JMIR Research Protocols

Impact Factor (2023): 1.4
Volume 6 (2017), Issue 8 ISSN 1929-0748 Editor in Chief: Xiaomeng (Simone) Ma, PhDc, MS, BS

Contents

Protocols

Proposal

Tutorial

Original Papers

Short Paper

Corrigenda and Addenda

Protocol

Pacing, Conventional Physical Activity and Active Video Games to Increase Physical Activity for Adults with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Protocol for a Pilot Randomized Controlled Trial

Katia Elizabeth Ferrar[1], B.App.Sc (PT), B. Hlth. Sc (Hons), PhD; Ashleigh E Smith[1], B. Sc (Biomedical Science), B.Sc (Hons), PhD; Kade Davison[1], B.App.Sc(HM), B.Hlth.Sc.(Hons), PhD

Alliance for Research in Exercise, Nutrition and Activity, School of Health Sciences, University of South Australia, Adelaide, Australia

Corresponding Author:
Katia Elizabeth Ferrar, B.App.Sc (PT), B. Hlth. Sc (Hons), PhD
Alliance for Research in Exercise, Nutrition and Activity
School of Health Sciences
University of South Australia
GPO Box 2471
CEA-14
Adelaide,
Australia
Phone: 61 8 8302 2554
Fax: 61 8 8302 2853
Email: katia.ferrar@mymail.unisa.edu.au

Abstract

Background: Myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) is a serious illness of biological origin characterized by profound physical and cognitive exhaustion and postexertion malaise. Pacing is a common strategy used to manage available energy and complete activities of daily living; yet little research has investigated this as a strategy to increase physical activity levels. Typically, people living with ME/CFS are faced by unique barriers to physical activity participation and are less physically active than healthy peers. As such they are at increased risk of physical inactivity–related health consequences. Active video games may be a feasible and acceptable avenue to deliver physical activity intervention by overcoming many of the reported barriers to participation.

Objective: The primary objective of this pilot study is to determine the feasibility and acceptability of active video games to increase physical activity levels of people with ME/CFS. The secondary aims are to explore the preliminary effectiveness of pacing and active video gaming to pacing alone and pacing plus conventional physical activity to increase the physical activity levels of adults with ME/CFS and explore the relationship between physical activity and cumulative inflammatory load (allostatic load).

Methods: This study will use a mixed method design, with a 3-arm pilot randomized controlled trial, exit interviews, and collection of feasibility and process data. A total of 30 adults with ME/CFS will be randomized to receive either (1) pacing, (2) pacing and conventional physical activity, or (3) pacing and active video gaming. The intervention duration will be 6 months, and participants will be followed up for 6 months postintervention completion. The intervention will be conducted in the participant's home, and activity intensity will be determined by continuously monitored heart rate and ratings of perceived exertion. Feasibility and acceptability and process data will be collected during and at the end of the intervention. Health-related outcomes (eg, physical activity, blood samples, quality of life, and functioning) will be collected at baseline, end of intervention, and 6 months after intervention completion.

Results: This protocol was developed after 6 months of extensive stakeholder and community consultation. Enrollment began in January 2017; as of publication, 12 participants were enrolled. Baseline testing is scheduled to commence in mid-2017.

Conclusions: This pilot study will provide essential feasibility and acceptability data which will guide the use of active video games for people with ME/CFS to increase their physical activity levels. Physical activity promotion in this clinical population

has been poorly and under-researched, and any exploration of alternative physical activity options for this population is much needed.

Trial Registration: Australia New Zealand Clinical Trials Registry: ACTRN12616000285459; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=370224 (Archived by WebCite at http://www.webcitation.org/6qgOLhWWf)

(JMIR Res Protoc 2017;6(8):e117) doi:10.2196/resprot.7242

KEYWORDS

fatigue syndrome, chronic; video games; exercise

Introduction

Myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) is a complex, serious illness of biomedical origin characterized by profound physical and cognitive exhaustion and postexertion malaise [1]. Symptoms such as fatigue and cognitive impairments are typically made worse to the point of collapse after physical exertion [1]. The combination of physical inability to be active due to the condition and behaviors to avoid potential flare-up results in people with ME/CFS being less physically active than healthy controls [2].

The mechanisms underlying ME/CFS remain unclear, and studies report associations with multiple individual proinflammatory markers and indicators of immune activity, but the findings are inconsistent [3]. A potential mechanism by which physical activity plays a role in managing chronic conditions is through allostasis: the body's response to stress orchestrated by the hypothalamic-pituitary-adrenal axis and the central nervous system via the production of stress hormones [4]. The detrimental negative effect of allostasis can be quantified as allostatic load (AL). The theoretical mechanism between AL and physical activity has recently been empirically supported, with higher levels of physical activity associated with lower AL in Mexican American adults [5] and middle-aged women [6]. Further, adults living with ME/CFS are reported to have higher AL than healthy controls [7,8]. It is widely accepted that cumulative biomarkers are stronger predictors of morbidity and mortality, yet recent research into ME/CFS continues to investigate individual neuroendocrine and immune biomarkers with limited success and inconsistent findings [3]. Further, few AL studies are longitudinal or interventional in nature, limiting the relationships and conclusions that can be drawn.

Physical activity is essential for health and well-being and has a role in prevention and management of muscle atrophy and many chronic conditions such as cardiovascular disease, osteoporosis, and some cancers [9]. There is much debate in the literature surrounding the effectiveness and safety of exercise or physical activity-based interventions for people with ME/CFS. In this population, the majority of physical activity–based interventions have employed a Graded Exercise Therapy (GET) approach. Typically, after setting baseline levels, the intention of GET protocols is to incrementally increase activity despite possible changes in symptom severity. One common criticism of GET protocols is they typically suggest stabilization of activity progression at the point of symptom flare-up but do not appear to include modification or cessation of activity to promote symptom resolution. The results of GET protocols are mixed, with some reporting symptom improvement [10,11]. The risk of adverse responses to GET is not clear from intervention studies; however, support group surveys of people living with ME/CFS suggest a significant proportion of these interventions result in negative experiences during the intervention [12].

An alternative approach to delivery of physical activity interventions is pacing. Pacing is a typical strategy used by people with ME/CFS to manage their activities of daily living within their available energy envelope by structuring and organizing their daily activities. Pacing is typically well accepted by people living with ME/CFS (according to patient surveys) and is used primarily to manage symptom flare-ups and work within the available energy, but it is also recommended that pacing can help to slowly increase the energy envelope and thus the level of activity performed. Pacing is a flexible, symptom-contingent process that can be used to increase physical activity levels. Survey data suggests a greater percent of people living with ME/CFS feel pacing improved their symptoms (45%, n=69) compared to GET (12%, n=77) [12]. In addition, anecdotal evidence suggests a large percentage of people living with ME/CFS already use pacing and heart rate (HR) monitoring to some extent (personal communication by ME/CFS Australia [SA] spokesperson, March 2016), and as such, pacing will be considered the control situation for this study.

Only one published study can be located that has employed a flexible symptom-contingent protocol (ie, pacing using HR and perceived exertion monitoring) to increase activity levels in people with ME/CFS [13]. The study demonstrated improvements in physiological outcomes (eg, blood pressure, cognitive ability) but no change in physical activity between the groups postintervention. The study used a self-report activity questionnaire with poor reliability at low intensity levels of activity, which is concerning because this accounts for almost all physical activity in most people with or without ME/CFS, and only low to moderate validity with other self-reported physical activity outcomes [14]. Overall, the flexible, symptom-contingent approach to increasing activity levels in people with ME/CFS has not been adequately researched. An important finding from the Wallman et al [13] study is that none of the participants in the exercise group rated themselves as any worse and 91% rated themselves some degree better overall. The current study will build on this small body of literature by investigating pacing alone versus pacing and physical activity using objective measures of physical activity and safely

monitoring activity intensity via continuous HR monitoring and ratings of perceived exertion (RPE).

People with ME/CFS report unique barriers to participation in physical activity. These barriers include extreme fatigue just getting to and from structured activity, greater difficulty being active standing compared to sitting, and ability to participate in only very short durations of activity [1]. Interventions using commercially available active video games, such as Xbox Kinect, can overcome many of these barriers for people with ME/CFS. Active video games can be played at home either while standing or seated, can be easily played for short periods of time (eg, 1-2 minutes), and provide choice of different gaming activities—factors which will potentially reduce patient dissatisfaction and drop out [15].

An inability to fully participate in their family's lives and community often results in people living with ME/CFS reporting depression, social isolation, and lack of connectedness [1]. Active video games have the potential to provide an avenue for family connectedness and promote a feeling of normality while enabling participation (potentially with others) in an everyday activity. People's experiences with active video gaming suggest substantial benefits including feelings of empowerment, reductions in depression, increased psychosocial well-being, and increased social connectedness by playing with family members [16].

Active video gaming has been shown to be feasible and acceptable to other chronic health populations such as those with cerebral palsy and stroke and is commonly played while sitting in these populations [17]. Evidence suggests active video gaming can improve a range of rehabilitative outcomes, increase physical activity levels among sedentary and frail elderly adults, and increase cognitive performance and reduce fatigue in women with systemic lupus erythematosus [18,19,20]. Despite the evidence to suggest active video gaming as effective and relevant, no studies have investigated the feasibility or acceptability of active video gaming in adults with ME/CFS.

The primary objective of this pilot study is to investigate the feasibility and acceptability of pacing and active video gaming to increase the physical activity levels of adults with ME/CFS.

The secondary objectives of this study are to (1) explore the preliminary effectiveness of pacing and active video gaming to pacing alone and pacing and conventional physical activity to increase the physical activity levels of adults with ME/CFS and (2) explore the relationship between allostatic load and physical activity in people with ME/CFS.

The related hypotheses are (1) preliminary evidence of benefit from pacing and active video gaming to increase physical activity compared to pacing alone will be demonstrated, (2) preliminary evidence of equal benefit of pacing and video gaming to increase physical activity levels compared to pacing and conventional exercise will be demonstrated, and (3) preliminary evidence of a negative relationship will be demonstrated between allostatic load and physical activity.

Methods

Trial Design

A mixed-method study will be conducted comprising an exploratory 3-armed randomized controlled pilot study (1:1:1 allocation ratio) of pacing versus pacing and conventional physical activity versus pacing and active video gaming; qualitative exit interviews and collection of process data to investigate the feasibility and acceptability of the approach. The study protocol has been approved by the University of South Australia Human Research Ethics Committee (protocol 0000035299). The protocol is registered with the Australia New Zealand Clinical Trials Registry [ACTRN12616000285459]. All participants will be required to provide written informed consent prior to enrolling in the study and report general practitioner consent to participate.

Protocol Development

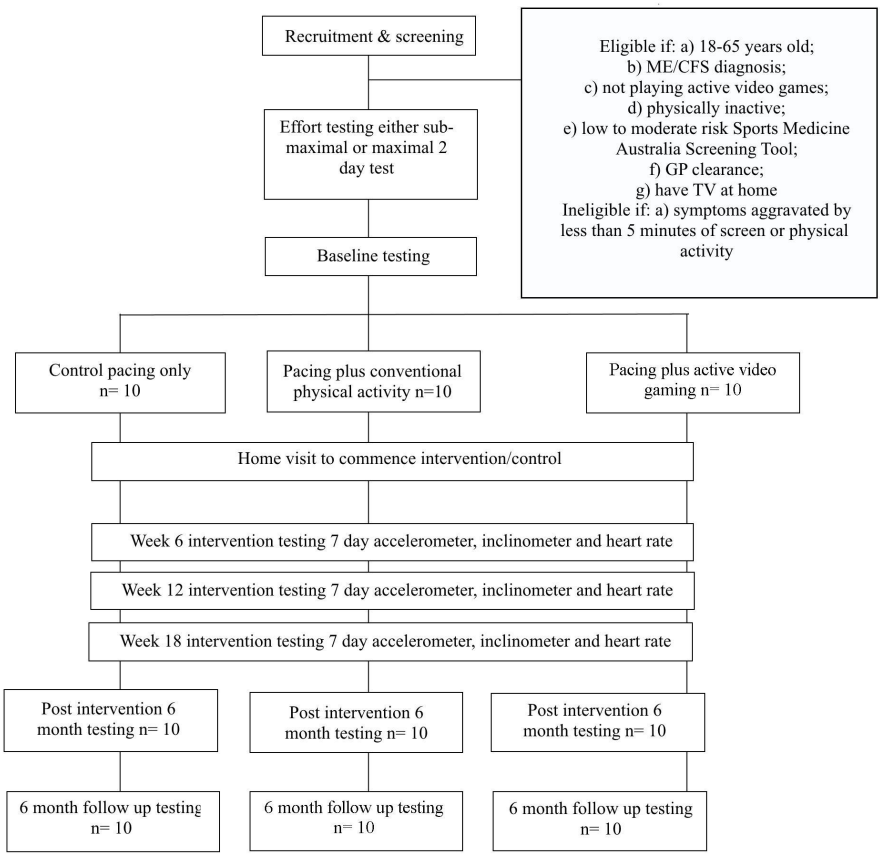
A stakeholder advisory group was convened including 2 representatives from the local state ME/CFS society, a rheumatologist with extensive clinical and research knowledge of ME/CFS, a general physician with extensive clinical expertise in managing ME/CFS, 2 adults currently living with ME/CFS, and the research team. The primary purpose of the stakeholder advisory group was to assist in the consumer-driven and evidence-based development of the final protocol. A 6-month consultative development phase was completed in December 2016. The stakeholder advisory group welcomed and considered comments and concerns from external bodies (eg, national and international support groups).

Intervention

After baseline assessment, participants will be block randomized (block size = 6, 1:1:1 allocation ratio) into pacing only, pacing plus conventional physical activity, or pacing plus active video gaming (Figure 1).

Participants will be randomized using a computer-generated random number sequence and opaque sealed envelopes generated by a University of South Australia researcher not involved in the study. All participants will receive telephone support from the research assistant once a week for the first 3 months of the intervention and every 2 weeks for the last 3 months of the intervention, and participants can contact the research assistant at any additional time. At the end of the 12 months, participants still enrolled will be provided with an Xbox Kinect unit in acknowledgement of their participation.

Figure 1. Study Flow Chart.



Control Condition

In the pacing only condition, participants will be educated about pacing to manage their daily energy levels and informed not to commence a physical activity program for 6 months. Pacing is a process whereby activities of daily living are organized and timed in order to not exceed the available energy for that particular day. Pacing will be self-monitored to avoid excessive exertion by means of continuous HR monitoring and RPEs maintained below a designated threshold. The HR and RPE

thresholds will be determined during the baseline effort test, which is described in detail in the Methods section. Participants in the pacing only control condition will be provided with no instruction to increase physical activity levels.

Intervention Conditions

The intervention will be 6 months in duration. In the pacing plus conventional activity condition, participants will be educated about pacing and prescribed a home-based participant-controlled symptom-contingent physical activity program involving a conventional activity type selected by the participant such as walking, elastic band resistance exercise, or cycling.

In the pacing and active video gaming condition, participants will be provided with an Xbox Kinect (set up in their home) and prescribed a home-based participant-controlled symptom-contingent physical activity program involving dance and sport games (Sports Rivals and Dance Central). Swapping passive screen time (such as television viewing) with the active gaming will be recommended. If the participants report no or minimal screen time, they will be encouraged to swap out other activities that require low cognitive or physical load. The active video gaming and conventional physical activity is intended to be additional to existing physical activities.

Participants in the two conditions involving physical activity (ie, pacing and conventional physical activity and pacing and active video gaming) will have their starting physical activity prescription developed by the research assistant who will be an accredited exercise physiologist with at least 2 years of clinical experience. The research assistant will develop the prescription based on participant self-reported amount of activity and symptom relationship from the Symptom History Questionnaire (detailed later in Methods section). Prescription will involve allocation to 1 of 3 tiers: light tolerance, moderate tolerance, or high activity tolerance (Table 1).

Table 1. Physical activity prescription triage categories.

Activity tolerance tier	Categorization	Baseline	Progress
Exclude	Self-reported flare-up of symptoms with less than 5 minutes of screen time; moderate to severe symptom flare-up with less than 5 minutes of self-reported light physical activity or household chores	—	—
Light	Flare-ups with predominantly moderate to high intensity tasks or longer bouts of light intensity tasks or pushing energy envelope with too many activities on a given day; able to participate in light activity for at least 5 minutes (but less than 10 minutes)	2 minutes per day (6-8 minutes per week)	30 seconds to 1 minute per day per intervention week
Moderate	Symptom flare-ups with activities predominantly considered moderate or high intensity (ie, minimal flare-ups with light intensity tasks), able to be active for between 10 to 15 minutes at a light intensity with minimal or no flare-up of symptoms	5 minutes per day (15-20 minutes per week)	1 to 2 minutes per day per intervention week
High	Reported symptom flare-ups with activities predominantly considered high intensity or longer bouts moderate tasks (ie, minimal flare-ups with light intensity and short bouts of moderate intensity tasks), able to be active for 15 or more minutes at a light intensity with minimal or no flare-up of symptoms	10 minutes per day (30-40 minutes per week)	2 to 4 minutes per day per intervention week

Physical activity will be prescribed every second day (at the most) and perceived exertion will be monitored by the participant remaining under the set HR and RPE thresholds at all times. Participants will be in full control of their program and will be able to stop physical activity sessions entirely; plateau; or reduce frequency, duration, or intensity of sessions to manage symptom occurrences. Activity sessions can be completed in one session (eg, 1 10-minute block) or split up throughout the day.

Any activity progression will occur every 2 weeks. To be eligible to progress, participants must demonstrate symptoms are stable. Stable equates to HR and RPE zones adhered to for the 2-week period and no increase in symptoms as a result of activity. Progression can be escalated (particularly as higher volumes demonstrated) with consultation with research team and research assistant only if symptoms are stable. The active gaming or conventional activity will be encouraged on those days that the participants consider there to be adequate energy to do all activities of daily living usual for that person. Participants will also be encouraged to consider cognitive activities as potentially draining in addition to physical activities.

Setting

Participants will be recruited from Adelaide, South Australia, Australia. Testing will occur during face-to-face sessions at the University of South Australia, City East Campus, Adelaide, Australia. Home visits will be conducted to educate the participants regarding their condition allocation and set up video games where applicable, and the remainder of the intervention will be home-based.

Participants

Inclusion criteria:

- Aged from 18 to 65 years
- General practitioner clearance
- Self-reported diagnosis of ME/CFS by a general practitioner or medical specialist based on one of the commonly accepted criteria (the Oxford Criteria is not acceptable)

- Able to complete exercise test (either maximal or submaximal)
- Not currently playing active video games
- Score low to moderate on Sports Medicine Australia Exercise Screening Tool [21]
- Self-report less than 150 minutes of moderate intensity activity each week (not meeting Australian National Physical Activity Guidelines for adults [22])
- Have a television in their home (Internet not required) and be willing to develop video game literacy

Exclusion criteria:

- Self-reported aggravation of symptoms with 5 or fewer minutes of screen time
- Self-reported aggravation of symptoms with 5 or fewer minutes of light intensity physical activity or movement

Sample Size

This is the first study to investigate the feasibility and acceptability of active video gaming to increase the physical activity levels of people living with ME/CFS. A sample of 30 participants was deemed appropriate to provide the process feedback regarding testing procedures, and feasibility and acceptability data will be collected from 10 participants regarding the active video gaming experience.

Recruitment

Participants will be recruited via local community support organizations, fliers, Facebook advertising, and the University of South Australia Clinical Trials Facility website.

Feasibility and Acceptability

To explore the feasibility of implementation, total testing time will be calculated at each testing point for each participant. Statistics regarding the total amount and type of resources required (eg, consumables, printing) will be collected. Recruitment feasibility will be investigated after recruitment phase and involve analysis of recruitment statistics (eg, duration, ineligibility vs eligibility ratio number of recruitment avenues). Comprehensive informal exit interviews will be conducted with the active video game intervention group participants at the end of the intervention to explore acceptability either face-to-face or over the telephone. Information will be obtained on general satisfaction with the program, perceived cost effectiveness, acceptability of the testing protocols, and desire to continue with active video gaming.

Effort Testing (Baseline Only)

Prior to randomization, participants will be screened for eligibility and will participate in either a maximal effort 2-day testing protocol or a submaximal effort protocol. A submaximal effort protocol has been included as an option as community feedback indicated unwillingness from some people due to the potential symptom flare-up predicted with the maximal protocol.

Effort testing on a bicycle ergometer is required to accurately determine the HR and RPEs at ventilatory threshold (VT) or 10% below the VT for each participant. This 10% below VT has been posited as a preferred limit to stay below to avoid symptom flare-up [23] and will be employed in this study for all 3 arms.

