science, clinical medicine, epidemiology, and public health [12,13]. Although these disciplines are involved at various stages, from the development process to the implementation of digital technologies, a shared understanding of key terms within the field of digitalization in health is still missing. More importantly, to develop, implement, integrate, and evaluate needs-based digital public health interventions (DiPHIs), a clear and mutual understanding of the specific properties of digital health technologies for public health purposes is required [14].

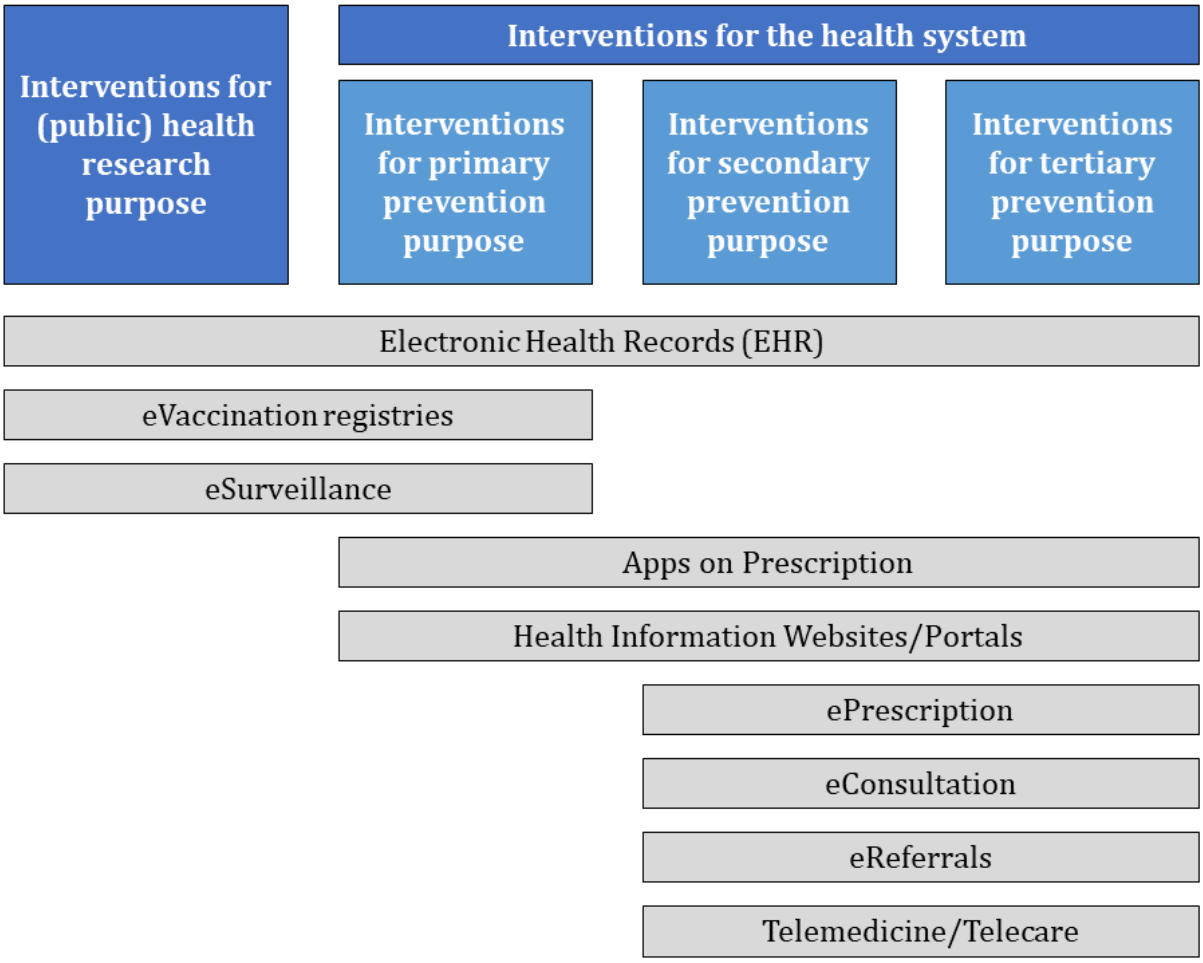
Definition of Digital Public Health

For this review, we will define digital public health following the understanding of the European Public Health Association as the use of digital technologies or tools to achieve public health goals. Therefore, digital public health is not a new discipline but rather represents the digitalization of public health [15]. Following the definition of Winslow [16], public health aims at "preventing disease, prolonging life, and promoting physical [and mental] health and efficiency through organized community efforts." These include health education, organizing health care (both medical and nursing services), or controlling infections on a community level (or on a global level as evident during the COVID-19 pandemic) from an individual to a community, national, or international level [16,17]. Although the definition provided by Winslow is over 100 years old, it is still the most referred to definition of public health. For instance, the World Health Organization still uses this definition with respect to essential public health functions. These functions place public health as the primary discipline in health governance (eg, planning of health services), financing health interventions and services, health information systems (eg, surveillance systems), health communication, universal health coverage, health education, or health regulations to protect and ensure the health of vulnerable groups [18].

This importance of the given definition of public health makes it suitable for our understanding of digital public health as the digitalization of Winslow's definition. According to the ongoing importance of Winslow's description in essential public health functions, we define DiPHIs as interventions that address "at least one essential public health function through digital means" [19]. Similar to public health interventions, DiPHIs aim to strengthen the population's health from the individual to the community and national levels. To achieve this, DiPHIs, similar to interventions for digital health, eHealth, or mobile health (mHealth), use information and communication technologies (ICTs) [12,20]. ICTs include the use of radio, television, smartphones, hardware, and software for computers and satellites for communication purposes. eHealth is defined as the use of ICT for health purposes, whereas the focus lies on delivering health services and not on health promotion itself [21,22]. mHealth is understood as using wireless and mobile technology to improve health [21]. The main difference between digital public health and digital health (including eHealth and mHealth) is that the former solely targets individuals' health, whereas digital public health targets groups of people and communities [14]. With the combination of public health goals (defined as the above-mentioned essential public health functions [18]) and ICT application, DiPHIs can monitor public health outcomes and disasters (as seen in tracing apps for the COVID-19 pandemic) [23]. In the best case, DiPHIs follow an evidence- and needs-based approach with a participatory user–targeted development design to improve the acceptance of the intervention within the population [12,14,19,24,25].

We expect that the described interventions of our finally chosen papers will target one of the health areas listed in Figure 1. Nevertheless, we will likely identify other DiPHIs throughout the review process and update this first proposed landscape of interventions in digital public health.

Figure 1. The proposed landscape of digital public health intervention classification.



Study Aim and Objective

This scoping review aims to serve as a blueprint for future DiPHIs to support countries in adapting digital public health to their health system. To fulfill this aim, the main objective is to outline real-world DiPHIs on all levels of health care (primary, secondary, and tertiary health care), prevention, and health promotion based on our predefined definition of a DiPHI. The second objective will be to map reported intervention characteristics of existing real-world DiPHIs.

For this scoping review, intervention characteristics will include the nontechnical elements (eg, the target group or the addressed level of health care) and the technical features of an intervention (eg, data exchange between health providers). Our review will not provide a detailed analysis of the cost-effectiveness of DiPHIs, their influence on health outcomes, the facilitators or barriers for implementation, or adaptation of interventions, as DiPHIs overall are too heterogeneous to be summarized within one literature review [26,27].

Methods

Inclusion Criteria

Overview

The inclusion criteria follow the PIOS (Participants, Intervention, Outcome, and Study design) format. The inclusion and exclusion criteria are summarized in Table 1 and described in detail below.

Table 1. Inclusion and exclusion criteria of the scoping review.

Level	Inclusion criteria	Exclusion criteria
Population	The study focuses on the geographical community level or above (regional or national)	The study population consists of veterans, armed forces, prisoners, inmates, refugees, or asylum seekers
Intervention	The paper describes a concrete DiPHI[a]; the DiPHI is paid or reimbursed by the government or health insurance; the intervention uses the internet and/or Bluetooth to facilitate health care, allows communication between providers or providers and patients, promotes its users' health, reuses the collected data for public health research, or enables digital surveillance of public health disasters	The paper offers a framework or overview of an intervention type but does not describe a concrete intervention in detail; the intervention does not use the internet, does not address public health functions, or needs to be privately bought by the user without reimbursement by the government or health insurance (ie, interventions for the private market); the intervention uses SMS text messaging or regular phone calls; the intervention focuses on background management processes; the described intervention does not match our definition of digital public health; the intervention targets only the individual user but not a group of people; the digital public health intervention is not the central research object of the publication
Study design	All original peer-reviewed studies, reports, books, book chapters, or peer-reviewed conference papers that have a description of a DiPHI as their primary intervention component	Study protocols, editorials, commentaries, conference proceedings, or reviews (narrative reviews, scoping reviews, systematic reviews, or meta-analyses); not peer-reviewed conference papers or original studies
Accessibility	The paper is available on the internet or after contact with the authors	The full text is not available on the internet or after contact with the authors
Language	The paper is published in English, Chinese, or German	The paper is published in a language other than English, Chinese, or German

[a]DiPHI: digital public health intervention.