The 2-day maximal effort testing protocol is a smaller study running concurrently at the University of South Australia looking to confirm the differing physiological responses demonstrated by people with ME/CFS due to postexertional malaise during repeat exercise testing when compared to healthy controls. The 2-day maximal effort testing protocol will involve attending the university on 2 consecutive days and participating in the same testing protocol on a cycle ergometer wearing a gas analysis mask. A subjective questionnaire regarding fatigue status will be completed, followed by a 10-minute period of rest while wearing an HR monitor. The incremental exercise test on the bike will progress from very light to maximal exercise over a period of 8 to 12 minutes. The test starts with a 5-minute, 40 watt warm-up, during which rate of HR increase is assessed. Following this 5 minutes, the test increases using a ramp protocol, with increases of 5 watts every 20 seconds (so 15 watts per minute on average). The test finishes when the participant reaches volitional exhaustion. Both testing sessions will be conducted at the same time of day.

The submaximal effort testing alternative will involve attending the university on 1 occasion. Participants will complete 1 test on a cycle ergometer while wearing a gas analysis mask. Participants will exercise on a cycle ergometer commencing at a work rate of 25 watts and increasing by 15 watts each minute until a HR of 85% of age predicted maximum is achieved (208–[0.7×age]) [24]. The test will be terminated prior to this point if participant reaches voluntary exhaustion or reports an RPE of 19 or higher, respiratory exchange ratio of >1.15, or a plateau in HR or VO_2 with increasing work rate. Recommended exercise intensity will be limited to an HR corresponding to 10% below measured VT or in the case that VT is not adequately detected in the exercise test then below 70% of age-predicted HR max.

Health-Related Outcomes

Overview

There will be at least 4 weeks between effort testing and commencement of the intervention with 2 weeks being stable at a pretest pattern of symptoms/activity. Participants may require longer recovery before commencing the intervention.

All health-related outcomes will be measured at 3 time points: baseline, 6 months after baseline at the end of the intervention, and at 6-month follow-up after the intervention has been completed. In addition, accelerometer, inclinometer, and continuous captured HR will be measured during week 6, week 12, and week 18 of the intervention. The Karnofsky Scale [25] will be collected more frequently, with the research assistant recording the Karnofsky Scale value each time telephone contact is made with a participant.

Health-related outcomes have been chosen after extensive consultation with the stakeholder advisory group which included people living with ME/CFS. Outcomes have been selected to first provide data on the primary health-related outcomes (ie, physical activity and allostatic load:11 variables). Outcomes

have also been selected to cover client-centered outcomes such as use of time, quality of life, and functioning. These outcomes are of utmost importance to the people with ME/CFS and capture the broader International Classification of Functioning, Disability, and Health framework concepts of activity limitations and participation restriction [26]. Finally, outcomes have been selected to capture the vast range of symptoms reported by people with ME/CFS to ensure the levels are monitored alongside the physical activity changes to ensure safety and investigate net benefit to the participants.

Adherence Diary

All 3 groups will be provided with a pen and paper adherence diary to complete daily for the 6-month duration of the intervention. The pacing only group will document adherence to pacing protocol and HR achieved (yes/no) and RPE limit achieved (yes/no). The conventional physical activity and video games groups will document minutes of activity, specific day, HR achieved (yes/no), and RPE limit achieved (yes/no).

Sociodemographic Questionnaire

A sociodemographic questionnaire developed for the purpose of this study will capture data such as age, gender, education level, working/volunteer status, income, and duration of conditions.

Symptom History Questionnaire

A questionnaire developed for this study will capture all symptoms related to ME/CFS and also the relationship between symptoms and activity to assist in the pacing and activity protocol prescription for each individual.

Accelerometer

GENEActiv (Active Insights Ltd) accelerometers will be worn on the nondominant hand for 7 days during each testing period. Accelerometers will be worn for 24 hours per day but will be removed for bathing and swimming activities. Participants will complete a log to report periods of nonwear time and sleep/nap times.

Inclinometer

ActivPAL (PAL Technologies Ltd) inclinometers are matchbox-sized devices that measure position relative to gravity. They are worn on the front of the thigh attached with hypoallergenic adhesive tape for a period of 7 days and report the minutes spent sitting, standing, and walking. Inclinometers are the current gold standard measurement of sitting/reclining time, a measure of sedentary time.

Multimedia Activity Recall for Children and Adults

The Multimedia Activity Recall for Children and Adults [27] is a computerized self-report 24-hour recall tool that allows detailed recording of time use behaviors that has been tested on various populations including adults and those with ME/CFS in a recent study conducted by several of the authors. Some participants with ME/CFS report some cognitive strain and burden completing more than 1 24-hour recall intervention over the telephone at once, so for this study, 4 interviews will be conducted each on separate days (during the week that the accelerometer is being worn) capturing at least 1 weekend and 1 weekday.

Height, Weight, and Waist Girth

Weight will be measured using the TANITA BC-418 bioelectrical impedance analysis scale. Participant weight will be taken with participant barefoot and clothed in lightweight clothing only. Height will be measured using the ECOMED seca 284 stadiometer. Two measurements will be taken for both weight and height. If the second weight measurement differs by >1%, a third measurement will be required. Similarly, if the second height measurement is >5 mm different then a third measurement is required. An average of the measurements will be obtained (2 measures) or a median (3 measures). Body mass index (BMI) will be calculated. Waist girth will be measured using an anthropometric tape measure and following standardized protocol by an accredited exercise physiologist.

Blood Pressure

Blood pressure will be taken using standard procedures and a sphygmomanometer while participant is sitting or reclining (same position each time). Systolic and diastolic blood pressure will be recorded.

Heart Rate Variability

RR intervals will be recorded during 5 minutes of supine rest prior to the exercise test using a personal HR monitor (RS800CX, Polar Electro Oy). RR interval data will be downloaded to Polar Protrainer 5 software (Polar Electro Oy) where artifacts or ectopic heart beats are removed through Polar's automatic filtering. Data will then be exported to heart rate variability (HRV) analysis software (Kubios HRV Analysis version 2.0 beta 1, The Biomedical Signals Analysis Group) where any remaining artifacts or ectopic heart beats can be removed and the final 2 minutes of recording analyzed for the HRV time domain variable of interest (root mean square of successive differences).

Allostatic Load

Blood will be taken after fasting to collect data on insulin and glucose, total cholesterol and high-density lipoprotein (HDL) (so we can calculate low-density lipoprotein [LDL]), C-reactive protein (CRP), and adrenaline and noradrenaline. These measures cover both the primary mediators and secondary outcomes of the AL. Other measures used to calculate AL will include HR variability, BMI, waist girth, and blood pressure.

The AL index will be calculated by determining whether, for each participant, their variables are in the healthy zone (coded 0) or the unhealthy zone (coded 1) and then summed to provide a composite AL index. Some variables will be coded based on clinical cut points; the remaining variables that do not have accepted clinical cut points will be analyzed by calculating the upper 25th percentile (code 1), except HRV where the lowest quartile will indicate risk (code 1). The following criteria will be used to determine the risk zones based on previous studies or published guidelines:

- BMI >30 kg/m^2 = 1

- Waist circumference >94 cm (men) and >80 cm (women) = 1
- Fasting lipid profile:
  - Total cholesterol ≥5.2 mmol/L=1
  - HDL <1.00 mmol/L=1
  - LDL ≥3.4 mmol/L=1
- Fasting glucose >5.5 mmol/L=1
- Systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg = 1
- CRP >3.00 mg/L (cardiovascular risk cut point) = 1
- Adrenaline (highest quartile) = 1
- Noradrenaline (highest quartile) = 1
- HRV (lowest quartile) = 1

Prescribed medication will be taken into account so that an individual is given the value 1 (indicating health risk) for diastolic and systolic blood pressure if they used blood pressure lowering medication and for HDL cholesterol if they used lipid lowering medication. The maximum value AL=12, minimum=0.

Diet Diary

A pen-and-paper 3-day food diary will be completed capturing all food and drink consumed during the dates including sweets, snacks, nibbles, sauces and dressings, water and supplements (vitamins). The diet data will be analyzed using the FoodWorks software (Xyris).

Body Temperature

Body temperature will be measured using an electric thermometer. This will be measured to assess if there is an acute inflammatory response (fever). In addition, participant will also be asked to self-report presence of acute infection during testing periods.

Karnofsky Scale

The Karnofsky Scale [25] is used to assess the current abilities of the participant based on a numerical scale from 0 to 100. These scales allow a quick assessment of the participant's function.

Chalder Fatigue Scale

The Chalder Fatigue Scale [28] is one of the most widely used measures for assessing physical and mental symptomatic fatigue experienced by ME/CFS patients. Four response options are available, ranging from "less than usual" to "much more than usual." The Likert system for scoring was used (0, 1, 2, 3), with a total possible score ranging from 0 to 33. A higher score indicates more fatigue.

Cognitive Failure Questionnaire

The Cognitive Failure Questionnaire [29] assesses self-reported deficits in attention, perception, memory, and motor functioning. The questionnaire measures the frequency of everyday cognitive failures or lapses by asking participants to rate how often they make mistakes on a 5-point Likert scale, from 0 (never) to 4 (very often). The instrument produces a global "cognitive complaints" score (ranging from 0-100), with higher scores indicating more cognitive failures.

Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index [30] is commonly used to collect data regarding subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medication, and daytime dysfunction. The seven component scores are rated from 0 to 3, with 0 being no difficulty and 3 indicating severe difficulty. The composite score of the 7 subscale scores provides a global score ranging from 0 to 21, where higher numbers indicated poorer sleep quality.

Short Form 36 Questionnaire

The Short Form Health Survey [31] is a 36-item patient-reported survey of patient health commonly used as a measure of quality of life. The questionnaire was developed at RAND Health as part of the medical Outcomes Study.

Cambridge Neuropsychological Test Automated Battery

Cognitive performance will be assessed at each visit using 8 well-validated tests assessing reaction time, attention, memory, and executive function from the Cambridge Cognition, suitable for people living with ME/CFS. Tests include motor training, simple and 5-choice reaction time, spatial working memory, pattern recognition memory immediate and delayed, one touch stockings of Cambridge, attention switching task, and rapid visual processing. The order that tests are delivered has been determined with key experts in cognitive testing at Cambridge Cognition so longer tests are separated with shorter tests. This will ensure adequate breaks between testing occur. All testing is automated on a tablet to reduce administration error between participants. Furthermore, testing will take place in a small quiet clinic room with a bed available if participant needs to lie down and rest for administration or in between tests. Snacks and water will be available to participants throughout the cognitive testing.

Adverse Events

A comprehensive adverse events reporting protocol has been developed for this study based on the stakeholder group feedback and previous literature. An adverse event (AE) will be defined as "any clinical change, disease, or disorder experienced by the participant during their participation in the trial, whether or not considered related to the intervention studied in the trial."

AE data will collected (1) during every preplanned phone contact between the participant and the research assistant, (2) if the participant contacts the research assistant to report an AE, and (3) during the 3 testing sessions (baseline, 6 months, and 12 months). Questions will be asked about the following 3 categories to encourage the participant to report AEs: any new comorbid medical conditions reported if not previously reported at baseline, any events for which the participant consulted their general practitioner or other medical advisor or took medication, any other events that might have affected the health status of the participant (eg, increased work stress).

The research assistant will have knowledge of participant allocation (ie, control or active intervention group) and as such will be able to use any AE data reported to provide activity counseling and advice regarding the intervention. If an AE is

reported, participants will be asked whether they consider the event to be directly caused by the intervention.

A serious adverse event (SAE) will be defined as one of the following outcomes:

- Death
- Threat to life (ie, an immediate, not hypothetical, risk of death at the time of the event)
- Required hospitalization except for elective treatment of a preexisting condition
- Increased severity and persistent disability, defined as (1) severe (ie, significant) deterioration in the participant's ability to carry out their important activities of daily living (eg, employed person no longer able to work, caregiver no longer able to give care, ambulant participant becoming bed bound) and (2) symptom and disability persistent (ie, of at least 4 weeks continuous duration)
- Any other important medical condition which, though not included in the above, might require medical or surgical intervention to prevent one of the outcomes listed

For any AE established as serious, a medical specialist member of the stakeholder advisory group will be asked to establish whether or not the event was a serious adverse reaction based on treatment allocation information. A serious adverse reaction will be considered to be a reaction to the trial. Both medical specialist and participant attribution will be reported.

A nonserious adverse event (NSAE) will be any health event that was not categorized as an serious AE or serious adverse reaction. NSAEs will be rated as mild, moderate, or severe by both the participant (at the time of reporting) and a medical specialist. The duration of any increases in severity or disability from these events will also be recorded.

Each NSAE will be ascribed a body system (gastroenterological, neurological, etc) by a health practitioner team member. NSAEs attributed to ME/CFS (ie, considered to be a symptom of ME/CFS) will be put into a separate category. To monitor disability, each time the research assistant calls a participant they will be requested to self-report their disability on the Karnofsky disability scale.

Testing Considerations

All equipment and testing facilities will be perfume-free (or as low as possible) as many people with ME/CFS also report multiple chemical sensitivities. Testing facilities will provide footstools and beds to rest. Rest periods will be provided at the request of the participant during the testing sessions to reduce fatigue. Paper questionnaires can be sent home with participants to be completed at their own time and mailed back to spread the effort. Cambridge Neuropsychological Test Automated Battery testing can be paused at one point to allow a rest period. Breakfast/food will be provided after the fasting bloods have been taken.

Analysis Plan

Exit questionnaires will be audiorecorded, transcribed verbatim, and sent back to participants for member checking. Themes will be extracted and summarized from interview transcripts.

Analysis of continuous data to assess preliminary evidence of benefit will include examination of mean differences and confidence intervals ranging from 75% to 95%. Regression analysis will be used to explore preliminary evidence of a relationship between allostatic load and physical activity. The alpha level will be set at 0.2, an appropriate level for a pilot studies [32]. Categorical questionnaire data will be analyzed using chi-square analysis. As this study is a pilot study, any inferential statistics will be purely exploratory (ie, not powered to detect statistical differences).

Results

A 6-month extensive consultative protocol development phase has been completed; 3 comprehensive rounds of protocol topics, comments, and plans for amendment occurred.

The first round focused on the following general topics: inclusion criteria, time frame between effort testing and intervention commencement, time and frequency of assessment, and the inclusion of continuous HR data collection. The Karnofsky Scale and alternative disability scales were discussed. The second round included discussion on the following broad topics: design modification to 3-arm trial, symptom history data collection, and the pacing and physical activity protocol. The third round involved discussion on the following topics: AEs protocol, adherence diary, and clarification of maximal and submaximal effort test procedures.

Virtual meetings allowed for interactive group discussion, and consensus was reached on all aspects of the protocol. All members of the stakeholder advisory group have endorsed the current protocol. Enrollment began in January 2017; as of publication, 12 participants were enrolled. Baseline testing is scheduled to commence in mid-2017.

Discussion

Summary

The proposed research will fill gaps in the existing literature by (1) providing only the second known study of symptom-contingent pacing as a strategy to increase physical activity in adults with ME/CFS, (2) providing the first data on feasibility and acceptability of using active video games as a strategy to increase physical activity in people with ME/CFS, and (3) providing longitudinal data regarding the relationship between allostatic load and physical activity in people living with ME/CFS.

Adults living with ME/CFS often suffer orthostatic intolerance, and as such, a physical activity modality that can be achieved in a sitting or reclining position is advantageous. In addition, many people living with ME/CFS report short bouts of activity to be easier to manage, hence an activity modality that allows short bouts of several minutes to then allows adequate rest addresses the consumer requirements. Active video games such as Nintendo WiiFit and Xbox Kinect have been demonstrated to demand low to moderate levels of physical activity in the elderly, even when sitting [18]. So playing these games while sitting/reclining/lying can still be useful. Active video gaming has been shown to be feasible and acceptable to other chronic

health populations such as those with cerebral palsy and stroke and is commonly played sitting in these populations [17].

Evidence suggests active video gaming can improve the health and well-being of people with a variety of chronic and debilitating health conditions [17]. Despite the evidence to suggest active video gaming as effective and relevant, no studies have investigated the effect, feasibility, or acceptability of active video gaming in adults with ME/CFS.

Conclusion

The results of this trial will provide the feasibility and acceptability data to suggest whether active video gaming can be pursued as a strategy to help people living with ME/CFS overcome physical inactivity. The pilot study will provide important data to guide future effectiveness trials and is intended to deliver safe, effective, evidence-based, consumer-driven physical activity interventions to the ME/CFS community.

Acknowledgments

This project is funded by the Mason Foundation National Medical Program (MAS2015F053).

Authors' Contributions

KEF led the writing of this manuscript and the development of this protocol and successful grant application. AS led the development of the cognitive testing and blood collection components and advised on the remainder. KD led the development of the exercise testing and physiological marker data collection and advised on the remainder. KEF led the Stakeholder Advisory Group process. AS and KD contributed to writing all sections of this manuscript. All authors read and approved the manuscript.

Conflicts of Interest

None declared.

References

1. Clayton EW. Beyond myalgic encephalomyelitis/chronic fatigue syndrome: an IOM report on redefining an illness. JAMA 2015 Mar 17;313(11):1101-1102. [doi: 10.1001/jama.2015.1346] [Medline: 25668027]
2. van der Werf SP, Prins JB, Vercoulen JH, van der Meer JW, Bleijenberg G. Identifying physical activity patterns in chronic fatigue syndrome using actigraphic assessment. J Psychosom Res 2000 Nov;49(5):373-379. [Medline: 11164063]
3. Brenu EW, van Driel ML, Staines DR, Ashton KJ, Ramos SB, Keane J, et al. Immunological abnormalities as potential biomarkers in Chronic Fatigue Syndrome/Myalgic Encephalomyelitis. J Transl Med 2011 May 28;9:81 [FREE Full text] [doi: 10.1186/1479-5876-9-81] [Medline: 21619669]
4. McEwen BS. Stress, adaptation, and disease. Allostasis and allostatic load. Ann N Y Acad Sci 1998 May 01;840:33-44. [Medline: 9629234]
5. Gay JL, Salinas JJ, Buchner DM, Mirza S, Kohl HW, Fisher-Hoch SP, et al. Meeting physical activity guidelines is associated with lower allostatic load and inflammation in Mexican Americans. J Immigr Minor Health 2015 Apr;17(2):574-581 [FREE Full text] [doi: 10.1007/s10903-013-9950-1] [Medline: 24242155]
6. Upchurch DM, Rainisch BW, Chyu L. Greater leisure time physical activity is associated with lower allostatic load in white, black, and Mexican American midlife women: findings from the National Health and Nutrition Examination Survey, 1999 through 2004. Womens Health Issues 2015;25(6):680-687 [FREE Full text] [doi: 10.1016/j.whi.2015.07.002] [Medline: 26344446]
7. Maloney EM, Gurbaxani BM, Jones JF, de Souza Coelho L, Pennachin C, Goertzel BN. Chronic fatigue syndrome and high allostatic load. Pharmacogenomics 2006 Apr;7(3):467-473. [doi: 10.2217/14622416.7.3.467] [Medline: 16610956]
8. Goertzel BN, Pennachin C, de Souza Coelho L, Maloney EM, Jones JF, Gurbaxani B. Allostatic load is associated with symptoms in chronic fatigue syndrome patients. Pharmacogenomics 2006 Apr;7(3):485-494. [doi: 10.2217/14622416.7.3.485] [Medline: 16610958]
9. Powell KE, Paluch AE, Blair SN. Physical activity for health: What kind? How much? How intense? On top of what? Annu Rev Public Health 2011;32:349-365. [doi: 10.1146/annurev-publhealth-031210-101151] [Medline: 21128761]
10. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramírez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. JAMA 2001 Sep 19;286(11):1360-1368. [Medline: 11560542]
11. Castell B, Kazantzis N, Moss-Morris R. Cognitive behavioral therapy and graded exercise for Chronic Fatigue Syndrome: a meta-analysis. Clin Psychol Sci Prac 2011;18:311-324. [doi: 10.1111/j.1468-2850.2011.01262.x]
12. ME Association. ME/CFS illness management survey results: no decisions about me without me, part 1. URL: http://www.meassociation.org.uk/2015/05/23959/ [accessed 2017-05-23] [WebCite Cache ID 6qgYnf4VK]
13. Wallman KE, Morton AR, Goodman C, Grove R, Guilfoyle AM. Randomised controlled trial of graded exercise in chronic fatigue syndrome. Med J Aust 2004 May 03;180(9):444-448. [Medline: 15115421]
14. O'Brien-Cousins S. An older adult exercise status inventory: reliability and validity. J Sport Behav 1996;19(4):288-306.