Study Populations/Participants

We will extract data from all studies that focus on the geographical community level and above (eg, regional or national). As public health concentrates on the population level and not the merely individual level, we will exclude all case studies that focus exclusively on single institutions or departments (eg, single hospitals or emergency departments) but do not address at least the community level.

Our scoping review will aim to describe DiPHIs and online technologies for the general public (ordinary citizens of a city/state country that form a society). We will include papers addressing health care workers (eg, physicians, nurses, or therapists) and people with access to health care and health insurance (both statutory and private health insurance) without further regulations or restrictions. In contrast, we will exclude groups with limited or special access to public health care (ie, having their own health care system). This applies to the following three groups with precisely regulated access to health care: (1) veterans and the armed forces (who receive treatment in many countries within the military health system, often paid by the ministry of defense), (2) prisons and inmates (for which a prison physician needs to approve the treatment by another physician outside the prison), and (3) refugees and asylum seekers (who are not health insured in most countries during the procedure for granting the right of asylum and therefore often only have limited access to health care).

Our review will include all participants in terms of age, gender, ethnicity, morbidities, education level, staff role, and occupation. We will include studies that examined specific interventions among health care providers (eg, physicians, nurses, or therapists) and studies that analyzed the use of interventions by laypersons (people with a profession outside medicine or health care or with no profession).

Interventions/Technologies

Following our definition of a DiPHI, we will include studies describing digital health interventions that use the internet (eg, mobile devices such as smartphones, sensors, or wearables with WiFi; computer-based solutions that use cellular services; cloud systems to store/allocate health data; or wireless medical devices) and/or Bluetooth (eg, mobile devices such as smartphones, sensors, or wearables) to address at least one essential public health function to (1) facilitate health care, (2) allow for communication between providers or between providers and patients, (3) promote one's health, (4) collect data in a way that enables their secondary use in public health research, or (5) allow for digital surveillance of public health disasters.

We will exclude every digital health intervention that does not address one of the five inclusion criteria listed above as they are not deemed to represent digital public health based on our definition.

To be included, studies should not exclusively focus on individuals (as is the case for digital health interventions) but should also focus on groups of people (eg, communities). They will also have to report on the DiPHI as the main object of research and offer a description of the intervention/technology. Such examples include studies explaining the implementation of a national digital surveillance system in a country or the social acceptance of online consultations in a public health system.

Studies that use a DiPHI only for secondary data analysis but do not have the DiPHI as the main object of research will also

be omitted. Examples of such papers are studies mainly interested in measuring obesity in a community and only use electronic health records (EHRs) as a data source without going into detail on the records themselves. Therefore, the EHR in such study types would not be seen as an intervention for the sake of this scoping review. We will also exclude studies that offer no precise details on the modules of the intervention or its implementation process, such as studies that only give a brief overview of telemedicine opportunities in general without a concrete adaptation case. We will further exclude studies that evaluate or present interventions for the private market, for which individuals will have to buy without the option of a refund from their public health system (eg, privately bought wearables or apps that are not provided/refunded by a public health institution, insurance, or the government). Interventions that do not use the internet or Bluetooth (eg, a regular phone call for remote counseling or SMS as text reminder interventions) will be excluded. Technologies that focus solely on background management processes (eg, hospital management systems) are not considered to be DiPHIs and are therefore not included.

Outcome Measures

To answer our research questions, we will map the description of DiPHIs and their characteristics presented in the selected papers. These elements will be clustered to form an overview of specific modules for DiPHIs.