15. Ekkekakis P. Let them roam free? Physiological and psychological evidence for the potential of self-selected exercise intensity in public health. Sports Med 2009;39(10):857-888. [doi: 10.2165/11315210-000000000-00000] [Medline: 19757863]
16. Verheijden KL, Jaarsma T, Strömberg A. Exergaming in older adults: a scoping review and implementation potential for patients with heart failure. Eur J Cardiovasc Nurs 2014 Oct;13(5):388-398 [FREE Full text] [doi: 10.1177/1474515113512203] [Medline: 24198306]
17. Staiano AE, Flynn R. Therapeutic uses of active videogames: a systematic review. Games Health J 2014 Dec;3(6):351-365 [FREE Full text] [doi: 10.1089/g4h.2013.0100] [Medline: 26192642]
18. Peng W, Crouse JC, Lin J. Using active video games for physical activity promotion: a systematic review of the current state of research. Health Educ Behav 2013 Apr;40(2):171-192. [doi: 10.1177/1090198112444956] [Medline: 22773597]
19. Yuen HK, Holthaus K, Kamen DL, Sword DO, Breland HL. Using Wii Fit to reduce fatigue among African American women with systemic lupus erythematosus: a pilot study. Lupus 2011 Oct;20(12):1293-1299 [FREE Full text] [doi: 10.1177/0961203311412098] [Medline: 21700656]
20. Pope Z, Zeng N, Gao Z. The effects of active video games on patients' rehabilitative outcomes: a meta-analysis. Prev Med 2017 Feb;95:38-46. [doi: 10.1016/j.ypmed.2016.12.003] [Medline: 27939260]
21. Sports Medicine Australia pre-exercise screening system. URL: https://www.essa.org.au/for-gps/adult-pre-exercise-screening-system/ [accessed 2016-12-31] [WebCite Cache ID 6n9Z2QUi2]
22. Australia's Physical Activity and Sedentary Behaviour Guidelines. URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-phys-act-guidelines [accessed 2016-12-30] [WebCite Cache ID 6n9UsbHNC]
23. Davenport TE, Stevens SR, VanNess MJ, Snell CR, Little T. Conceptual model for physical therapist management of chronic fatigue syndrome/myalgic encephalomyelitis. Phys Ther 2010 Apr;90(4):602-614 [FREE Full text] [doi: 10.2522/ptj.20090047] [Medline: 20185614]
24. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol 2001 Jan;37(1):153-156 [FREE Full text] [Medline: 11153730]
25. Karnofsky DA, Burchenal JH. The clinical evaluation of chemotherapeutic agents in cancer. In: MacLeod CM, editor. Evaluation of Chemotherapeutic Agents. New York: Columbia University Press; 1949.
26. World Health Organization. International Classification of Functioning, Disability, and Health. 2001. URL: http://www.who.int/classifications/icf/en/ [accessed 2017-05-23] [WebCite Cache ID 6qgZNOc2g]
27. Ridley K, Olds TS, Hill A. The Multimedia Activity Recall for Children and Adolescents (MARCA): development and evaluation. Int J Behav Nutr Phys Act 2006 May 26;3:10 [FREE Full text] [doi: 10.1186/1479-5868-3-10] [Medline: 16725055]
28. Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, Wright D, et al. Development of a fatigue scale. J Psychosom Res 1993;37(2):147-153. [Medline: 8463991]
29. Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The Cognitive Failures Questionnaire (CFQ) and its correlates. Br J Clin Psychol 1982 Feb;21 (Pt 1):1-16. [Medline: 7126941]
30. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [Medline: 2748771]
31. Ware JE, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. J Clin Epidemiol 1998 Nov;51(11):903-912. [Medline: 9817107]
32. Lee EC, Whitehead AL, Jacques RM, Julious SA. The statistical interpretation of pilot trials: should significance thresholds be reconsidered? BMC Med Res Methodol 2014 Mar 20;14:41 [FREE Full text] [doi: 10.1186/1471-2288-14-41] [Medline: 24650044]

Abbreviations

AE: adverse event
AL: allostatic load
BMI: body mass index
CRP: C-reactive protein
GET: Graded Exercise Therapy
HDL: high-density lipoprotein
HR: heart rate
HRV: heart rate variability
LDL: low-density lipoprotein
ME/CFS: myalgic encephalomyelitis/chronic fatigue syndrome
NSAE: nonserious adverse event
RPE: ratings of perceived exertion
VT: ventilatory threshold

Edited by G Eysenbach; submitted 30.12.16; peer-reviewed by B Cao, M Vollenbroek; comments to author 22.03.17; revised version received 24.04.17; accepted 24.04.17; published 01.08.17.

Please cite as:
Ferrar KE, Smith AE, Davison K
Pacing, Conventional Physical Activity and Active Video Games to Increase Physical Activity for Adults with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2017;6(8):e117
URL: http://www.researchprotocols.org/2017/8/e117/
doi:10.2196/resprot.7242
PMID:28765100

Protocol

A Web-Based Public Health Intervention to Reduce Functional Impairment and Depressive Symptoms in Adults With Type 2 Diabetes (The SpringboarD Trial): Randomized Controlled Trial Protocol

Judy Proudfoot[1,2*], PhD; Janine Clarke[1*], MPsychol (Clin), PhD; Jane Gunn[3*], MD, PhD; Susan Fletcher[3*], PhD; Samineh Sanatkar[1], PhD; Kay Wilhelm[2], MBBS, MD, FRANZCP; Lesley Campbell[4], MBBS, FRCP (UK), FRACP AM; Nicholas Zwar[5], MBBS, PhD; Mark Harris[6], MBBS, PhD; Helen Lapsley[2], PhD; Dusan Hadzi-Pavlovic[2], BSc, MPsychol; Helen Christensen[1,2], PhD

[1]Black Dog Institute, Randwick, Australia

[2]School of Psychiatry, University of New South Wales (UNSW) Sydney, Sydney, Australia

[3]Department of General Practice, University of Melbourne, Melbourne, Australia

[4]Diabetes and Metabolism Division, Garvan Institute of Medical Research, Sydney, Australia

[5]School of Medicine, University of Wollongong, Wollongong, Australia

[6]Centre for Primary Health Care and Equity, University of New South Wales (UNSW) Sydney, Sydney, Australia

[*]these authors contributed equally

Corresponding Author:
Janine Clarke, MPsychol (Clin), PhD
Black Dog Institute
Hospital Road
Randwick, 2031
Australia
Phone: 61 2 9382 3767 ext 23767
Fax: 61 2 9382 8208
Email: janine.clarke@unsw.edu.au

Abstract

Background: Depressive symptoms are common in people with type 2 diabetes and contribute to adverse health consequences that substantially impact social and vocational function. Despite the existence of effective depression treatments, the majority of people with type 2 diabetes do not access these when needed. Web-based alternatives to more traditional psychotherapies offer a potential solution to reducing the personal and economic burdens of type 2 diabetes.

Objective: This paper outlines the protocol for a randomized controlled trial (RCT) of myCompass, a Web-based public health psychotherapy intervention, in people with type 2 diabetes. Fully automated, interactive, and delivered via the Internet without clinician support, myCompass teaches cognitive behavioral therapy-based skills and supports symptom monitoring to improve daily functioning and reduce mild-to-moderate mental health symptoms.

Methods: A two-arm RCT will be conducted. People with type 2 diabetes and mild-to-moderately severe depressive symptoms will be recruited from the community and general practice settings. Screening and enrollment is via an open-access website. Participants will be randomized to use either myCompass or an active placebo program for 8 weeks, followed by a 4-week tailing-off period. The placebo program is matched to myCompass on mode of delivery, interactivity, and duration. Outcomes will be assessed at baseline and at 3-month, 6-month, and 12-month follow-up. The primary study outcome is work and social functioning. Secondary study outcomes include depressive and anxious symptoms, diabetes-related distress, self-care behaviors, and glycemic control.

Results: Nationwide recruitment is currently underway with the aim of recruiting 600 people with type 2 diabetes. Recruitment will continue until October 2017.

Conclusions: This is the first known trial of a Web-based psychotherapy program that is not diabetes specific for improving social and vocational function in people with type 2 diabetes and mild-to-moderately severe depressive symptoms. With the

increasing prevalence of type 2 diabetes and depression, a potentially scalable public health intervention could play a very large role in reducing unmet mental health need and ameliorating the personal and societal impact of illness comorbidity.

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR) Number: ACTRN12615000931572; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=368109 (Archived by WebCite at http://www.webcitation.org/6rh3imVMh)

(JMIR Res Protoc 2017;6(8):e145) doi:10.2196/resprot.7348

KEYWORDS

type 2 diabetes; depression; Web-based intervention

Introduction

Background

Type 2 diabetes is a common chronic and disabling disease and a major contributor to global disease burden [1]. Depression is frequently comorbid with type 2 diabetes and contributes independently to a range of adverse health outcomes that substantially compromise social and vocational functioning. These include greater diabetes symptom burden, poorer self-care, and increased risk of micro- and macrovascular complications and mortality [2]. In addition, the economic impact of disease comorbidity is considerable with functional disability, functional dependence, workplace productivity losses, health service use, and health care costs higher for people with both conditions than those with diabetes alone [2,3].

Research supports a bidirectional relationship between type 2 diabetes and depression [4]. Findings supporting increased risk of type 2 diabetes in people with depression [5] are generally explained with reference to physiological (eg, hypothalamic-pituitary-adrenal axis dysfunction [6]), motivational (eg, poorer self-care, adiposity, and inactivity [6,7]), and/or pharmacologic (eg, impact of antidepressant medication on glycemic control [5]) factors. At the same time, increased risk of depression in people with type 2 diabetes is generally attributed to the emotional burden of living with a complex and demanding disease that often intrudes into normal lifestyle [8]. The likelihood of a reciprocal relationship between type 2 diabetes and depression makes the personal and societal impact of illness comorbidity potentially immeasurable [4].

Despite the existence of evidence-based therapies for depression in diabetes, including cognitive behavioral therapy (CBT) and antidepressant medication, upward of 60% of people with comorbid conditions do not receive mental health treatment [9]. In the primary health care setting, where most people with type 2 diabetes access medical treatment, low screening rates mean that depressive symptoms are often missed [9] and only a minority of patients who screen positive accept a referral for face-to-face support [10]. At the same time, personal, social, and structural barriers to seeking help—including privacy concerns and stigma, lack of support and poor relationships with health care providers, time and lifestyle constraints, financial cost, and lack of service availability—compromise access to satisfactory mental health care for many patients [11,12]. There is considerable opportunity, therefore, to reduce the personal and societal burden of illness by facilitating greater access to more flexible and cost-efficient mental health treatments for this patient group.

Internet delivery of evidence-based psychological therapies is a popular, clinically effective, and cost-effective means of increasing treatment reach; a small number of Web-based diabetes-specific interventions that directly target depression have been evaluated [13]. These include van Bastelaar et al's [14] adaptation of Lewinsohn's Coping with Depression Program and Nobis et al's [15] application of systematic behavioral activation and problem-solving therapy (ie, GesundheitsTraining.Online Mood Enhancer Diabetes [GET.ON MED]). Findings from these studies establish the symptom benefits of Internet-delivered self-help for diabetes patients with clinically relevant levels of depressive symptomatology. However, subthreshold depression is more prevalent in type 2 diabetes than clinical depression and is associated with increased functional limitation and disability, including reduced social and vocational performance [16,17]. As such, the effectiveness of electronically delivered psychotherapy as a treatment approach for type 2 patients with mild-to-moderate depressive symptoms also warrants rigorous scientific attention.

Proudfoot et al [18] have previously published controlled data showing that mild-to-moderate mental health symptoms—including symptoms of depression, anxiety, and stress—are reduced following use of the broadly available Web-based program, myCompass, with treatment benefits extending to work and social functioning. More recently, in an uncontrolled pilot study, myCompass showed promise as an acceptable and effective treatment for depression and functional disability in people with type 1 and type 2 diabetes [19]. myCompass differs from the interventions described in van Bastelaar et al [14] and Nobis et al [15] in that it is a fully automated (ie, no therapist input) public health program that is generic in its therapeutic content (ie, not diabetes sensitive). An intervention approach that is capable of treating depressive symptoms without disease-specific modification is potentially a more efficient and accessible alternative to meeting the unmet mental health needs of people with type 2 diabetes. Generic skills may also assist the increasing number of individuals experiencing multi-morbidity, for whom depression co-occurs with somatic symptoms of multiple illnesses (eg, diabetes, heart disease, hypertension, and kidney disease) [20,21]. From a primary care perspective, a public health program may have important pragmatic advantages in the primary care setting, where time pressures often impede dissemination of, and prohibit practitioner training in, multiple disease-specific tools, and

where treatment of multi-morbidity and undifferentiated physical and mental health symptoms is particularly relevant.

Objectives

This paper was prepared using the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidance for presenting clinical trial protocols [22] and the Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth (CONSORT-EHEALTH) checklist [23]. The paper describes a randomized controlled trial (RCT) of the Web-based intervention, myCompass, in patients with comorbid type 2 diabetes and mild-to-moderately severe depressive symptoms. The primary aim of this RCT, called SpringboarD, is to determine whether treatment with myCompass improves work and social function for people with type 2 diabetes and mild-to-moderate depressive symptoms. As functional disability predicts further functional deterioration, functional dependence, and increased use of health services and health care costs [16], depression treatments that improve work and social functioning may substantially reduce the personal and economic burden of comorbid depression and type 2 diabetes. Specifically, we hypothesize that the intervention group will show significant improvement in self-reported functioning socially and in the workplace compared with a placebo-controlled condition.

Our secondary aim is to evaluate the impact of myCompass on a range of symptom- and disease-related variables known to impact a patient's disease management and blood glucose control. Specifically, in addition to depressive symptoms, we will examine whether myCompass is more effective than a placebo condition in improving anxiety symptoms and diabetes-related distress. Defined as a person's emotional adjustment to the chronic stressors specific to diabetes, diabetes-related distress is of particular interest due to its greater prevalence and potential role in mediating the relationship between depression and glycemic control [24]. While we have preliminary evidence suggesting that diabetes-related distress may improve following treatment with myCompass [19], replication of this finding in a controlled study is required.

Trial Design

The study is designed as a two-arm individually randomized RCT and is conducted entirely online. Participants allocated to the intervention group have full access to the myCompass program for 8 weeks, followed by a 4-week tailing-off period to facilitate maintenance. Participants randomized to the control group have access to a placebo Internet-delivered program that is matched to myCompass on mode and duration of delivery and interactivity, but is void of therapeutic content. Participants in both groups have uninterrupted access to treatment as usual for their diabetes during the study.

Methods

Participants, Interventions, and Outcomes

Study Setting

The setting for this study is Australia. An estimated 1 million Australian adults—5% of the population—had self-reported type 2 diabetes in September 2016, and rates are similar across metropolitan, regional, and remote areas [25].

Eligibility Criteria

This study focuses on adults with type 2 diabetes and mild-to-moderately severe depressive symptoms. People are eligible to take part if they are aged 18-75 years, screen positive for depression on the 2-item Patient Health Questionnaire (PHQ-2) (ie, score ≥2) [26], and have access to an Internet-enabled device (eg, computer, tablet, and/or mobile phone). People who screen positive for depression complete the full 9-item Patient Health Questionnaire (PHQ-9) at screening so that the level of symptom severity can be determined.

Exclusion criteria include inability to read English with ease, severe depressive symptoms on the full PHQ-9 (ie, score >19), probable psychosis as measured by the psychosis screener developed for the Australian National Mental Health and Wellbeing Survey [27], currently receiving face-to-face counseling or therapy for depression, changed antidepressant medication in the previous 2 months, high suicide risk as assessed by item 9 of the PHQ-9, and previous use of the myCompass program.

Interventions

Active Intervention

The active intervention, myCompass [28], is a fully automated, self-help, public health intervention that is tailored to the user and has no therapist involvement in its delivery. Program tailoring occurs via users' responses to a symptom profiler completed at registration. In-built algorithms target the user's most salient symptoms and provide recommendations about the symptoms and/or behaviors they might consider monitoring and the treatment modules likely to be of greatest therapeutic benefit. There is flexibility, however, for users to choose their own set of self-monitoring dimensions and treatment modules (see Figures 1-4).

Self-monitoring of symptoms and lifestyle factors occurs in real time via mobile (eg, phone or tablet) and stationary (eg, desktop computer) devices. Users can self-monitor up to three symptoms and behaviors of their choice at any one time—selected from a list of 20—or three that are recommended to them by the program. Each symptom is rated on a 10-point scale. At the time of rating, users also provide contextual information about where they are, what they are doing, and who they are with using a series of drop-down menus. Users can schedule short message service (SMS) text message or email reminders to facilitate self-monitoring—frequency of reminders is determined by the user; receive and print graphical feedback about their monitoring, including contextual information, on their phone or computer to monitor change and assist identification of triggers; and elect to receive helpful facts, mental health care tips, or motivational statements by SMS text message or email.

myCompass treatment modules were developed by mental health professionals at the Black Dog Institute in Sydney, Australia. The program contains 14 skill-building modules derived from CBT, interpersonal psychotherapy, problem-solving therapy,

and positive psychology that are interactive and available for completion on computer and tablet devices. Module content covers topics such as *Managing Fear and Anxiety*, *Tackling Unhelpful Thinking*, *Managing Loss and Major Life Change*, and *Solving Problems*. Each module comprises three 10-minute sessions and includes home practice tasks for completion between the weekly sessions to promote skill generalization. The module targeting stress in diabetes has been deactivated for trial participants, to ensure that the intervention is a generic public health program. Participants are encouraged to complete at least three modules during the intervention period, during their own time.

User privacy is managed by a password-protected log-on and by ensuring that user-generated data (ie, self-monitoring ratings) are not stored on the user's phone, but are instead transferred via the Internet using Secure Sockets Layer protocol, which encrypts transmitted data by rendering it unreadable to anyone other than the intended recipient, and by storing the data in secure servers. The data is reidentifiable only with the list of study participant codes, to which only the named researchers will have access and which will be stored in a password-protected file separate from the study data. Registering to use myCompass is free and users are not billed for the SMS text messages they receive.

Participants randomized to myCompass have access to the full program for 8 weeks, followed by a 4-week tailing-off period during which only the symptom-monitoring function will be accessible. Research has shown that adherence with online interventions is improved when users receive program feedback that is personalized in its content [29]. For this reason, myCompass users receive automated personalized feedback via email about their use of the program's self-monitoring and module functions at fixed intervals (ie, weeks 1, 3, 5, and 7).

Figure 1. Screenshot of myCompass landing page.

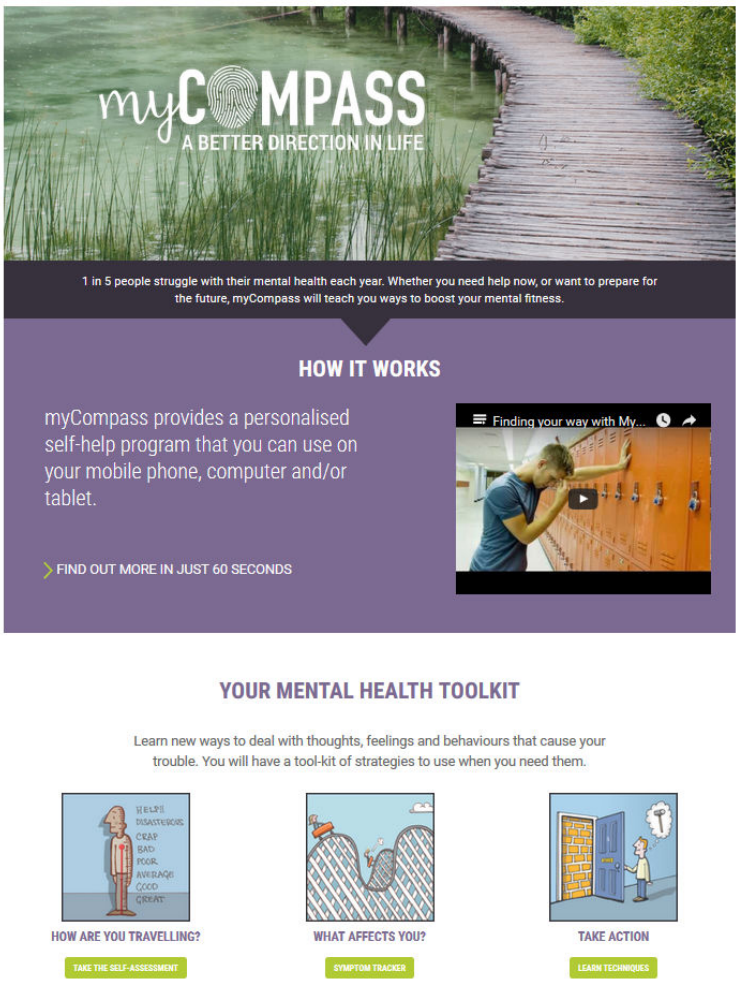


Figure 2. Screenshot of myCompass self-monitoring page.