Study Designs

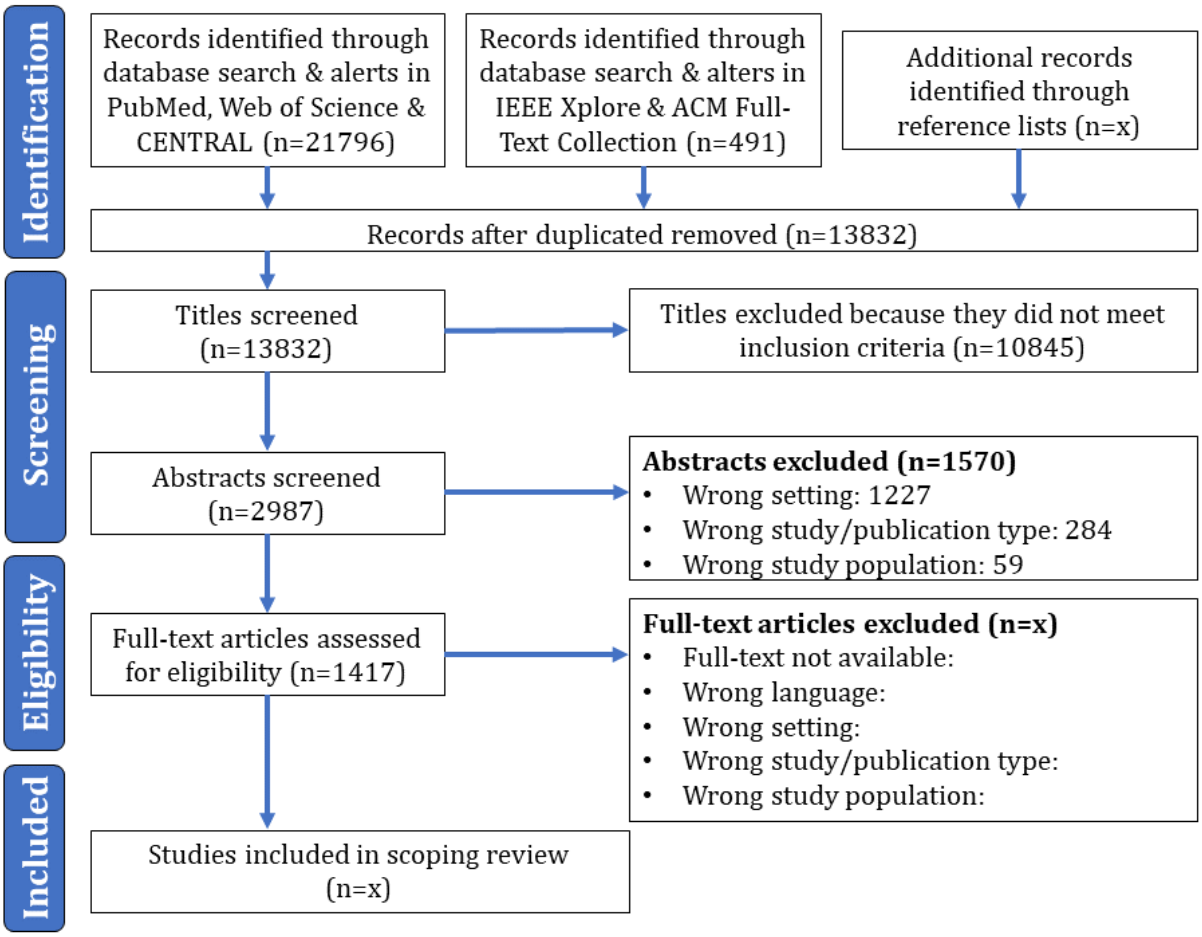
The scoping review will include all suitable published papers to gather information on all digital tools classified as DiPHIs. This includes all peer-reviewed original studies (observational studies, experimental trials, qualitative studies, and health-economic analyses), governmental reports, books, book chapters, or peer-reviewed conference papers that have the evaluation and description of a digital health intervention as their primary intervention component.

We will exclude study protocols, editorials, and commentaries based on their limited containment of an intervention description and concrete original study results. Further, we will exclude review study types such as narrative, scoping, and systematic reviews and meta-analyses to avoid duplications. We will also exclude conference proceedings and not peer-reviewed publications (original studies and conference papers). Lastly, publications will be excluded in the screening process if the described intervention does not match our definition of digital public health.

Literature Search

We will use a two-part search strategy to identify publications that meet our inclusion criteria. For the first part, we searched three electronic bibliographic databases for published work on February 19, 2021, using a comprehensive search strategy for possible DiPHIs: PubMed, CENTRAL (Cochrane Central Register of Controlled Trials), and Web of Science (see Figure 2, left box in the identification phase). We added two databases, IEEE (Institute of Electrical and Electronics Engineers) Xplore and the Association for Computing Machinery (ACM) Full-Text Collection, to the systematic search to identify more publications from computer science. Both databases were searched on December 1, 2021 (see Figure 2, middle box in the identification phase). All five search strategies are based on the PubMed search string but have been adapted to consider differences in the vocabulary used by the database (Medical Subject Heading [MeSH] terms) and its syntax specifications. To ensure that the systematic search results are not outdated by the time of data extraction, we set up alerts for all databases about new publications on the topic of interest. We will include all alerts that fit our inclusion criteria and that appeared until the start of the full-text screening. For the second part of our search strategy, we will manually screen the reference lists of studies included in the scoping review (see Figure 2, right box in the identification phase). This will ensure that relevant studies are not overlooked.