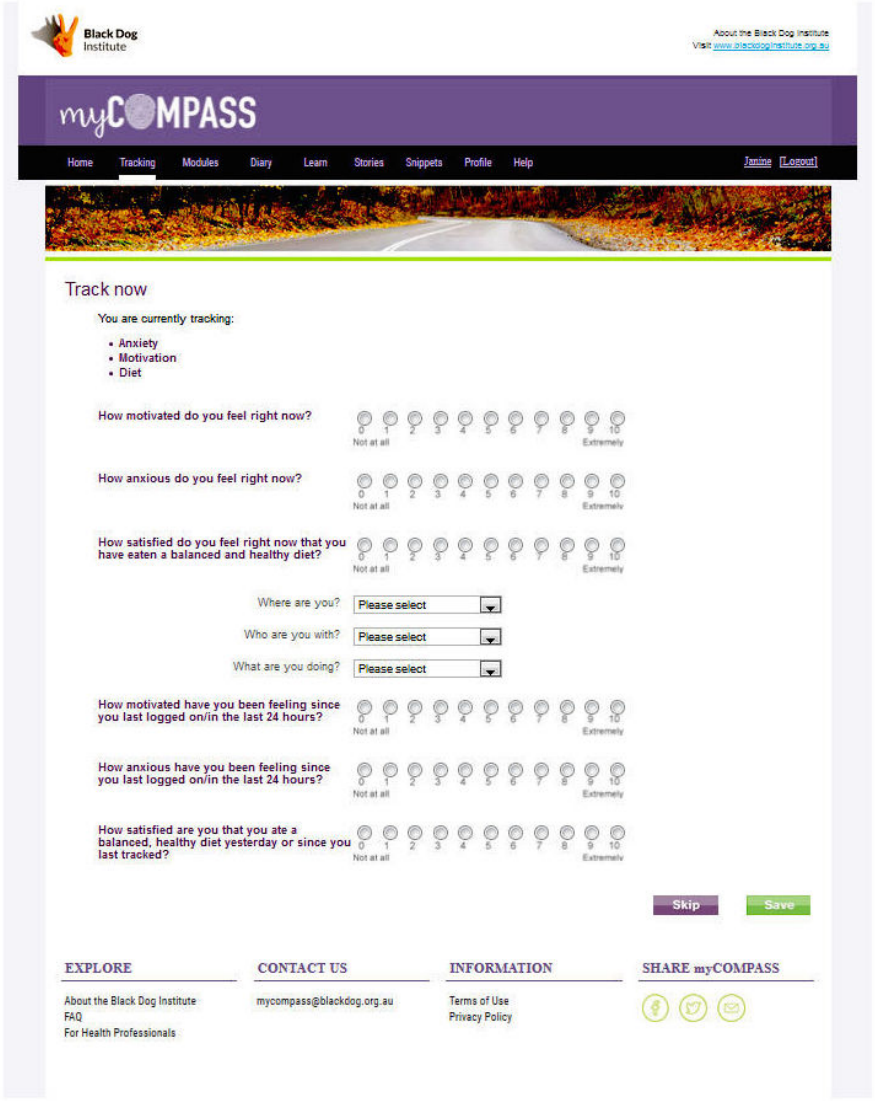
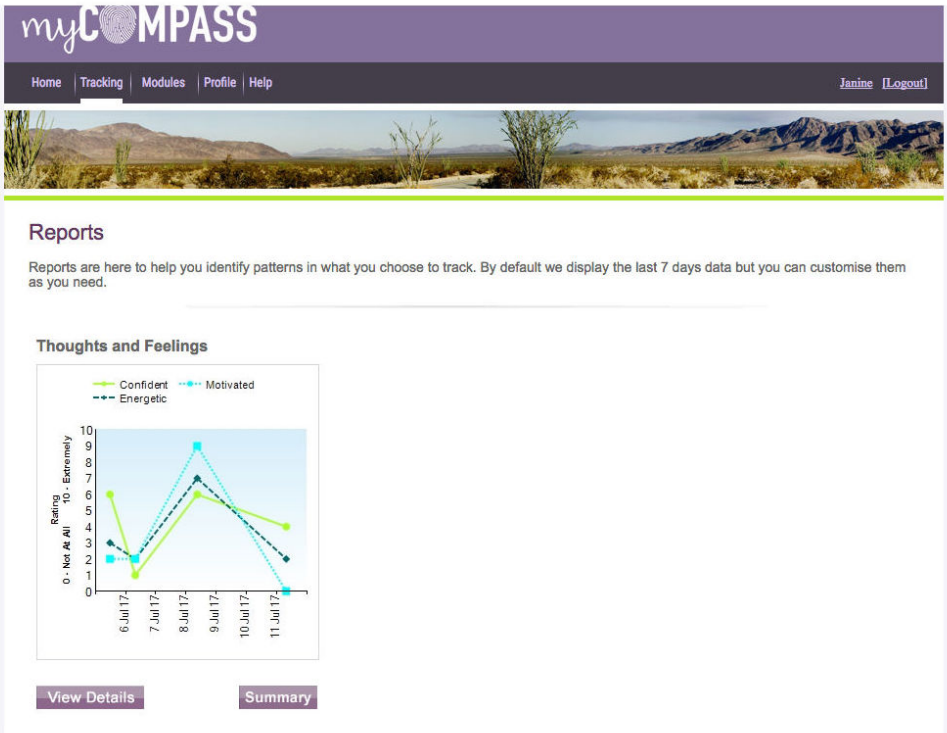


Figure 3. Screenshot of a page from the myCompass module Tackling Unhelpful Thinking.

Figure 4. Screenshot of myCompass graphical feedback.



Placebo Intervention

The placebo program, Healthy Lifestyles, was adapted from a control program used in previous studies by members of the research team. Like myCompass, the program offers users program tailoring at the outset, followed by access to a range of interactive modules containing health- and lifestyle-related information, including skin care and eye health. It contains no therapeutic content, has high face validity as a health and lifestyle intervention, and has previously shown no symptom impact [30]. Participants have full access to the placebo condition for 8 weeks, plus a 4-week tailing-off period (see Figures 5-7).

To further replicate the interactivity of myCompass, users of the placebo program receive an email at weeks 1, 3, 5, and 7 containing a brief reminder to log into the program, but no feedback about their use of the program. They also receive a weekly email containing a factual statement about a health or lifestyle issue.

Figure 5. Screenshot of landing page for placebo program, Healthy Lifestyles.

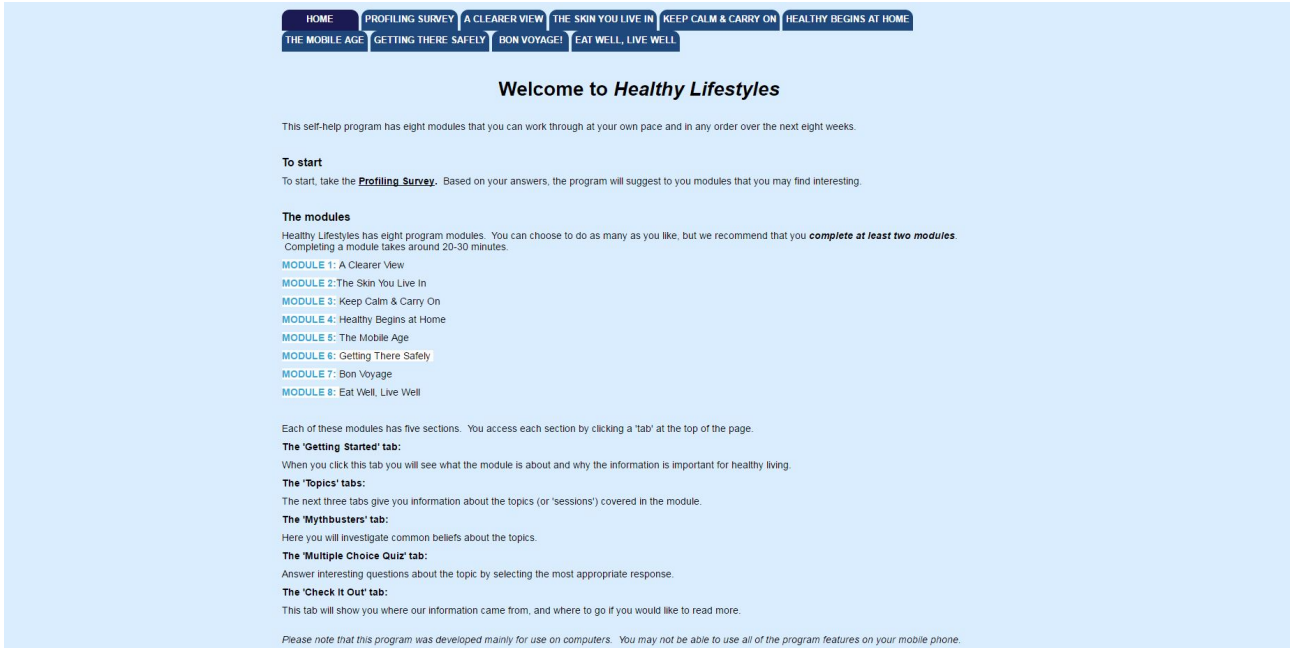


Figure 6. Page from the Healthy Lifestyles module, The Skin You Live In.

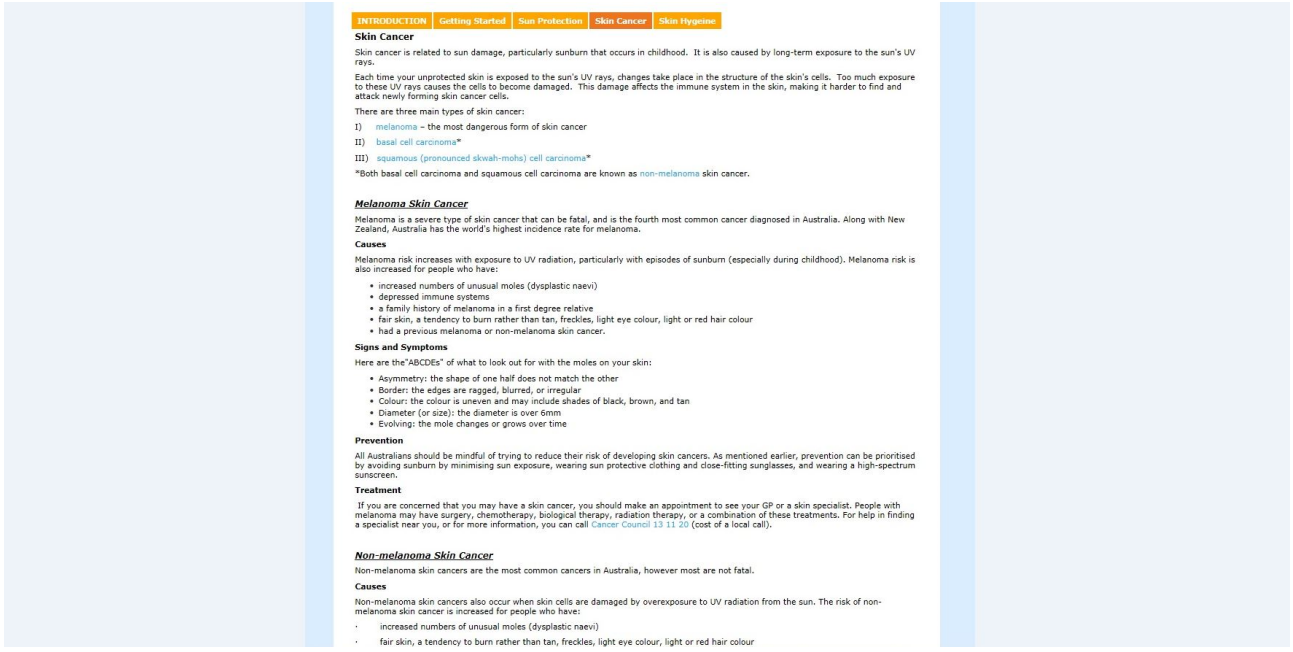


Figure 7. Sample interactive task from the Healthy Lifestyles module.

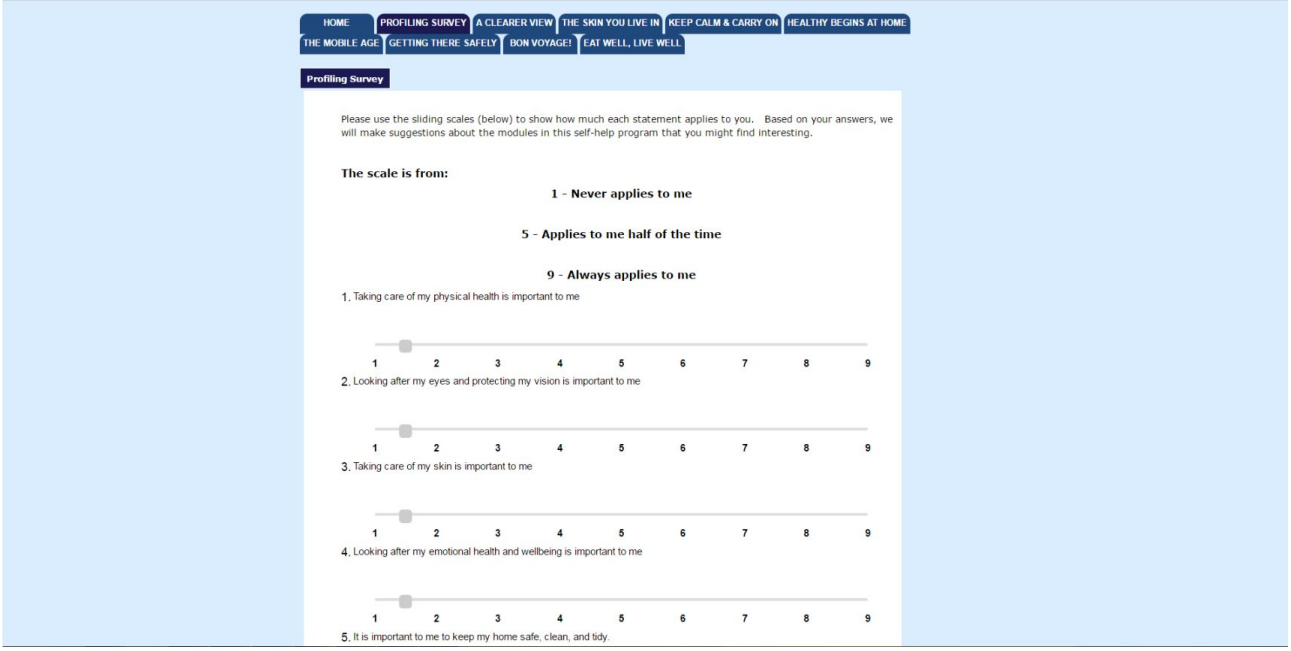


Table 1. Assessments completed at assessment phases.

Assessments		Assessment phase			
		T1[a]	T2[b]	T3[c]	T4[d]
Baseline					
	Demographic data	X			
	Disease/treatment data	X	X	X	X
	Mental health history	X			
Primary outcome	Work and Social Adjustment Scale [31]	X	X	X	X
Secondary outcomes					
	Patient Health Questionnaire-9 [26]	X	X	X	X
	Diabetes Distress Scale [32]	X	X	X	X
	Generalized Anxiety Disorder Scale [33]	X	X	X	X
	Self-Management Profile for Type 2 Diabetes Scale (behavior items only) [34]	X	X	X	X
	Glycosylated hemoglobin (HbA1c)	X		X	X
	Days out of role [35]	X	X	X	X
	Health service utilization	X	X	X	X

[a]T1: baseline assessment and allocation to intervention or placebo group.

[b]T2: completion of intervention period and online postintervention assessment.

[c]T3: completion of 6-month follow-up assessment.

[d]T4: completion of 12-month follow-up assessment.

Outcomes

A summary of assessments completed online by participants at baseline, postintervention, and follow-up is presented in Table 1 [26,31-35].

The primary outcome is work and social functioning, which is measured by the Work and Social Adjustment Scale (WSAS) [31]. The WSAS is a psychometrically sound measure of the impact of mental health problems on daily functioning in five domains: work, social leisure activities, private leisure activities, home management, and personal relationships [31,36]. Scores range from 0 to 40, with higher scores indicating poorer adjustment.

Secondary outcomes include the PHQ-9, the Diabetes Distress Scale [32], the 7-item Generalized Anxiety Scale [33], and the Self-Management Profile for Type 2 Diabetes Scale [34].

At baseline and again at 6- and 12-month follow-up, glycosylated hemoglobin (HbA1c) data will also be collected

from participants' medical records, with their consent. The HbA1c test is a blood test showing a person's average blood glucose level over the previous 3 months and is measured as part of routine clinical care to monitor long-term blood sugar control in people with diabetes. In all cases, the most recent result will be obtained.

Additional measurements include collection of disease-related data (eg, age of onset and treatment regimen), sociodemographic data (eg, age, gender, education, and occupation), and mental health history data (eg, service use and previous diagnoses) at baseline. Service utilization and supports, including medication, received for problems related to mental health and diabetes are also assessed at each assessment point. Days out of role are measured using an item from the World Health Organization Disability Assessment Schedule that asks people to note the number of days in the previous 30 that they were completely unable to perform their work or normal activities because of problems with their physical or mental health [35].

At the conclusion of the trial, data indicating the extent of participant engagement with myCompass will be extracted from the program, including frequency of log-ins, number of modules started and completed, and self-monitoring frequency.

Participant Timeline

Participant consent, screening, and assessment takes place online via a secure study-specific website [37]. After providing informed consent, potential participants complete the online screening tool to determine eligibility. Unsuitable candidates receive automated feedback explaining the reason for their ineligibility; details of crisis supports and other self-help and face-to-face resources are provided to individuals, as appropriate.

Eligible candidates are registered in a secure Web-based platform where they complete the baseline questionnaire and are immediately allocated randomly to either the intervention program or control program for a period of 12 weeks. Subsequent assessment points coincide for both groups at postintervention (12 weeks) as well as 6 and 12 months after randomization (see Figure 8).

Sample Size

The primary study outcome is work and social functioning, which is measured by the WSAS [31]. In a previous study comparing myCompass with an active control intervention, Proudfoot et al [18] obtained an average between-group effect size—based on estimated marginal means—of Cohen d=0.3 for work and social functioning. We calculated the sample size for this study using a series of t tests (cross-sectional comparison of arms) with alpha=.05 and desired power=80%. The sample size required was N=350 (175 per arm).

We cross-checked the sample size calculation using a reduction of 5 points on our secondary outcome, the PHQ-9, and got a similar outcome. A 5-point reduction on the PHQ-9 is considered an adequate treatment response [26].

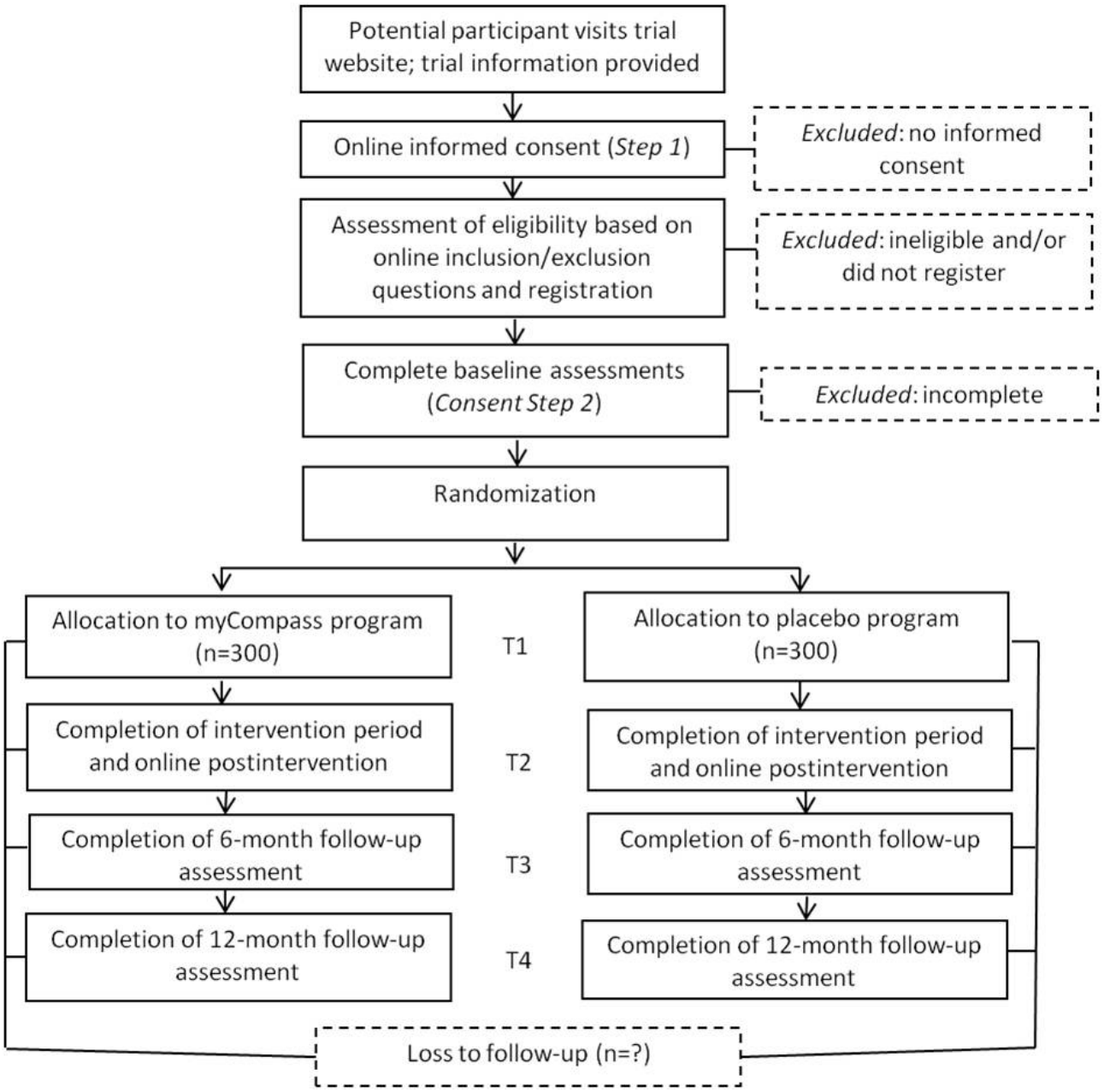
Previous studies indicate attrition rates of approximately 40% in eHealth studies, generally [38]. As such, we aim to recruit 300 participants in each arm of the study (ie, 600 in total) for sufficient statistical power for completer analyses.

Recruitment

Study participants are being recruited via promotional materials distributed in general practice settings in New South Wales and Victoria—where the majority of Australians with type 2 diabetes reside—and disseminated nationally via print advertisements; social media posts, including Facebook and Twitter; clinical research registries; and other publicity channels of state and local diabetes stakeholder groups and the Black Dog Institute.

Promotional materials invite interested candidates to visit the SpringboarD Project website [37] to provide consent and complete the screening tool. The recruitment message for the study focuses on learning new ways to deal with stress and distress and live active and emotionally healthier lives with type 2 diabetes. Recruitment has commenced and will continue until October 2017 or until our sample target is reached.

Figure 8. SpringboarD participant flow diagram.



Assignment of Interventions

Allocation

We are using computerized blocked randomization with blocks of eight to assign participants to the two treatment conditions. Randomization to the intervention and placebo program occurs immediately after a participant completes the baseline measures using an automated system built into the study software. In this way, the allocation sequence is applied without the researchers' knowledge.

Blinding

The placebo program has been developed to replicate the mode of delivery, interactivity, and duration of myCompass and participants in each group will be treated equally by the research team. It is not possible, however, to assume that participants will remain blind to study allocation during the intervention and follow-up periods.

Data Collection, Management, and Analysis

Data Collection

The majority of primary and secondary outcome data is being collected electronically via standardized self-report questionnaires that are completed by logging into the SpringboarD study website (described above). At each assessment point, the website sends a unique link to the study questionnaire via email. Questionnaire data is maintained on a secure server at the University of New South Wales (UNSW) Sydney and is downloaded periodically for storage in a password-protected data file accessible by two project personnel (JC and SS). HbA1c data is collected from each participant's general practitioner (GP) via phone, mail, and/or facsimile.

Retention

Participant attrition is a major concern in studies of unguided interventions [38]. To facilitate trial retention, we are utilizing a combination of strategies covering various themes that have been used with success in previous studies. First, a systematic schedule of *push* messages, including personalized email and telephone prompts and reminders, is being implemented to motivate questionnaire completion at each of the postintervention and follow-up assessments [39]. Second, Aus $30 gift vouchers are provided to participants who return completed questionnaires at each time point to compensate them for the time and effort they have expended [40]. Third, a sense of project identity is maintained by using a study logo and design template (ie, set fonts, formatting, and colors) to create the SpringboarD *brand*, and by distributing quarterly newsletters to participants to update them on trial progress [41]. Finally, a phone call from the research team within a week of registration provides participants a minimum level of human contact and the team an opportunity to express thanks, provide encouragement, and confirm personal details [42].

Statistical Analysis

Participants in the intervention and control groups will be compared at baseline using chi-square tests for categorical data and *t* tests for continuous data to assess randomization success. Treatment effects on primary and secondary outcomes will be evaluated by intention-to-treat analysis using mixed-models repeated measures (MMRM). In MMRM, no participant is removed from the analysis because all available data are used to obtain parameter estimates. Effect size will be measured using Cohen *d*. For all outcome measures, within- and between-group differences will be standardized to Cohen *d* using the pooled standard deviation of the observed scores obtained at baseline. We plan to analyze contrasts between intervention and control groups at postintervention and at 3-month, 6-month, and 12-month follow-up.

Supplementary analyses will use data for completers and will also investigate whether there are any differences by recruitment source, duration of diabetes, and presence of comorbid conditions.

Monitoring

The integrity of the trial, including data collection and monitoring, trial progress, adverse events, and compliance with UNSW Sydney Human Research Ethics Committee (HREC) reporting procedures, is overseen by the Steering Committee consisting of the chief (JP) and associate investigators. The Steering Committee will meet biannually over the lifespan of the project. Adverse events may include unfavorable changes to mental health or diabetes control and may be related or unrelated to the study. As the study does not impact routine diabetes care and is examining the effect of an evidence-based intervention for people with mild-to-moderately severe depression (ie, serious mental illness is an exclusion criteria), no serious adverse events are anticipated and no interim analyses are planned.

Ethics and Dissemination

Research Ethics Approval

The SpringboarD study protocol and materials have been approved by the HREC at UNSW Sydney and registered with the Australia and New Zealand Clinical Trials Register (ACTRN12615000931572). Annual reports and substantive amendments to this protocol will be submitted to the HREC for approval by the chief investigator. The study coordinator (JC) is responsible for communicating protocol changes to relevant stakeholders, including the Australian New Zealand Clinical Trials Registry.

Consent or Assent

Information about the study is provided on the SpringboarD project website; individuals can choose to read the information online or download a PDF to keep. Consent is obtained online in a two-stage process. First, individuals consent to the study by checking a box at the end of the study information page and progressing to the page containing the eligibility screen. Eligible individuals are then provided the option of registering an account with the study website (ie, username and password); those who opt not to register are also considered to have not consented to the trial.

Participants consent to the project team informing their treating GP of his or her involvement in the study to facilitate HbA1c data collection; they also provide a point of emergency contact should a participant score in the severely distressed range of the PHQ-9 (ie, score >19) or be at risk of self-harm (ie, score 3 on item 9 of the PHQ-9) at any assessment point. GPs are informed by mail within 2 weeks of their patient's enrolment, at which time they are requested to inform the research team if a diagnosis of type 2 diabetes has not been given.

Confidentiality

The eligibility screen is conducted anonymously such that no personal information about potential participants is collected. Only eligible individuals provide identifying information that is downloaded and stored separately from study data in a password-protected file.

Ancillary and Posttrial Care

At the conclusion of the trial, the active intervention will be made available to all participants in the control group.

Results

Nationwide recruitment is currently underway with the aim of recruiting 600 people with type 2 diabetes. Recruitment will continue until October 2017.

Discussion

Principal Findings

Treatment of depressive symptoms in people with type 2 diabetes might help to improve short- and long-term social and vocational functioning. Internet-delivered psychotherapy is an effective treatment for depression in people with type 2 diabetes; however, few studies have focused on mild-to-moderately severe

depressive symptoms where treatment need in diabetes patients is greatest. Rarely has social and vocational function been evaluated in studies of online depression treatments in type 2 diabetes; the effectiveness of a generic Internet-based program has not been studied in this patient group. This study will shed light on whether an Internet-delivered public health program has the potential to reduce unmet treatment need and lessen the personal and societal impact of mild-to-moderately severe depressive symptoms within the context of type 2 diabetes.

Limitations

Trials of self-guided interventions frequently report high rates of attrition, including study dropout (ie, questionnaire noncompletion) and/or disengagement from the program [38]. Study attrition introduces selection bias and potential misrepresentation of treatment effects. To reduce the impact of actual participant dropout, we will recruit a substantially larger sample than is required on the basis of our sample size calculation. To minimize potential study dropout, we will utilize a combination of strategies that have been shown to positively impact retention in previous trials, including email and telephone prompts and reminders [39], recompense for questionnaire completion [40], and activities aimed at keeping the study "front of mind" and engaging for older participants [41,43]. In addition, so that questionnaire completion is not contingent upon program use, participants will access the study questionnaire outside the intervention (ie, via a link sent to a nominated email account). To maximize program use, automated program reminders will be sent by email to participants at biweekly intervals [29].

Volunteer bias is another possible weakness of this study, as our recruitment techniques may yield a sample that is healthier [44] and more highly motivated to learn new skills and engage with self-guided therapy. Problems occur, for example, if participation in our study is reflective of better health and a broader personal commitment to self-improvement of diabetes outcomes and daily functioning, as this is likely to result in within-group changes that are larger than for a less motivated and more representative sample of type 2 diabetes patients. However, since myCompass is broadly available for self-referral and use whenever and wherever people choose, the recruitment processes are consistent with both the self-help nature of the intervention and its eventual use in the type 2 diabetes population.

Conclusions

The personal and economic costs of comorbid type 2 diabetes and depressive symptoms are substantial. However, the development of Internet-delivered interventions offers a potential solution to mitigate these impacts [13-15]. myCompass is a broadly available and efficacious public health intervention that can be delivered at minimal cost [45]. It therefore presents itself as a potentially effective and timely option for reducing unmet mental health need and ameliorating the personal and societal impact of co-occurring depression and type 2 diabetes.

Acknowledgments

The authors acknowledge support from the Australian Government's National Health and Medical Research Council (grant No. 1083116).

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.1 [23].

[PDF File (Adobe PDF File), 7MB - resprot_v6i8e145_app1.pdf]

References

1. Roy T, Lloyd CE. Epidemiology of depression and diabetes: A systematic review. J Affect Disord 2012 Oct;142 Suppl:S8-S21. [doi: 10.1016/S0165-0327(12)70004-6] [Medline: 23062861]
2. Katon WJ. The comorbidity of diabetes mellitus and depression. Am J Med 2008 Nov;121(11 Suppl 2):S8-S15 [FREE Full text] [doi: 10.1016/j.amjmed.2008.09.008] [Medline: 18954592]
3. Egede LE. Effects of depression on work loss and disability bed days in individuals with diabetes. Diabetes Care 2004 Jul;27(7):1751-1753. [Medline: 15220260]
4. Manigault KR. The bidirectional relationship between depression and diabetes. US Pharm 2016 Nov 17;41(11):26-29 [FREE Full text]
5. Semenkovich K, Brown ME, Svrakic DM, Lustman PJ. Depression in type 2 diabetes mellitus: Prevalence, impact, and treatment. Drugs 2015 Apr;75(6):577-587. [doi: 10.1007/s40265-015-0347-4] [Medline: 25851098]
6. Holt RI, de Groot M, Golden SH. Diabetes and depression. Curr Diab Rep 2014 Jun;14(6):491 [FREE Full text] [doi: 10.1007/s11892-014-0491-3] [Medline: 24743941]
7. Renn BN, Feliciano L, Segal DL. The bidirectional relationship of depression and diabetes: A systematic review. Clin Psychol Rev 2011 Dec;31(8):1239-1246. [doi: 10.1016/j.cpr.2011.08.001] [Medline: 21963669]

8. Fisher EB, Chan JC, Nan H, Sartorius N, Oldenburg B. Co-occurrence of diabetes and depression: Conceptual considerations for an emerging global health challenge. J Affect Disord 2012 Oct;142 Suppl:S56-S66. [doi: 10.1016/S0165-0327(12)70009-5] [Medline: 23062858]
9. Hermanns N, Caputo S, Dzida G, Khunti K, Meneghini LF, Snoek F. Screening, evaluation and management of depression in people with diabetes in primary care. Prim Care Diabetes 2013 Apr;7(1):1-10 [FREE Full text] [doi: 10.1016/j.pcd.2012.11.002] [Medline: 23280258]
10. Fleer J, Tovote KA, Keers JC, Links TP, Sanderman R, Coyne JC, et al. Screening for depression and diabetes-related distress in a diabetes outpatient clinic. Diabet Med 2013 Jan;30(1):88-94. [doi: 10.1111/dme.12001] [Medline: 22924587]
11. Griffiths KM, Christensen H. Internet-based mental health programs: A powerful tool in the rural medical kit. Aust J Rural Health 2007 Apr;15(2):81-87. [doi: 10.1111/j.1440-1584.2007.00859.x] [Medline: 17441815]
12. Blixen CE, Kanuch S, Perzynski AT, Thomas C, Dawson NV, Sajatovic M. Barriers to self-management of serious mental illness and diabetes. Am J Health Behav 2016 Mar;40(2):194-204 [FREE Full text] [doi: 10.5993/AJHB.40.2.4] [Medline: 26931751]
13. Hadjiconstantinou M, Byrne J, Bodicoat DH, Robertson N, Eborall H, Khunti K, et al. Do Web-based interventions improve well-being in type 2 diabetes? A systematic review and meta-analysis. J Med Internet Res 2016 Oct 21;18(10):e270 [FREE Full text] [doi: 10.2196/jmir.5991] [Medline: 27769955]
14. van Bastelaar KM, Pouwer F, Cuijpers P, Riper H, Snoek FJ. Web-based depression treatment for type 1 and type 2 diabetic patients: A randomized, controlled trial. Diabetes Care 2011 Feb;34(2):320-325 [FREE Full text] [doi: 10.2337/dc10-1248] [Medline: 21216855]
15. Nobis S, Lehr D, Ebert DD, Berking M, Heber E, Baumeister H, et al. Efficacy and cost-effectiveness of a Web-based intervention with mobile phone support to treat depressive symptoms in adults with diabetes mellitus type 1 and type 2: Design of a randomised controlled trial. BMC Psychiatry 2013 Nov 15;13:306 [FREE Full text] [doi: 10.1186/1471-244X-13-306] [Medline: 24238346]
16. Egede LE. Diabetes, major depression, and functional disability among U.S. adults. Diabetes Care 2004 Feb;27(2):421-428. [Medline: 14747223]
17. Schmitz N, Gariépy G, Smith KJ, Malla A, Boyer R, Strychar I, et al. Longitudinal relationships between depression and functioning in people with type 2 diabetes. Ann Behav Med 2014 Apr;47(2):172-179. [doi: 10.1007/s12160-013-9534-2] [Medline: 24046149]
18. Proudfoot J, Clarke J, Birch M, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and Web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: A randomised controlled trial. BMC Psychiatry 2013 Nov 18;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
19. Clarke J, Proudfoot J, Ma H. Mobile phone and Web-based cognitive behavior therapy for depressive symptoms and mental health comorbidities in people living with diabetes: Results of a feasibility study. JMIR Ment Health 2016 May 31;3(2):e23 [FREE Full text] [doi: 10.2196/mental.5131] [Medline: 27245948]
20. de Groot M, Anderson R, Freedland KE, Clouse RE, Lustman PJ. Association of depression and diabetes complications: A meta-analysis. Psychosom Med 2001;63(4):619-630. [Medline: 11485116]
21. Gunn JM, Ayton DR, Densley K, Pallant JF, Chondros P, Herrman HE, et al. The association between chronic illness, multimorbidity and depressive symptoms in an Australian primary care cohort. Soc Psychiatry Psychiatr Epidemiol 2012 Feb;47(2):175-184. [doi: 10.1007/s00127-010-0330-z] [Medline: 21184214]
22. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: Guidance for protocols of clinical trials. BMJ 2013 Jan 08;346:e7586 [FREE Full text] [Medline: 23303884]
23. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
24. Fisher L, Gonzalez JS, Polonsky WH. The confusing tale of depression and distress in patients with diabetes: A call for greater clarity and precision. Diabet Med 2014 Jul;31(7):764-772 [FREE Full text] [doi: 10.1111/dme.12428] [Medline: 24606397]
25. Type 2 Diabetes. Turner, Australia: National Diabetes Services Scheme (NDSS), Diabetes Australia; 2016 Sep 30. URL: https://static.diabetesaustralia.com.au/s/fileassets/diabetes-australia/6eb4569c-dabf-4187-8eab-a571eb3dc3f6.pdf [accessed 2017-01-17] [WebCite Cache ID 6nZnpdASe]
26. Kroenke K, Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. Psychiatr Ann 2002 Sep;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
27. Degenhardt L, Hall W. The Association Between Psychosis and Problematic Drug Use Among Australian Adults: Findings From the National Survey of Mental Health and Well-being. Randwick, Australia: National Drug and Alcohol Research Centre, University of New South Wales; 2000. URL: https://ndarc.med.unsw.edu.au/sites/default/files/ndarc/resources/TR.093.pdf [accessed 2017-07-04] [WebCite Cache ID 6riYMjFUO]
28. myCompass. URL: https://www.mycompass.org.au/ [accessed 2017-07-02] [WebCite Cache ID 6rfMEogzd]

29. Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN) (2): A randomized trial of a self-help depression skills program with reminders. J Med Internet Res 2005 Jun 21;7(2):e16 [FREE Full text] [doi: 10.2196/jmir.7.2.e16] [Medline: 15998607]
30. Christensen H, Griffiths KM, Mackinnon AJ, Kalia K, Batterham PJ, Kenardy J, et al. Protocol for a randomised controlled trial investigating the effectiveness of an online eHealth application for the prevention of generalised anxiety disorder. BMC Psychiatry 2010 Mar 21;10:25 [FREE Full text] [doi: 10.1186/1471-244X-10-25] [Medline: 20302678]
31. 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 [FREE Full text] [Medline: 11983645]
32. Polonsky WH, Anderson BJ, Lohrer PA, Welch G, Jacobson AM, Aponte JE, et al. Assessment of diabetes-related distress. Diabetes Care 1995 Jun;18(6):754-760. [Medline: 7555499]
33. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
34. Peyrot M, Bushnell DM, Best JH, Martin ML, Cameron A, Patrick DL. Development and validation of the self-management profile for type 2 diabetes (SMP-T2D). Health Qual Life Outcomes 2012 Oct 05;10:125 [FREE Full text] [doi: 10.1186/1477-7525-10-125] [Medline: 23039868]
35. Üstün TB, Kostanjsek N, Chatterji S, Rehm J, editors. Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0). Geneva, Switzerland: World Health Organization; 2010.
36. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): Randomized controlled trial. J Med Internet Res 2009 May 11;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
37. SpringboarD. URL: https://springboard.blackdoghealth.org.au [accessed 2017-07-03] [WebCite Cache ID 6rfZH13S4]
38. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
39. Christensen H, Batterham P, Mackinnon A, Griffiths KM, Kalia HK, Kenardy J, et al. Prevention of generalized anxiety disorder using a Web intervention, iChill: Randomized controlled trial. J Med Internet Res 2014 Sep 02;16(9):e199 [FREE Full text] [doi: 10.2196/jmir.3507] [Medline: 25270886]
40. Brueton VC, Tierney J, Stenning S, Harding S, Meredith S, Nazareth I, et al. Strategies to improve retention in randomised trials. Cochrane Database Syst Rev 2013 Dec 03(12):MR000032 [FREE Full text] [doi: 10.1002/14651858.MR000032.pub2] [Medline: 24297482]
41. Robinson KA, Dinglas VD, Sukrithan V, Yalamanchilli R, Mendez-Tellez PA, Dennison-Himmelfarb C, et al. Updated systematic review identifies substantial number of retention strategies: Using more strategies retains more study participants. J Clin Epidemiol 2015 Dec;68(12):1481-1487 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.04.013] [Medline: 26186981]
42. Andrews G, Williams A. Up-scaling clinician assisted Internet cognitive behavioural therapy (iCBT) for depression: A model for dissemination into primary care. Clin Psychol Rev 2015 Nov;41:40-48. [doi: 10.1016/j.cpr.2014.05.006] [Medline: 25043445]
43. Davies K, Kingston A, Robinson L, Hughes J, Hunt J, Barker S, et al. Improving retention of very old participants in longitudinal research: Experiences from the Newcastle 85+ study. PLoS One 2014;9(10):e108370 [FREE Full text] [doi: 10.1371/journal.pone.0108370] [Medline: 25302500]
44. Stoop CH, Nefs G, Pop VJ, Pouwer F. Screening for and subsequent participation in a trial for depression and anxiety in people with type 2 diabetes treated in primary care: Who do we reach? Prim Care Diabetes 2017 Jun;11(3):273-280. [doi: 10.1016/j.pcd.2017.02.006] [Medline: 28330680]
45. Solomon D, Proudfoot J, Clarke J, Christensen H. e-CBT (myCompass), antidepressant medication, and face-to-face psychological treatment for depression in Australia: A cost-effectiveness comparison. J Med Internet Res 2015 Nov 11;17(11):e255 [FREE Full text] [doi: 10.2196/jmir.4207] [Medline: 26561555]

Abbreviations

CBT: cognitive behavioral therapy
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth
GET.ON MED: GesundheitsTraining.Online Mood Enhancer Diabetes
GP: general practitioner
HbA1c: glycosylated hemoglobin
HREC: Human Research Ethics Committee
MMRM: mixed-models repeated measures
PHQ-2: 2-item Patient Health Questionnaire
PHQ-9: 9-item Patient Health Questionnaire
RCT: randomized controlled trial
SMS: short message service

SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
UNSW: University of New South Wales
WSAS: Work and Social Adjustment Scale

Edited by G Eysenbach; submitted 18.01.17; peer-reviewed by L Mayberry, J Fleming, T Penders, P Musiat; comments to author 22.03.17; revised version received 22.05.17; accepted 24.05.17; published 03.08.17.