Figure 2. Flow chart of the search and screening process. ACM: Association for Computing Machinery; CENTRAL: Cochrane Central Register of Controlled Trials; IEEE: Institute of Electrical and Electronics Engineers.

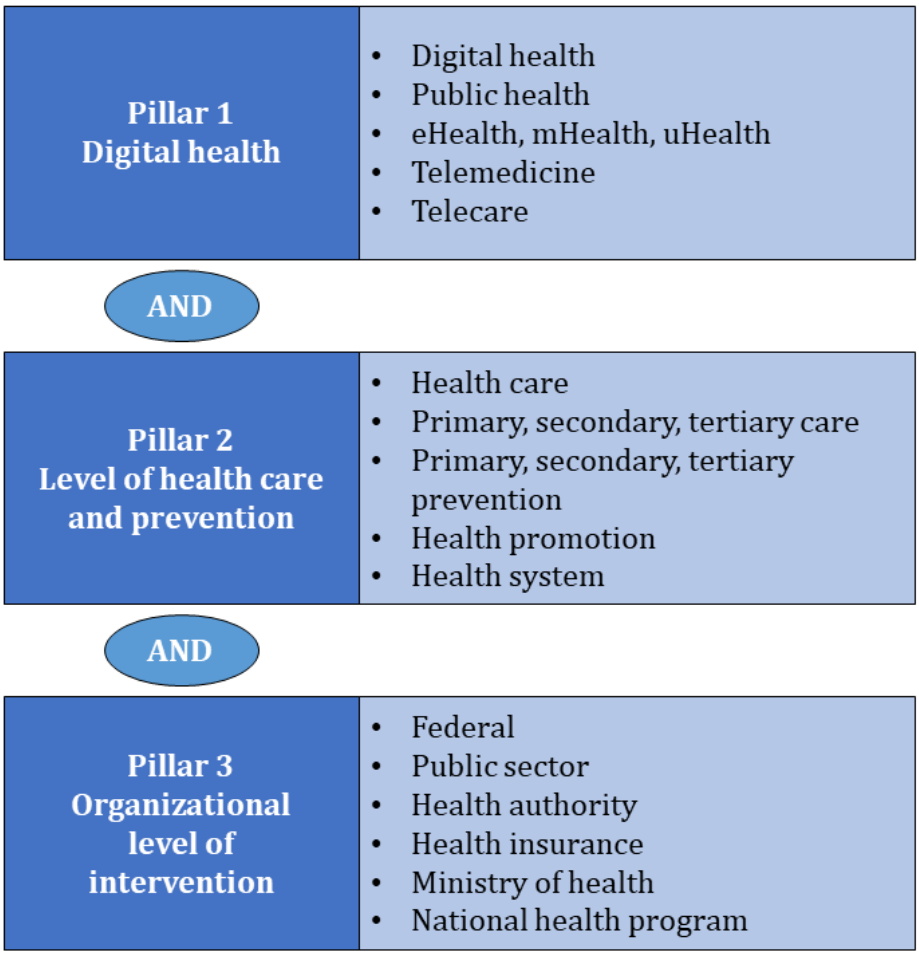


The search strategy for each database consists of three pillars, each combined with "AND." Individual terms within the bodies are connected with "OR." The first pillar includes terms related to digital health and the second contains different health care and prevention levels. The third pillar finally lists search terms to describe the organizational level of interventions (see Figure 3 for the short version and see Multimedia Appendix 1 for the extended version for all five databases). We decided against naming concrete interventions such as EHR, mHealth apps, or electronic prescriptions to reduce the risk of confirmation bias.

For PubMed and CENTRAL, the search bodies also included MeSH terms to identify publications listed within those categories. All search terms will be limited to title, abstract, and keywords. There will be no use of additional filters such as language, geographical, or year of publication restrictions. Exclusion of publications due to another language will be made during full-text screening. Our scoping review will not have any unpublished paper or grey literature.

Identified studies will be screened independently by two authors (LM and MF). Following the aforementioned inclusion criteria, we will separate the screening process into title screening, abstract screening, and full-text screening. In case of disagreements over the eligibility of specific papers, the two screening authors (LM and MF) will discuss whether or not to include a publication for each of the three stages of the screening process (title, abstract, and full-text screening). If the disagreement cannot be resolved through discussion, the third author (CCP) will have the final decision on the eligibility of the publication. We will calculate the Cohen κ value for each step of a screening process to illustrate the agreement between the two screening authors (LM and MF).

Figure 3. Short version of the search strategy. mHealth: mobile health.



Data Extraction

Following quality assessment, we will extract data from the included studies. Two review authors (LM and CCP) will independently extract data in a Microsoft Excel 2019 sheet. These data will contain information on: (1) participants, including recruitment, study completion rates, study population, and participant demographics (age, gender, insurance status, race, ethnicity, education level, income, geography, and language); (2) intervention, including a description of the intervention, target group, use case, addressed level of prevention/health care, geographical level of intervention (local, regional, national), and core function/modules of the intervention; (3) outcomes, including indicators of user acceptability, indicators of health economic evaluation, and other outcomes and indicators; and (4) setting, including the country, year of publication, publication type, study methodology, and year of data collection.

We will request missing data from the authors of our included studies via email. We will resolve any discrepancies through discussion with the third author, MF.

Synthesis of Results

All included studies will be summarized in a narrative synthesis. We will group all included papers in a table according to the described intervention type (see Figure 1). Based on the interventions' descriptions, we will map out each intervention type's characteristics and technological functions. As an example, many countries offer an EHR system at the national level. However, most of them have a different understanding of EHRs (which makes empirical research on this issue complicated). Our scoping review will therefore extract characteristics of the EHR from one publication (reporting on a whole country or a state/region within a county) and compare those with the attributes of an EHR from a different report (country, state, region). This approach will result in overlaps of characteristics, which can be defined as core characteristics if the majority of all EHRs share this attribute. Some EHRs might have features that none or only a few others have. These can be defined as "added characteristics." To display common functions within each intervention type, we will create a table with all mentioned characteristics and rank each intervention based on how many features it includes (see Table 2).

Table 2. Example table of core characteristics for a given intervention group based on electronic health records.

Study	Characteristic 1	Characteristic 2	Characteristic 3	Characteristic 4
Study 1	✓	✓	✓	
Study 2		✓		✓
Study 3	✓		✓	✓

For each intervention type, we will analyze the following: the country setting by income level according to the World Bank [28], level of prevention and health care (primary, secondary, or tertiary prevention), target group (eg, medical provider, health insurance, researchers, or general population), and use case (eg, communication facilitator, health education, tracking, tracing, surveillance, or self-management of chronic disease).

We will further analyze differences between locally implemented interventions and national digital public health policies for the same intervention group. The clusters for categorization of the interventions' features will be developed during data extraction based on our research evidence. This approach is used mainly in qualitative research. Based on the aim of our scoping review, we decided against predefined clusters for categorization as these would require predefined descriptions of intervention characteristics, which we want to explore with this review.

Results

The systematic search in the three databases, Web of Science, CENTRAL, and PubMed, was performed on February 19, 2021, and produced 18,363 results. A total of 13,383 papers were included in the review after deduplication. Of these, 2962 titles were included for abstract screening. We performed an additional systematic search in the IEEE Xplore and ACM Full-Text Collection databases on December 1, 2021, which identified another 491 titles, 38 of which were included for abstract screening. In addition, 22 abstracts from the second search were included for full-text screening. In total, we have included 1417 publications for full-text screening and expect to complete this scoping review in April 2022. The results will be published in peer-reviewed journals and conferences based on the identified outcomes.

Discussion

General Aims and Significance

This paper presents the protocol for a scoping review following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) reporting standards [29]. This scoping review of DiPHIs will provide an extensive report of the literature about worldwide existing DiPHIs that take advantage of ICT. To our knowledge, no review has been conducted with the broad scope proposed here. We will show what characteristics they have in common and the areas they cover from a population-health level. This information will be useful as a blueprint for future DiPHI development to support countries in adapting digital public health to all areas of their health system.

Strengths

The chosen method of a scoping review is an effective technique for mapping comprehensive and interdisciplinary topics such as digital public health. The search is performed in the three largest databases for digital public health topics (ie, PubMed, CENTRAL, and the Web of Science) and the two largest databases for computer science (IEEE Xplore and ACM Full-Text Collection), which reduces the risk of missed articles. Two researchers (LM and MF) will systematically and independently select the studies that fit the inclusion criteria. The third author (CCP) will settle disagreements to ensure the reliability of the results. Additionally, we explained our study identification process in detail, including inclusion criteria, to ensure reproducibility. The review will follow the PRISMA-ScR checklist that is specific for scoping reviews [29] to ensure the high level of quality and transparency of this study.