Please cite as:
Proudfoot J, Clarke J, Gunn J, Fletcher S, Sanatkar S, Wilhelm K, Campbell L, Zwar N, Harris M, Lapsley H, Hadzi-Pavlovic D, Christensen H
A Web-Based Public Health Intervention to Reduce Functional Impairment and Depressive Symptoms in Adults With Type 2 Diabetes (The SpringboarD Trial): Randomized Controlled Trial Protocol
JMIR Res Protoc 2017;6(8):e145
URL: http://www.researchprotocols.org/2017/8/e145/
doi:10.2196/resprot.7348
PMID:28778848

Protocol

Efficacy of Lactobacillus Reuteri DSM 17938 for the Treatment of Acute Gastroenteritis in Children: Protocol of a Randomized Controlled Trial

Henryk Szymański[1], MD; Hania Szajewska[2], MD

[1]Department of Pediatrics, St Hedwig of Silesia Hospital, Trzebnica, Poland

[2]Department of Pediatrics, The Medical University of Warsaw, Warsaw, Poland

Corresponding Author:
Hania Szajewska, MD
Department of Pediatrics
The Medical University of Warsaw
Zwirki i Wigury 63A
Warsaw, 02-091
Poland
Phone: 48 223179421
Email: hania@ipgate.pl

Abstract

Background: Acute gastroenteritis (AGE) is one of the most common diseases among children. Oral rehydration therapy is the key treatment. However, despite proven efficacy, it remains underused. This is because oral rehydration solution neither reduces the frequency of bowel movements and fluid loss nor shortens the duration of illness. Hence, there is interest in adjunctive treatments. According to the 2014 guidelines developed by the European Society for Paediatric Gastroenterology, Hepatology and Nutrition, the use of the following probiotics may be considered in the management of children with AGE in addition to rehydration therapy: *Lactobacillus rhamnosus* GG (low quality of evidence; strong recommendation) and *Saccharomyces boulardii* (low quality of evidence; strong recommendation). Less compelling evidence is available for *Lactobacillus reuteri* DSM 17938 (very low quality of evidence; weak recommendation).

Objective: Considering that evidence on *L reuteri* remains limited, the goal of the study is to assess the effectiveness of *L reuteri* DSM 17938 in the treatment of AGE in children. Children vaccinated and not vaccinated against rotavirus will be evaluated separately.

Methods: This will be a double-blind, placebo-controlled, randomized trial. Children between 1 and 60 months of age with AGE, defined as a change in stool consistency to loose or liquid form (according to the Bristol Stool Form scale or Amsterdam Stool Form scale) and/or an increase in the frequency of evacuations (typically ≥3 in 24 h) lasting for no longer than 5 days, will be recruited. A total of 72 children will receive either *L reuteri* DSM 17938 at a dose of 2×10^8colony-forming units twice daily or matching placebo for 5 consecutive days. A similar sample size for rotavirus vaccinated and nonvaccinated children is planned. The primary outcome measure is the duration of diarrhea. Two separate studies and reports for rotavirus vaccinated and nonvaccinated children are planned.

Results: The recruitment started in January 2017 and is planned to be finalized in June 2018 for rotavirus nonvaccinated children. The recruitment of rotavirus-vaccinated children may be slower due to a relatively low coverage rate in Poland. Data analysis and submission to a peer-reviewed journal is expected within 3 months after completion of the study.

Conclusion: This study will add to current knowledge on the efficacy of *L reuteri* DSM 17938 for the management of AGE.

Trial registration: ClinicalTrials.gov NCT02989350; https://clinicaltrials.gov/ct2/show/NCT02989350 (Archived by WebCite at http://www.webcitation.org/6slOFkyTH)

(JMIR Res Protoc 2017;6(8):e164) doi:10.2196/resprot.7924

KEYWORDS

probiotics; diarrhea; infants; RCT

Introduction

Acute gastroenteritis (AGE) is one of the most common diseases among children. Generally, this is a self-limited illness lasting 5 to 7 days, and thus, the main aim of treatment is to prevent dehydration, metabolic acidosis, and electrolyte disturbances. In the vast majority of cases of AGE with mild or moderate dehydration, this can be achieved with oral rehydration solutions. Despite the proven efficacy of oral rehydration, it remains underused [1]. The main reason for this is that an oral rehydration solution neither reduces the frequency of bowel movements and fluid loss nor shortens the duration of illness, which decreases its acceptance and prompts interest in adjunctive treatments.

In 2014, the European Society for Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) provided recommendations for the use of probiotics for the treatment of AGE in previously healthy infants and children based on a systematic review. The use of the following probiotics may be considered in the management of children with AGE in addition to rehydration therapy: *Lactobacillus rhamnosus* GG (low quality of evidence; strong recommendation) and *Saccharomyces boulardii* (low quality of evidence; strong recommendation). Less compelling evidence is available for *Lactobacillus reuteri* DSM 17938 (very low quality of evidence; weak recommendation) [2,3].

L reuteri is named after Gerhard Reuter, a German microbiologist [4]. Among a number of *L reuteri* strains, *L reuteri* DSM 17938 is probably the best studied. *L reuteri* DSM 17938 is the daughter strain of *L reuteri* ATCC 55730. The latter was originally isolated from the breast milk of a Peruvian mother, and it may be present in normal humans on the mucosa of the gastric corpus and antrum, duodenum, and ileum [4,5]. As *L reuteri* ATCC 55730 was found to carry potentially transferable resistance traits for tetracycline and lincomycin, it was replaced by *L reuteri* DSM 17938, a strain without unwanted plasmid-borne resistance [6]. Like other *L reuteri* strains, *L reuteri* DSM 17938 has several different types of mechanisms of action. One of the best-documented mechanisms is its antimicrobial activity. *L reuteri* strains produce reuterin, a broad-spectrum antibacterial substance [7,8], which is capable of inhibiting the growth of a wide spectrum of microorganisms such as Gram-positive or negative bacteria, yeast, fungi, or parasites [9]. *L reuteri* strains may also regulate immune responses [10-12].

The clinical efficacy of *L reuteri* DSM 17938 has been studied in a number of clinical trials, including studies aimed at the assessment of *L reuteri* DSM 17938 in the management of AGE.

A 2012 study by Francavilla et al [13] was carried out in 74 children with acute diarrhea randomized to receive *L reuteri* DSM 17938 (at a dose of 4×10^8colony-forming units [CFU]) or placebo for 7 days. Compared with the placebo group, in the *L reuteri* group there was a significant reduction in the duration of diarrhea (3.3 [SD 2.1] vs 2.1 [SD 1.7] days, respectively; *P*<.03), the risk of watery diarrhea on day 2 (81% vs 55%, respectively, *P*<.02) and on day 3, 73% vs 46%, respectively, *P*<.03), and the risk of relapse of diarrhea (42% vs 15%; respectively, *P*<.03). There was not a significant difference in hospital stay between the groups.

A 2014 multicenter, randomized, single-blind clinical trial by Dinleyici et al [14] was performed in 127 hospitalized children with AGE lasting 12 to 72 hours. Conventional therapy plus the administration of *L reuteri* DSM 17938 at a dose 1×10^8CFU for 5 days compared with conventional therapy alone (control group) reduced the duration of diarrhea (70.7 [SD 26.1] vs 103.8 [SD 28.47] hours, respectively; *P*<.001), increased the number of diarrhea-free children after 24 and 48 hours (50% vs 5%, respectively, *P*<.001) and 72 hours (69% vs 11%, respectively, *P*<.001), and reduced the hospital stay (4.3 [SD 1.3] days vs 5.46 [SD 1.77] days, respectively, *P*<.001).

A 2015 trial by the same authors performed in 64 outpatient children with acute infectious diarrhea found that compared with oral rehydration therapy alone (control group), the additional administration of *L reuteri* at a dose 1×10^8CFU for 5 days significantly reduced the duration of diarrhea (74.3 [SD 15.3] hours vs 60.4 [SD 24.5] hours, respectively, *P*<.05). The percentage of children with diarrhea after 48 hours was lower in the *L reuteri* group than the control group (13/29 vs 27/31, respectively; relative risk [RR] 0.51, 95% CI 0.34 to 0.79, *P*<.01). From the 72nd hour of the intervention onward, there was no difference between the 2 groups in the percentage of children with diarrhea [15].

A 2015 meta-analysis of these 3 randomized controlled trials (n=256) found that compared with placebo or no treatment, *L reuteri* DSM 17938 administration significantly reduced the duration of diarrhea (mean difference [MD] –24.82 hours, 95% CI –38.8 to –10.8) and increased the chance of cure on day 1 (RR 11.3, 95% CI 2.2 to 59) and day 2 (RR 4.54, 95% CI 2.0 to 10.2) [16]. However, heterogeneity and wide confidence intervals call for caution in interpreting the results. To the best of our knowledge, there are no data on the use of probiotics use in rotavirus-vaccinated children with AGE.

More studies are needed to establish the efficacy of *L reuteri* DSM 17938 in the management of AGE in children. We aim to conduct a well-designed study with sufficient power, an adequate follow-up period, and relevant clinical end points. The study will be led by a team experienced in performing clinical trials in children with AGE [17,18] and follow the same protocol as one of our recent trials on AGE [19].

The objective of this study is to assess the effectiveness of *L reuteri* DSM 17938 in the treatment of AGE in children. Children vaccinated and not vaccinated against rotavirus will be evaluated separately.

Methods

Study Design

This study is designed as a randomized, double-blinded, placebo-controlled trial with 1:1 allocation, and it is described in more detail in the subsequent sections. The trial was registered at ClinicalTrials.gov [NCT02989350], and any important changes in the protocol will be implemented there.

Ethics and Dissemination

The Ethical Committee of the Lower Silesia Medical Chamber issued approval for the study before recruitment commenced. Verbal and written information regarding informed consent will be presented to the caregivers. Any modifications to the protocol that may affect the conduct of the study will be presented to the Ethical Committee. The full protocol will be available freely due to open-access publication. The findings of this randomized controlled trial will be submitted to a peer-reviewed journal. Abstracts will be submitted to relevant national and international conferences. The need for separate reporting of data on rotavirus-vaccinated and nonvaccinated populations is foreseen. The Ethics Committee did not require auditing for this study.

Setting and Participants

The recruitment will take place in the emergency room and the Department of Paediatrics of the St Hedwig of Silesia Hospital in Trzebnica, Poland. However, inclusion of outpatients and involvement of other recruiting sites are under consideration provided the study staff is experienced, adequately trained, and competent in conducting clinical trials. The start of the recruitment is planned in January 2017, and it should be completed within the following 2 years. Participants will be randomized after their first visit to the emergency room or after admission to the hospital. Caregivers will receive oral and written information on the study. Written informed consent will by obtained by physicians involved in the study. See Textbox 1 for selection criteria.

Interventions

The intervention under investigation is *L reuteri* DSM 17938 . The placebo drops consist of a mixture of pharmaceutical grade medium chain triglycerides and sunflower oil together with pharmaceutical grade silicon dioxide to give the product the correct rheological properties. The formulation is identical to the active product but without *L reuteri* DSM 17938. In our trial, we choose to use a placebo for a comparator, as it is widely regarded as the gold standard for testing the efficacy of new treatments [20,21].

The study products (*L reuteri* DSM 17938 and placebo) will be manufactured and supplied by BioGaia as bottle with drops, free of charge. The manufacturer will have no role in the conception, protocol development, design, or conduct of the study or in the analysis or interpretation of the data.

The dose of *L reuteri* DSM 17938 will be 2×10^8CFU. Both *L reuteri* DSM 17938 and placebo will be taken orally, 5 drops twice daily for a consecutive 5 days. Caregivers will be instructed to administer the study products at the same time of the day. If needed, discontinuation or modification of the treatment may be considered at the discretion of the physician.

Table 1 presents a timetable of activities planned during the study. For initial rehydration, all children will be treated according to the 2014 ESPGHAN recommendations (fast oral rehydration over 3 to 4 hours by mouth or via nasogastric tube with the recommended hypotonic solution) [3]. After all signs of dehydration have disappeared, oral rehydration solution will be given for ongoing losses until the diarrhea stops. Rapid reintroduction of the previous diet after successful rehydration will be recommended.

Textbox 1. Selection criteria.

Inclusion criteria:

- Acute gastroenteritis defined as a change in stool consistency to loose or liquid form (according to the Bristol Stool Form scale or in the case of infants, the Amsterdam Stool Form scale) and/or an increase in the frequency of evacuations (typically ≥3 in 24 hours) lasting for no longer than 5 days
- Aged older than 1 month and younger than 60 months
- Caregiver must provide written informed consent

Exclusion criteria:

- Use of antibiotics within 2 weeks prior to enrollment
- Use of gelatine tannate, diosmectite, probiotics, racecadotril, or zinc (including zinc-containing oral rehydration solution) within 1 week prior to enrollment (a single dose is allowed)
- Breastfeeding (>50%).
- Chronic diarrheal gastrointestinal disease (eg, inflammatory bowel disease, cystic fibrosis, celiac disease, food allergy)
- Immunodeficiency
- Malnutrition (weight/height/length under 3rd percentile) (World Health Organization Child Growth Standards will be used) [20]

Table 1. Timetable of activities planned during the study.

Activities	Days							
	1	2	3	4	5	6	7	8
Inclusion/exclusion criteria	x							
Informed consent	x							
Enrollment	x							
Randomization	x							
Stool sample taken	x							
Intervention	x	x	x	x	x			
Follow-up						x	x	x
Return of diary								x
Concomitant medication	x	x	x	x	x	x	x	x
Adverse events	x	x	x	x	x	x	x	x

After rechecking the inclusion and exclusion criteria, participants will be assigned into 1 of 2 groups (experimental or control). Caregivers will receive a diary of symptoms to record the number of stools and stool consistency during the intervention (including recording of the timing of stools); Bristol Stool Form (BSF) and Amsterdam Stool Form (ASF) scales will be provided. Additionally, caregivers will be asked to write down any adverse events and concomitant medication during the intervention period. In line with current ESPGHAN guidelines [3], children presenting with AGE do not require routine etiological investigation; however, before the first dose of study products, a stool sample will be taken to determine the rotavirus etiology of the diarrhea with a standard immunoassay method. More specific microbiological investigations with standard stool cultures will be performed only if needed [3]. At any time, caregivers will have the right to withdraw the participating child from the study; they will be not obliged to give reasons for this decision, and there will be no effect on subsequent physician and/or institutional medical care.

Concomitant Medications

The concomitant administration of any other medication, including antipyretics and antiemetics, will be at the discretion of the physician to provide adequate care. However, it is recommended that no unnecessary concomitant medication be used. In particular, the use of antibiotics, diosmectite, probiotics, or racecadotril (all included in the exclusion criteria) should be avoided.

Follow-Up

All study participants will be followed up for the duration of the intervention (5 days) and then for an additional 3 days or until the cessation of diarrhea.

Outcomes

The primary outcome will be the duration of diarrhea, defined as the time until the normalization of stool consistency according to the BSF or ASF scale (in BSF scale, numbers 1, 2, 3, 4, and 5; in ASF scale, letters B or C) or the time until the normalization of the number of stools (compared with the period before the onset of diarrhea) and the presence of normal stools for 48 hours.

Secondary outcomes will include:

- Need for intravenous rehydration
- Duration of intravenous rehydration
- Need for hospitalization of outpatients
- Number of watery stools per day
- Vomiting
- Recurrence of diarrhea (48 hours after intervention)
- Severity of diarrhea according to Vesikari scale [22]
- Use of concomitant medications
- Adverse events

Randomization

A computer-generated randomization list will be prepared by a person with no clinical involvement in the trial using a computer program (StatsDirect, StatsDirect Ltd.) with an allocation ratio of 1:1 and a block of 6. The randomization lists will be stratified according to rotavirus vaccination status (nonvaccinated and vaccinated; the latter will be defined as receiving at least 1 dose of rotavirus vaccine).

Allocation Concealment

The allocation sequence will be concealed from the researchers enrolling and assessing participants in sequentially numbered, white, opaque, sealed and stapled envelopes which will be opened only after getting informed consent and registering the basic demographic data to case report form (CRF). Consecutive randomization numbers will be given to participants at enrollment. The study product will be packaged as product A or product B and will be given according to the randomization list.

Blinding

The study products (active and placebo) will be packaged in identical bottles. Contents will look and taste the same. Researchers, caregivers, outcome assessors, and a person responsible for the statistical analysis will be blinded to the intervention until the completion of the study. The information on intervention assignments will be stored in a sealed envelope

in a safe place in the administrative part of the department. The personal information about potential and enrolled participants will be stored in a locker within the study site, accessible to the involved researchers only.

Compliance

The caregivers will be asked to bring the remaining study product and diary to the study site at the end of the intervention period. Compliance with the study protocol will be checked by measuring the volume left unused. Based on previously published trials, it seems appropriate to consider those participants receiving <75% of the recommended doses as noncompliant.

Power Calculation

The primary outcome of the study is the duration of diarrhea. Based on available data in the literature, the average duration of gastroenteritis in children is 5 to 7 days [3]. We assume that a clinically significant difference in the effectiveness of *L reuteri* DSM 17938 versus placebo will shorten the duration of symptoms by 24 hours (±24 hours). To detect such a difference in the duration of diarrhea between the study groups with a power of 90% and alpha of 0.01, a sample of 60 children is needed. Assuming approximately 20% loss to follow-up, we aim to recruit a total of 72 children for this study. A similar sample size for rotavirus vaccinated and nonvaccinated children is planned. Sample size calculations were performed with StatsDirect version 2.3.8 (StatsDirect Ltd).

At the Department of Paediatrics of St Hedwig of Silesia Hospital, there are 300 admissions of children with diarrhea per year and the same number of such patients who present to the emergency room. Assuming that 20% of these children will be eligible for the study, we will achieve adequate participant enrollment to reach the target sample size during 1 year of recruiting for rotavirus nonvaccinated children. Recruitment of rotavirus-vaccinated children may be slower (up to 2 years). In Poland, rotavirus vaccination has been recommended since 2006; however, vaccine reimbursement is not available. Hence, currently, the rotavirus vaccination coverage rate is low (12.7% [23] to 30% [24]), likely due to the high vaccine cost.