Limitations

One limitation of our scoping review is that we do not include study protocols, editorials, or commentaries. It is also worth mentioning that although we include research performed worldwide, our review is limited to publications in English, German, or Chinese. This could limit the completeness of the identified publications. Publications might also be missed as we will restrict the manual search to the reference lists of included studies. We will not assess the quality of the included literature in this review, which may lead to concerns about the accuracy of the literature and affect the generalizability and evidence of the results. However, as we are not assessing the outcome results of the selected studies but only examining their description of a DiPHI, the missing quality analysis is not relevant to our scoping review. The last limitation is that this review is not going to assess the impact of the implementation of DiPHIs in a health system, but will rather only provide an overview of existing technologies and their characteristics.

Acknowledgments

The authors gratefully acknowledge the support of Leibniz ScienceCampus Bremen Digital Public Health, which is jointly funded by the Leibniz Association (W4/2018), Federal State of Bremen, and Leibniz Institute for Prevention Research and Epidemiology. This scoping review is part of a doctoral thesis funded by the Leibniz ScienceCampus Bremen.

Authors' Contributions

LM initiated and designed the study, provided a first version of the search strategy, and drafted the manuscript. LM was one of two independent literature reviewers. CCP helped finalize the search strategy, settled disagreements in the screening process, and reviewed the manuscript. MF was one of two independent literature reviewers and reviewed the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete search strategies for the PubMed, Web of Science, CENTRAL, IEEE Xplore, and ACM Full-Text Collection databases.
[DOCX File , 51 KB - resprot_v11i3e33404_app1.docx]

References

1. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. J Diabetes Sci Technol 2017 Sep;11(5):1015-1027 [FREE Full text] [doi: 10.1177/1932296817713506] [Medline: 28560898]
2. Bailey J, Murray E, Rait G, Mercer C, Morris R, Peacock R, et al. Interactive computer-based interventions for sexual health promotion. Cochrane Database Syst Rev 2010 Sep 08(9):CD006483. [doi: 10.1002/14651858.CD006483.pub2] [Medline: 20824850]
3. Strain T, Wijndaele K, Brage S. Physical activity surveillance through smartphone apps and wearable trackers: examining the UK potential for nationally representative sampling. JMIR Mhealth Uhealth 2019 Jan 29;7(1):e11898 [FREE Full text] [doi: 10.2196/11898] [Medline: 30694198]
4. Zeeb H, Pigeot I, Schüz B. Digital public health-rapid technological progress, but many open public health questions. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Feb;63(2):135-136 [FREE Full text] [doi: 10.1007/s00103-020-03092-0] [Medline: 31960072]
5. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive health communication applications for people with chronic disease. Cochrane Database Syst Rev 2005 Oct 19(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
6. Alwashmi M, Hawboldt J, Davis E, Marra C, Gamble J, Abu Ashour W. The effect of smartphone interventions on patients with chronic obstructive pulmonary disease exacerbations: a systematic review and meta-analysis. JMIR Mhealth Uhealth 2016 Sep 01;4(3):e105 [FREE Full text] [doi: 10.2196/mhealth.5921] [Medline: 27589898]
7. Gunasekeran D, Tseng R, Tham Y, Wong T. Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. NPJ Digit Med 2021 Feb 26;4(1):40. [doi: 10.1038/s41746-021-00412-9] [Medline: 33637833]
8. Azzopardi-Muscat N, Sørensen K. Towards an equitable digital public health era: promoting equity through a health literacy perspective. Eur J Public Health 2019 Oct 01;29(Supplement_3):13-17 [FREE Full text] [doi: 10.1093/eurpub/ckz166] [Medline: 31738443]
9. Dalton C. Enablers of innovation in digital public health surveillance: lessons from Flutracking. Int Health 2017 May 01;9(3):145-147 [FREE Full text] [doi: 10.1093/inthealth/ihx009] [Medline: 28582555]
10. Desai AN, Kraemer MUG, Bhatia S, Cori A, Nouvellet P, Herringer M, et al. Real-time epidemic forecasting: challenges and opportunities. Health Secur 2019;17(4):268-275 [FREE Full text] [doi: 10.1089/hs.2019.0022] [Medline: 31433279]
11. Shakeri Hossein Abad Z, Kline A, Sultana M, Noaeen M, Nurmambetova E, Lucini F, et al. Digital public health surveillance: a systematic scoping review. NPJ Digit Med 2021 Mar 03;4(1):41. [doi: 10.1038/s41746-021-00407-6] [Medline: 33658681]
12. Darmann-Finck I, Rothgang H, Zeeb H. Digitalization and health sciences - White Paper Digital Public Health. Gesundheitswesen 2020 Jul;82(7):620-622 [FREE Full text] [doi: 10.1055/a-1191-4344] [Medline: 32698204]
13. Taj F, Klein M, van Halteren A. Digital health behavior change technology: bibliometric and scoping review of two decades of research. JMIR Mhealth Uhealth 2019 Dec 13;7(12):e13311 [FREE Full text] [doi: 10.2196/13311] [Medline: 31833836]
14. Zeeb H, Pigeot I, Schüz B, Leibniz-WissenschaftsCampus Digital Public Health Bremen. Digital public health-an overview. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Feb;63(2):137-144. [doi: 10.1007/s00103-019-03078-7] [Medline: 31919531]
15. Odone A, Buttigieg S, Ricciardi W, Azzopardi-Muscat N, Staines A. Public health digitalization in Europe. Eur J Public Health 2019 Oct 01;29(Supplement_3):28-35 [FREE Full text] [doi: 10.1093/eurpub/ckz161] [Medline: 31738441]
16. Winslow CE. The untilled fields of public health. Science 1920 Jan 09;51(1306):23-33. [doi: 10.1126/science.51.1306.23] [Medline: 17838891]
17. What is public health? CDC Foundation. URL: https://www.cdcfoundation.org/what-public-health [accessed 2021-03-21]
18. World Health Organization. Essential public health functions, health systems and health security. Developing conceptual clarity and a WHO roadmap for action. Geneva: World Health Organization; 2018.

19. Wienert J, Jahnel T, Maaß L. What are digital public health interventions? First steps towards a definition and an intervention classification. JMIR Preprints. 2022 Jan 04. URL: https://preprints.jmir.org/preprint/31921 [accessed 2022-03-24]
20. Peek N, Sujan M, Scott P. Digital health and care in pandemic times: impact of COVID-19. BMJ Health Care Inform 2020 Jun;27(1):E100166 [FREE Full text] [doi: 10.1136/bmjhci-2020-100166] [Medline: 32565418]
21. Shaw T, McGregor D, Brunner M, Keep M, Janssen A, Barnet S. What is eHealth (6)? Development of a conceptual model for eHealth: qualitative study with key informants. J Med Internet Res 2017 Oct 24;19(10):e324 [FREE Full text] [doi: 10.2196/jmir.8106] [Medline: 29066429]
22. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
23. Murray C, Alamro N, Hwang H, Lee U. Digital public health and COVID-19. Lancet Public Health 2020 Sep;5(9):e469-e470 [FREE Full text] [doi: 10.1016/S2468-2667(20)30187-0] [Medline: 32791051]
24. Hochmuth A, Exner A, Dockweiler C. Implementation and participatory design of digital health interventions. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Feb;63(2):145-152. [doi: 10.1007/s00103-019-03079-6] [Medline: 31938837]
25. Jahnel T, Schüz B. Participatory development of digital public health: tension between user perspectives and evidence. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Feb;63(2):153-159. [doi: 10.1007/s00103-019-03082-x] [Medline: 31915866]
26. Bashi N, Fatehi F, Mosadeghi-Nik M, Askari M, Karunanithi M. Digital health interventions for chronic diseases: a scoping review of evaluation frameworks. BMJ Health Care Inform 2020 Mar;27(1):E100066 [FREE Full text] [doi: 10.1136/bmjhci-2019-100066] [Medline: 32156751]
27. Nicholl B, Sandal L, Stochkendahl M, McCallum M, Suresh N, Vasseljen O, et al. Digital support interventions for the self-management of low back pain: a systematic review. J Med Internet Res 2017 May 21;19(5):e179 [FREE Full text] [doi: 10.2196/jmir.7290] [Medline: 28550009]
28. World Bank Country and Lending Groups. The World Bank. 2021. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519 [accessed 2022-03-21]
29. Tricco A, Lillie E, Zarin W, O'Brien K, 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 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]

Abbreviations

ACM: Association for Computing Machinery
CENTRAL: Cochrane Central Register of Controlled Trials
DiPHI: digital public health intervention
EHR: electronic health record
ICT: information and communication technology
IEEE: Institute of Electrical and Electronics Engineers
MeSH: Medical Subject Headings
mHealth: mobile health
PIOS: Participants, Intervention, Outcome, and Study design
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by T Leung; submitted 06.09.21; peer-reviewed by H De Loof, K Craig, L Chen; comments to author 11.11.21; revised version received 02.01.22; accepted 05.02.22; published 31.03.22.

Please cite as:
Maaß L, Pan CC, Freye M
Mapping Digital Public Health Interventions Among Existing Digital Technologies and Internet-Based Interventions to Maintain and Improve Population Health in Practice: Protocol for a Scoping Review
JMIR Res Protoc 2022;11(3):e33404
URL: https://www.researchprotocols.org/2022/3/e33404
doi:10.2196/33404
PMID:35357321

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org