Data Collection and Management

All study participants will be assigned a study identification number. CRFs will be completed on paper forms. Data will then be entered and stored in a password-protected electronic database. The original paper copies of CRFs and all study data will be stored in a locker within the study site, accessible to the involved researchers only.

Statistical Analysis

All analysis will be conducted on an intention-to-treat basis, including all patients in the groups to which they are randomized for whom outcomes are available (including drop-outs and withdrawals). Descriptive statistics will be used to summarize baseline characteristics. The Student *t* test will be used to compare mean values of continuous variables for approximating a normal distribution. For nonnormally distributed variables, the Mann-Whitney U test will be used. The chi-square or Fisher exact test will be used, when appropriate, to compare percentages. For continuous outcomes, differences in means or differences in medians (depending on the distribution of the data), and for dichotomous outcomes, the RR and number needed to treat, all with a 95% CI, will be calculated. The difference between study groups will be considered significant when the *P* value is <.05, when the 95% CI for RR does not include 1.0, or when the 95% CI for MD does not include 0. All statistical tests will be 2-tailed and performed at the 5% level of significance. Two independent reports (rotavirus-vaccinated and nonvaccinated children) are planned.

Monitoring

The study will be carried out in accordance with the approved protocol. *L reuteri* DSM 17938 is being safely used worldwide for a number of indications, and the US Food and Drug Administration applied to it the Generally Recognized as Safe status [25]. Still, an independent Data and Safety Monitoring Board (DSMB) will be set up prior to the start of the study. The DSMB will review data after recruitment of 25%, 50%, and 75% of subjects to review the study progress and all adverse events.

Harms

Although the occurrence of adverse events as a result of participation in the current trial is not expected, data on adverse events will be collected. All serious adverse events will be immediately reported to the study coordinators who will be responsible for notifying the Ethics Committee, all participating investigators, and the manufacturer of the study products.

Results

The recruitment started in January 2017 and is planned to be finalized in June 2018 for rotavirus nonvaccinated children. The recruitment of rotavirus-vaccinated children may be slower due to a relatively low coverage rate in Poland. Data analysis and submission to a peer-reviewed journal is expected within 3 months after completion of the study.

Discussion

Summary

This study will add to the current knowledge on the efficacy of *L reuteri* DSM 17938 for the management of AGE. Children vaccinated and not vaccinated against rotavirus will be evaluated separately.

Strengths and Limitations of This Study

A strength of the study is its study design (randomized controlled trial), which is the most robust methodology to assess the effectiveness of therapeutic interventions. The findings of this randomized controlled trial, whether positive or negative, will contribute to the formulation of further recommendations on the use of *L reuteri* DSM 17938 for the treatment of AGE in children. A limitation is that stool volume, one of the objective ways of assessing the efficacy of antidiarrheal drugs, will not be assessed. In addition, the recruitment of rotavirus-vaccinated children may be slow due to a relatively low coverage rate in Poland, likely due to the high vaccine cost.

Authors' Contributions

H Szajewska conceptualized the study. H Szymański developed the first draft of the manuscript. Both authors contributed to the development of the study protocol and approved the final draft of the manuscript.

Conflicts of Interest

H Szymański declares no conflicts of interest. H Szajewska has served as a speaker for BioGaia, the manufacturer of *L reuteri* DSM 17938.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (v1.6.1).

[PDF File (Adobe PDF File), 601KB - resprot_v6i8e164_app1.pdf]

References

1. Guarino A, Albano F, Guandalini S. Oral rehydration: toward a real solution. J Pediatr Gastroenterol Nutr 2001 Oct;33 Suppl 2:S2-S12. [Medline: 11698779]
2. Szajewska H, Guarino A, Hojsak I, Indrio F, Kolacek S, Shamir R, et al. Use of probiotics for management of acute gastroenteritis: a position paper by the ESPGHAN Working Group for Probiotics and Prebiotics. J Pediatr Gastroenterol Nutr 2014 Apr;58(4):531-539. [doi: 10.1097/MPG.0000000000000320] [Medline: 24614141]
3. Guarino A, Ashkenazi S, Gendrel D, Lo Vecchio A, Shamir R, Szajewska H. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition/European Society for Pediatric Infectious Diseases evidence-based guidelines for the management of acute gastroenteritis in children in Europe: update 2014. J Pediatr Gastroenterol Nutr 2014 Jul;59(1):132-152. [doi: 10.1097/MPG.0000000000000375] [Medline: 24739189]
4. Reuter G. The Lactobacillus and Bifidobacterium microflora of the human intestine: composition and succession. Curr Issues Intest Microbiol 2001 Sep;2(2):43-53. [Medline: 11721280]
5. Valeur N, Engel P, Carbajal N, Connolly E, Ladefoged K. Colonization and immunomodulation by Lactobacillus reuteri ATCC 55730 in the human gastrointestinal tract. Appl Environ Microbiol 2004 Feb;70(2):1176-1181 [FREE Full text] [Medline: 14766603]
6. Rosander A, Connolly E, Roos S. Removal of antibiotic resistance gene-carrying plasmids from Lactobacillus reuteri ATCC 55730 and characterization of the resulting daughter strain, L. reuteri DSM 17938. Appl Environ Microbiol 2008 Oct;74(19):6032-6040 [FREE Full text] [doi: 10.1128/AEM.00991-08] [Medline: 18689509]
7. Axelsson LT, Chung TC, Dobrogosz WJ, Lindgren SE. Production of a broad spectrum antimicrobial substance by Lactobacillus reuteri. Microb Ecol Health Dis 1989 Jan;2(2):131-136. [doi: 10.3109/08910608909140210]
8. Talarico TL, Casas IA, Chung TC, Dobrogosz WJ. Production and isolation of reuterin, a growth inhibitor produced by Lactobacillus reuteri. Antimicrob Agents Chemother 1988 Dec;32(12):1854-1858 [FREE Full text] [Medline: 3245697]
9. Chung TC, Axelsson L, Lindgren SE, Dobrogosz WJ. In vitro studies on reuterin synthesis by Lactobacillus reuteri. Microb Ecol Health Dis 1989 Jan;2(2):137-144. [doi: 10.3109/08910608909140211]
10. Lin YP, Thibodeaux CH, Peña JA, Ferry GD, Versalovic J. Probiotic Lactobacillus reuteri suppress proinflammatory cytokines via c-Jun. Inflamm Bowel Dis 2008;14(8):1068-1083. [doi: 10.1002/ibd.20448]
11. Liu Y, Fatheree NY, Mangalat N, Rhoads JM. Human-derived probiotic Lactobacillus reuteri strains differentially reduce intestinal inflammation. Am J Physiol-Gastr L 2010 Aug 26;299(5):G1087-G1096. [doi: 10.1152/ajpgi.00124.2010]
12. Liu Y, Fatheree NY, Mangalat N, Rhoads JM. Lactobacillus reuteri strains reduce incidence and severity of experimental necrotizing enterocolitis via modulation of TLR4 and NF- B signaling in the intestine. Am J Physiol-Gastr L 2011 Dec 29;302(6):G608-G617. [doi: 10.1152/ajpgi.00266.2011]
13. Francavilla R, Lionetti E, Castellaneta S, Ciruzzi F, Indrio F, Masciale A, et al. Randomised clinical trial: Lactobacillus reuteri DSM 17938 vs. placebo in children with acute diarrhoea - a double-blind study. Aliment Pharmacol Ther 2012 Jun 11;36(4):363-369. [doi: 10.1111/j.1365-2036.2012.05180.x]
14. Dinleyici EC, PROBAGE Study Group, Vandenplas Y. Lactobacillus reuteri DSM 17938 effectively reduces the duration of acute diarrhoea in hospitalised children. Acta Paediatr 2014 Jul;103(7):e300-e305. [doi: 10.1111/apa.12617] [Medline: 24579935]
15. Dinleyici EC, Dalgic N, Guven S, Metin O, Yasa O, Kurugol Z, et al. Lactobacillus reuteri DSM 17938 shortens acute infectious diarrhea in a pediatric outpatient setting. J Pediatr (Rio J) 2015 Jul;91(4):392-396. [doi: 10.1016/j.jped.2014.10.009]
16. Urbańska M, Gieruszczak-Białek D, Szajewska H. Systematic review with meta-analysis: Lactobacillus reuteri DSM 17938 for diarrhoeal diseases in children. Aliment Pharmacol Ther 2016 May;43(10):1025-1034. [doi: 10.1111/apt.13590] [Medline: 26991503]

17. Szymański H, Pejcz J, Jawień M, Chmielarczyk A, Strus M, Heczko PB. Treatment of acute infectious diarrhoea in infants and children with a mixture of three Lactobacillus rhamnosus strains--a randomized, double-blind, placebo-controlled trial. Aliment Pharmacol Ther 2006 Jan 15;23(2):247-253 [FREE Full text] [doi: 10.1111/j.1365-2036.2006.02740.x] [Medline: 16393304]
18. Guandalini S, Pensabene L, Zikri MA, Dias JA, Casali LG, Hoekstra H, et al. Lactobacillus GG administered in oral rehydration solution to children with acute diarrhea: a multicenter European trial. J Pediatr Gastroenterol Nutr 2000 Jan;30(1):54-60. [Medline: 10630440]
19. Michałek D, Kołodziej M, Konarska Z, Szajewska H. Efficacy and safety of gelatine tannate for the treatment of acute gastroenteritis in children: protocol of a randomised controlled trial. BMJ Open 2016 Feb 19;6(2):e010530 [FREE Full text] [doi: 10.1136/bmjopen-2015-010530] [Medline: 26895988]
20. Castro M. Placebo versus best-available-therapy control group in clinical trials for pharmacologic therapies: which is better? Proc Am Thorac Soc 2007 Oct 01;4(7):570-573 [FREE Full text] [doi: 10.1513/pats.200706-073JK] [Medline: 17878471]
21. Benson H, Friedman R. Harnessing the power of the placebo effect and renaming it remembered wellness. Annu Rev Med 1996;47:193-199. [doi: 10.1146/annurev.med.47.1.193] [Medline: 8712773]
22. Schnadower D, Tarr PI, Gorelick MH, O'Connell K, Roskind CG, Powell EC, et al. Validation of the modified Vesikari score in children with gastroenteritis in 5 US emergency departments. J Pediatr Gastroenterol Nutr 2013 Oct;57(4):514-519 [FREE Full text] [doi: 10.1097/MPG.0b013e31829ae5a3] [Medline: 23676445]
23. Ganczak M, Dmytrzyk-Daniłów G, Karakiewicz B, Korzeń M, Szych Z. Determinants influencing self-paid vaccination coverage, in 0-5 years old Polish children. Vaccine 2013 Nov 19;31(48):5687-5692. [doi: 10.1016/j.vaccine.2013.09.056] [Medline: 24120549]
24. Pokorna-Kalwak D, Gwiazda E, Mastalerz-Migas A. Rotavirus vaccine coverage in one primary care setting. Family Med Prim Care Rev 2009;11:60.
25. Heimbach J, Center for Food Safety and Applied Nutrition, Food and Drug Administration. GRAS Notice 000254: Generally Recognized as Safe (GRAS) Determination of Lactobacillus reuteri Strain DSM 17938. 2008. URL: https://www.fda.gov/downloads/food/ingredientspackaginglabeling/gras/noticeinventory/ucm287952.pdf [accessed 2017-08-16] [WebCite Cache ID 6slYhukav]

Abbreviations

AGE: acute gastroenteritis
ASF: Amsterdam Stool Form
BSF: Bristol Stool Form
CFU: colony forming units
CRF: case report form
DSMB: Data Safety and Monitoring Board
ESPHGAN: European Society of Pediatric Gastroenterology, Hepatology, and Nutrition
MD: mean difference
RR: relative risk

Edited by G Eysenbach; submitted 26.04.17; peer-reviewed by B Singh, M Rezai; comments to author 08.06.17; revised version received 19.06.17; accepted 10.07.17; published 23.08.17.

Please cite as:
Szymański H, Szajewska H
Efficacy of Lactobacillus Reuteri DSM 17938 for the Treatment of Acute Gastroenteritis in Children: Protocol of a Randomized Controlled Trial
JMIR Res Protoc 2017;6(8):e164
URL: http://www.researchprotocols.org/2017/8/e164/
doi:10.2196/resprot.7924
PMID:28835355

Protocol

Smartphone-Enabled Health Coaching Intervention (iMOVE) to Promote Long-Term Maintenance of Physical Activity in Breast Cancer Survivors: Protocol for a Feasibility Pilot Randomized Controlled Trial

Paul Ritvo[1,2], PhD; Maya Obadia[3], PhD; Daniel Santa Mina[3,4], PhD; Shabbir Alibhai[5,6], MD; Catherine Sabiston[4], PhD; Paul Oh[7,8], MD; Kristin Campbell[9], PhD; David McCready[10,11], MD; Leslie Auger[12], PhD; Jennifer Michelle Jones[3,13], PhD

[1]School of Kinesiology and Health Science, York University, Toronto, ON, Canada
[2]Department of Psychology, University of Toronto, Toronto, ON, Canada
[3]Cancer Rehabilitation and Survivorship Program, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada
[4]Kinesiology and Physical Education, University of Toronto, Toronto, ON, Canada
[5]General Internal Medicine, University Health Network, Toronto, ON, Canada
[6]Department of Medicine, University of Toronto, Toronto, ON, Canada
[7]Cardiovascular Prevention and Rehabilitation Program, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
[8]Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada
[9]Department of Physical Therapy, University of British Columbia, Vancouver, BC, Canada
[10]Surgical Oncology, University Health Network, Toronto, ON, Canada
[11]Division of Surgery, University of Toronto, Toronto, ON, Canada
[12]Kinesiology Program, University of Guelph—Humber, Toronto, ON, Canada
[13]Department of Psychiatry, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Jennifer Michelle Jones, PhD
Cancer Rehabilitation and Survivorship Program
Princess Margaret Cancer Centre
University Health Network
PMB-B-148
200 Elizabeth St
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 581 8603
Email: jennifer.jones@uhn.ca

Abstract

Background: Although physical activity has been shown to contribute to long-term disease control and health in breast cancer survivors, a majority of breast cancer survivors do not meet physical activity guidelines. Past research has focused on promoting physical activity components for short-term breast cancer survivor benefits, but insufficient attention has been devoted to long-term outcomes and sustained exercise adherence. We are assessing a health coach intervention (iMOVE) that uses mobile technology to increase and sustain physical activity maintenance in initially inactive breast cancer survivors.

Objective: This pilot randomized controlled trial (RCT) is an initial step in evaluating the iMOVE intervention and will inform development of a full-scale pragmatic RCT.

Methods: We will enroll 107 physically inactive breast cancer survivors and randomly assign them to intervention or control groups at the University Health Network, a tertiary cancer care center in Toronto, Canada. Participants will be women (age 18 to 74 years) stratified by age (55 years and older/younger than 55 years) and adjuvant hormone therapy (AHT) exposure (AHT vs no AHT) following breast cancer treatment with no metastases or recurrence who report less than 60 minutes of preplanned physical activity per week. Both intervention and control groups receive the 12-week physical activity program with weekly

group sessions and an individualized, progressive, home-based exercise program. The intervention group will additionally receive (1) 10 telephone-based health coaching sessions, (2) smartphone with data plan, if needed, (3) supportive health tracking software (Connected Wellness, NexJ Health Inc), and (4) a wearable step-counting device linked to a smartphone program.

Results: We will be assessing recruitment rates; acceptability reflected in selective, semistructured interviews; and enrollment, retention, and adherence quantitative intervention markers as pilot outcome measures. The primary clinical outcome will be directly measured peak oxygen consumption. Secondary clinical outcomes include health-related quality of life and anthropometric measures. All outcome measures are administered at baseline, after exercise program (month 3), and 6 months after program (month 9).

Conclusions: This pilot RCT will inform full-scale RCT planning. We will assess pilot procedures and interventions and collect preliminary effect estimates.

Trial Registration: ClinicalTrials.gov NCT02620735; https://clinicaltrials.gov/ct2/show/NCT02620735 (Archived by WebCite at https://clinicaltrials.gov/ct2/show/NCT02620735)

(JMIR Res Protoc 2017;6(8):e165) doi:10.2196/resprot.6615

KEYWORDS

breast neoplasm; exercise; health coaching; RCT; telehealth

Introduction

Background

Breast cancer, the most frequently diagnosed female cancer, accounts for approximately 25% of new Canadian cancer diagnoses [1]. With improvements in early detection and treatment, breast cancer mortality has decreased significantly in the last 30 years despite increasing incidence [2,3]. Most breast cancer cases diagnosed at localized stages are associated with a mean 5-year survival rate of 96% [2,4].

Despite these improvements, breast cancer treatments can result in the long-term effects of chronic pain, fatigue, neuropathy, functional limitations, sleep disturbance, sexual dysfunction, infertility, cognitive impairment, cardiorespiratory dysfunctions, and generally reduced well-being [5-14]. Breast cancer survivors also confront elevated risks for local or distal recurrence, metastases, second primary cancers, type 2 diabetes, and cardiovascular disorders [3,15-23].

Physical activity can improve cancer outcomes and quality of life while reducing adverse effects and risks. Moderate-to-vigorous physical activity (MVPA) during or after breast cancer treatment is specifically associated with reductions in cancer-specific and all-cause mortality [23,24]. As reported in a recent systematic review of 17 breast cancer–specific observational studies, breast cancer–specific mortality reductions of 13% to 51% were observed when the highest-to-lowest physical activity categories were compared [25-32].

The American College of Sports Medicine (ACSM), the American Cancer Society, and Worldwide Cancer Research (among other national and international agencies) recommend 150 minutes per week of moderate intensity physical activity for cancer survivors [33-37]. While most breast cancer survivors believe in exercise benefits [37], physical activity levels generally reduce after a breast cancer diagnosis with the large majority of breast cancer survivors (more than 80%) not meeting recommended physical activity levels [29,37-50].

Nonetheless, the impact of a cancer diagnosis often stimulates patients to reconsider lifestyle modification [43,46,50-52], providing clinicians the opportunity to introduce physical activity promotion [53]. Substantial evidence supports the efficacy of several intervention approaches in short-term physical activity change [54-57], with findings from systematic reviews and meta-analyses of exercise studies involving cancer survivors indicating that MVPA (1) is safe and well-tolerated [35,58], (2) can significantly improve quality of life [58-61] and, (3) can improve aerobic and musculoskeletal fitness, body composition, social functioning, and mental health and reduce fatigue [36,55,56,58,61-69].

MVPA benefits following breast cancer diagnosis are only maintained for as long as exercise behaviors continue [70-72]. Therefore, the longitudinal assessment of MVPA maintenance following interventions is critical [57,72,73]. In noncancer populations, physical activity intervention effects are infrequently maintained [74-84].

Despite varying reports of barriers, the long-term maintenance of and adherence to MVPA protocols in cancer survivors has not been adequately studied [74-84]. For example, in a recent systematic review of physical and/or dietary interventions in breast cancer populations [85], only 10 of 63 trials assessed the postintervention maintenance of behavioral outcomes [85]. Of these, 4 of 10 achieved successful maintenance (defined as longer than 3 months) [85]. In one recent study of 488 long-term (more than 5 years) cancer survivors with mixed tumor types, participants were randomized to a wait-list control or to a combined diet and physical activity intervention consisting of mailed print material and 15 telephone counseling sessions over 12 months. In the intervention group, weekly physical activity levels increased significantly from baseline (37.5 minutes) to 1-year assessment (postintervention, 101.0 minutes), and these elevated levels were maintained at the 2-year follow-up assessment [86]. While these findings are encouraging, reliance on self-reported physical activity measures and low recruitment rates warrant additional studies with improved designs. Altogether, longitudinal assessments of physical activity maintenance using objective measures in breast cancer survivors are rare, and more are needed to inform physical activity

interventions aimed at achieving stabilized, long-term health outcomes [85,87].

Not surprisingly, the health behavior change methods guiding counseling in long-term MVPA maintenance have been inadequately tested in breast cancer survivors. Patient-centered interventions affecting multiple factors (eg, intrinsic motivations, perceived costs and benefits, barriers, ability to change) [88,89] have been derived from evidence-based models (eg, transtheoretical model, social cognitive theory, cognitive behavior theory, and theory of planned behavior) and, in the past, their related efficacies in changing multiple lifestyle behaviors (eg, smoking, diet, chronic sedentariness) have been demonstrated [87,90-93]. While evidential support does not favor one behavior change model, successful physical activity promotion programs have included self-directed physical activity guided by a counselor, follow-up behavioral prompts [56,94-99], and more than 4 sessions of related counseling.

Implementation of theory-based behavior change models for breast cancer survivors aimed at longitudinally maintained MVPA must account for treatment-related sequelae, including adaptations that distract from or discourage health behaviors (eg, avoiding physical pain and discomfort). Accordingly, healthy physical activity promotion requires a cognitive component that emphasizes protective MVPA effects (eg, prevention of breast cancer recurrence) and a cognitive-behavioral component that assists incremental physical activity increases. Exercise prescriptions identify protective goals while carefully incremented training programs assist breast cancer survivors with immediate experiences of improved fitness, well-being, and achievement.

Counseling strategies for improving MVPA can benefit from Internet linkage, smartphone use, and wearable technologies. As of 2012, high proportions of Canadian households access broadband Internet, with mobile services adopted by nearly 80% [100]. Concomittantly, smartphone use has increased from 33% to 56% in all adult Canadians [101]. In the United States, by 2011 78% of adults used the Internet [102] and at least 64% use smartphones [103]. With increasing use, mobile technology has a rapidly increasing health care role via clinical decision making and data collection supporting chronic disease self-management [102].

In support of health behavior change, Internet linkage can provide timely reminders, assessments, behavior-tracking, and "just in time" reinforcement [101]. Supportive communications between patients and providers can occur during the critical periods of dynamic change rather than hours, days, weeks, or months later. Wearable fitness technologies have become more user-friendly with integrated feedback [104] accessed through mobile devices and Internet-linked computers [105] with reliable monitoring of physical activity at lower costs than research accelerometers [106]. Although few in number, Fitbit studies have reported 95% to 99% validity when Fitbit step counts (measured through smartphone apps) are compared with directly measured steps in healthy participants and stroke [106] and traumatic brain injury patients [107].

Despite the accumulating evidence of improved health outcomes with mobile technologies in diabetes, asthma, cardiovascular disease [102,103,108], and physical activity promotion [108], these technologies have been understudied in cancer populations [109]. To advance adoption of long-term physical activity in breast cancer survivors, our innovative health coaching intervention (iMOVE) includes applications of smartphone, computer, and wearable technologies. The pilot study will evaluate iMOVE and inform the design of a larger pragmatic randomized controlled trial (RCT).

Aims of the Pilot Study

Aim 1: To evaluate recruitment, retention, and adherence with a goal of recruiting more than 40% of eligible, contacted patients, retaining more than 75% of participants until the 6-month assessment, and seeing more than 70% of intervention components completed.

Aim 2: To evaluate acceptability feedback for intervention modification in the anticipated full-scale RCT.

Aim 3: To determine pilot estimates of intervention efficacy on fitness (primary outcome) and patient-reported, anthropometric, physical, and psychosocial outcomes (secondary outcomes).

Methods

Recruitment

This pilot RCT will enroll physically inactive breast cancer survivors stratified by age (55 years and older/younger than 55 years) and adjuvant hormone therapy (AHT) exposure (AHT or no AHT). Recruitment will be undertaken through the Princess Margaret Cancer Centre (PMCC), and interventions will occur at the Electronic Living Laboratory for Interdisciplinary Cancer Survivorship Research (ELLICSR), the Cancer Survivorship and Wellness Centre located at the Toronto General Hospital. Both institutions are members of the University Health Network in Toronto, Ontario, and research ethics board approval was obtained from the University Health Network (13-6157-DE). The trial is registered at ClinicalTrials.gov [NCT02620735].

Participants

Adult (aged 18 to 75 years) female breast cancer survivors deemed disease-free after primary cancer treatment are eligible. See Textbox 1 for selection criteria.

Recruitment and Randomization

After identification from weekly generated clinic lists and chart reviews, patients will be approached by a member of their clinical team, and interested patients will meet with a research assistant for additional study explanation and eligibility screening. Participants will also be recruited by advertisement flyers located in hospital common areas. Eligibility will be ascertained in person when possible, with written consent obtained in person prior to randomization. After participants complete baseline questionnaires and initial physiological assessments, stratification-related data (age, AHT status) will be emailed to a biostatistician in the Department of Biostatistics at PMCC who will perform randomization and then send a study identification with intervention or control group allocation.

Textbox 1. Inclusion and exclusion criteria for study.

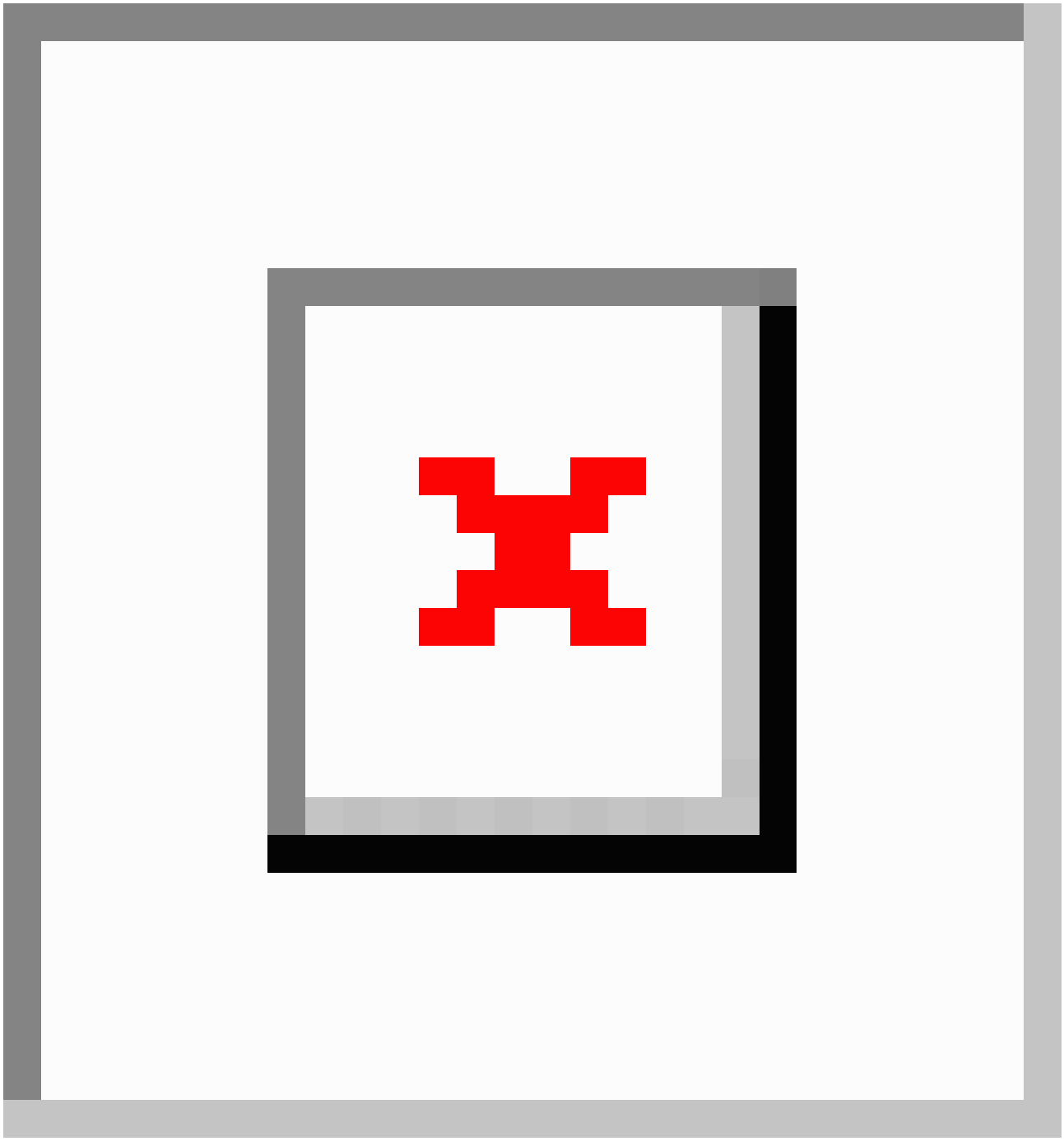
Inclusion criteria:

- Less than 2 years of completion of adjuvant therapy with the exception of hormone therapy for stage 0 to IIIA
- Self-report of fewer than 60 minutes of weekly preplanned physical activity
- Physician clearance for moderate-to-vigorous physical activity
- English proficiency
- Ability to attend exercise training sessions and study assessments at prescribed intervals for 9 months

Exclusion criteria:

- Plans to join a weight loss or exercise program within 9 months
- Current pregnancy or planned pregnancy within 9 months
- Planned surgery during study duration
- Unwillingness to be randomized

Figure 1. Exercise volume progression.



Description of Treatment Arms

Exercise Training Program (Intervention and Control Arms)

Participants in the intervention and control groups will receive individualized exercise programming progressing toward the ACSM guidelines of 150 minutes per week of moderate intensity aerobic exercise, 2 to 3 days of resistance training, and routine flexibility training.

In the first 12 weeks, intervention and control participants have 1 supervised, facility-based group exercise class plus instruction for 2 additional unsupervised, home-based exercise sessions. Facility-based sessions will be offered at a variety of times and days weekly to accommodate schedules and increase accessibility. From weeks 13 to 26, participants will complete 3 to 5 home-based exercise sessions per week and no facility-based sessions. Each exercise program is individualized based on the initial fitness assessment, physical limitations, and exercise preferences. The exercise prescriptions are developed and monitored by a certified exercise physiologist (CEP) and a registered kinesiologist (RKin) with volume progression over the intervention course (see Figure 1). Facility-based aerobic training consists of low-impact exercises or aerobic training machines. Home-based aerobic training includes participant-selected aerobic exercises including brisk walking and cycling. Resistance exercise (facility- and home-based) is completed using resistance bands and stability balls provided

to participants. The training intensity is based on exercise performance during group sessions observed by the CEP and RKin and is self-monitored via the 20-point Borg Scale for Rating of Perceived Exertion, with a target zone of 14 to 15, varying in accord with the subjective exercise experience of each participant. Adaptations to the exercise prescription based on the participant's experiences, preferences, and changes (improvements or decrements) in physical capacity will be made by the CEP and RKin to optimize intervention efficacy and safety. All participants receive exercise manuals with exercise descriptions (eg, instructive photographs, exercise safety guidelines, stretching instructions) and a weekly exercise log to review with the CEP and RKin. This program is based on the ACSM guidelines [36] and modeled after a successful theory-based program developed by research team members [110-112].

iMOVE Health Coach (Intervention Arm)

Intervention group participants are additionally provided with a technology-enabled health coach intervention (iMOVE) with 3 components: one-on-one telephone-based counseling, supportive health tracking smartphone software (Connected Wellness, NexJ Health Inc), and use of Fitbit and associated software (FitBit Flex, Fitbit Inc). The iMOVE intervention is intended to enhance sustained behavior change (physical activity) by integrating several active ingredients outlined in the cancer-survivor behavior change literature [113,114] and based on multiple theories, specifically motivational interviewing (MI) [115], cognitive behavioral therapy (CBT) [116], and relapse prevention therapy [117,118].

Theoretical constructs focus on promoting motivation and establishing exercise self-efficacy, exercise social support, and positive exercise-related feelings during the acute intervention (12 weeks) that are sustainable during the postexercise program period (6 months). The telephone-based health coaching component of iMOVE includes 10 30-minute telephone calls with a trained health coach, scheduled at weeks 1, 2, 3, 4, 5, 6, 8, and 12 (during the exercise program) and at weeks 20 and 28 (postexercise program booster sessions). Health coaches with a counseling background in MI and CBT are trained and supervised by a registered clinical psychologist and a Motivational Interviewing Network of Trainers–certified trainer. The training of health coaches involves instructions on breast cancer and related survivorship issues and continual exposure to the multitheoretical approach. Training proceeds weekly from the trial start to end based on case review and participant responses to the approaches implemented. Each health coach call to participants focuses on the assessment and enhancement of motivation, promotion of self-efficacy, and collaborative problem solving. Telephone-based counseling provides several advantages over face-to-face, notably the potential to reach multiregional populations, as telephone access is widely available [119] and requires no user or provider transport. The schedule provides support while building autonomy and independent motivation [120]. See Textbox 2 for a theoretical base summary.

In the scheduled telephone-based sessions, participants will interact with the Connected Wellness software (NexJ Health Inc) by smartphone or Internet-linked computer. This software, previously found effective with participants diagnosed with type 2 diabetes [121-123], has now been tailored for breast cancer. It tracks physical activity, nutrition, pain, and psychological well-being (eg, mood, energy) and supports goal setting (with selective daily or hourly reminders). All software entries are time-stamped, allowing for graph creation that combines multiple trackers, enabling participants and health coaches to see change indicators in relation to the physical activity levels undertaken. Every initiation of contact by participants with their health coach via text messaging is recorded. Confirmations of received text messages are provided by coaches to participants. While patients are encouraged to further discuss their texts during the next phone session, there is also provision for the health coaches to text message responses immediately, responding to questions and issues raised.

Use of the Fitbit Flex provides further assistance to participants in adhering to recommended physical activity routines and tracking physical activity, notably providing real-time feedback (light-emitting diode device lights indicate percent completion of preset, daily step goals). Additional connectivity in the Health Coach program allows the participant and health coach to jointly explore daily physical activity experiences. Fitbit Flex vibrates when preset goals (eg, 10,000 steps) are reached and records the steps taken, combining them with user data to calculate distance walked, calories burned, and the duration and intensity of activity. Fitbit Flex also measures sleep quality by tracking periods of restlessness (ie, how long it takes the wearer to fall asleep per detected body movement) and the estimated sleep duration. The user can monitor their own activity on the Health Coach platform and create summaries and periodic analyses.

As is common with behavioral interventions, a handbook specifies sessional objectives and provides clinical tools for health coaches to use each session. The health coach creates a session-by-session agenda based on patient goals, monitored activity, and motivations as collected with the software during intervals between sessions and at each session. MI and CBT are the core health behavior change theories employed. MI is a collaborative counseling method that elicits and strengthens motivation for change by addressing and resolving ambivalence [124]. MI has demonstrated effectiveness in increasing physical activity in cancer survivors and those with other chronic conditions [41,95,125-132], and positive MI-related effects have been longitudinally detected (eg, at 2 years postintervention) [126,133]. In instances when self-efficacy is impeded by distorted cognitions, CBT principles will be applied, particularly to influence affect-balance through cognitive modifications that prevent or ameliorate negative mood fluctuations [117]. Telephone-based interventions have been effective and acceptable to breast cancer patients [134-137] and useful in delivering MI and CBT interventions [138-140]. Intervention fidelity will be assessed by routine reviews of implementation variables.

While there are multiple theoretical models integrated within our intervention, these models are consolidated in the focus on addressing and resolving motivational ambivalence and identification and modification of the cognitive distortions that

maintain motivational ambivalence and prevent adoption of appropriate health behaviors.

Textbox 2. Theoretical base of iMOVE intervention.

Multifactor focus:

- Intrinsic motivations
- Perceived costs-benefits
- Identification of barriers
- Abilities to change
- Exercise self-efficacy
- Exercise social support
- Positive exercise-related feelings

Application of evidence-based theory:

- Transtheoretical model
- Social cognitive theory
- Motivational interviewing
- Cognitive behavioral therapy
- Relapse prevention theory
- Theory of planned behavior

Successful physical activity promotion features:

- Self-directed physical activity with more than 4 sessions of counseling guidance
- Follow-up behavioral prompts

Unique tailoring to breast cancer survivors:

- Program pacing per treatment sequelae (eg, physical pain and discomfort)
- Cognitive emphasis on protective physical activity effects (eg, prevention of breast cancer recurrence)
- Cognitive behavioral emphasis on paced, regulated physical activity increases
- Flexible exercise prescription for protective goals and incremental increases that optimize fitness and well-being

Technological assistance:

- Rapidly increasing role for mobile technology in health management
- Enabling patient-provider contacts during critical periods
- User-friendly wearable technologies (95%-99% validity on step counts)

Outcome Measures

Pilot Outcome Measures

These measures reflect appropriateness and effectiveness of design features:

1. Recruitment rate: based on Consolidated Standards of Reporting Trials criteria [141] via a screening log that enables data collection on eligible consented (pre- and post-initial screen) and eligible but nonrecruited individuals with nonrecruitment reasons documented.
2. Retention rate over the trial duration: the percentage of participants who complete the interventions and each data point; with reasons for drop out documented.
3. Capture of outcomes: recording of the proportion of participants at each time assessment point with complete or missing data.
4. Treatment implementation and fidelity: implementation of telephone sessions for the intervention group will be assessed by use documentation of the health coaching techniques and tools and identified barriers. Data from the health coaching software is stored on secure server and used to measure and analyze self-report and health coach activity.
5. Acceptability: telephone interviews will be conducted with a randomly selected subsample (n=25) of intervention participants following intervention completion. The goals are to explore participant perspectives of intervention feasibility and acceptability and to gain an understanding of experiences among those successful and unsuccessful at

physical activity maintenance over differing time periods (eg, during initial 3 months of intervention, 6 months of intervention, 9 months of follow-up). An interpretive descriptive qualitative methodology will be used [142], and a record of interview participation will be kept to distinguish participants from those who don't participate. The semistructured interviews will be about 45 minutes in duration and preceded by verbal informed consent. Interviews will be audiorecorded and transcribed verbatim.

Clinical Outcomes

Measures for fitness (primary), self-report (secondary), and anthropometric and physical outcomes (exploratory outcomes) are repeated at baseline, T1 (immediately after exercise program, month 3), and T2 (6 months after exercise program, month 9).

Primary Clinical Outcome

Cardiorespiratory fitness will be assessed by a graded exercise test using the modified Bruce protocol [143]. Directly measured peak volume of oxygen (mL/kg/minute) and anaerobic threshold will be obtained using a metabolic cart (TrueOne 2400, Parvo Medics) with continuous gas exchange analysis during incremental treadmill walking to volitional peak capacity. Blood pressure and arterial oxygen saturations are measured at rest and during exercise. Absolute and relative test termination criteria are based on standardized guidelines[144].

Secondary Clinical Outcomes

We will gather preliminary data on a number of exploratory variables which have been identified as important to understanding the potential impact of the intervention on patient-relevant and clinically-relevant outcomes. They are being collected to examine whether they are feasible to collect in a larger trial and whether they are responsive (sensitive to change) to the intervention [145,146].

Patient-Reported Clinical Outcomes

1. Godin-Shepherd Leisure-Time Exercise Questionnaire: a brief validated 3-item questionnaire that asks respondents to report on typical weekly exercise habits [147]
2. Functional Assessment of Cancer Therapy—Breast: generic quality of life measured with 44 self-report items [148]
3. Spielberger's State-Trait Anxiety Inventory—State [149]: a widely used 20-item measure of state anxiety
4. Center for Epidemiological Studies—Depression Scale short form: a 10-item self-report measure of depression
5. Functional Assessment of Cancer Therapy—Fatigue subscale [150]: a 13-item measure of fatigue in cancer patients
6. Breast Cancer Prevention Trial Symptoms Scale: a 42-item scale to assess side effects associated with the treatment of breast cancer
7. Fear of Recurrence Questionnaire: assesses anxiety about breast cancer recurrence [151]
8. Physical Activity Group Environment Questionnaire [152]: assesses group cohesion during exercise
9. Brief Pain Inventory [153]: a widely used measure to rapidly assess the severity of pain and its impact on functioning
10. Multiple Intervention Satisfaction Survey: an investigator-generated instrument that facilitates intervention participants in rank-ordering discrete intervention components with respect to how helpful they are in achieving outcomes during study participation. Additional items facilitate participant suggestions for deleting intervention components deemed (by participants) as of negligible benefit.

Anthropometric Clinical Outcomes

1. Body composition is assessed via body mass index, waist circumference, and body fat percentage
2. Waist circumference is measured according to the World Health Organization protocol (midpoint between lowest rib and iliac crest)
3. Body fat percentage is measured using bioelectrical impedance analysis [144]
4. Grip strength is measured using a Jamar dynamometer according to the Canadian Society for Exercise Physiology 2004 protocol

Results

Primary-Secondary Outcome Assessment

Recruitment and retention rates will be assessed [154] with estimates for participants with complete data per outcome and time point divided by the total number of study participants. Interpretation of the interview output (acceptability) will be based on inductive and deductive analyses and use of the constant comparative method [155].

Variability of the main and interaction effects will be examined in the primary clinical outcome (cardiovascular fitness) and each secondary outcome using separate repeated measures analysis of covariance models with Bonferroni corrections applied to the models. Hedges' *g* and associated confidence intervals [156] will be calculated as an estimate of the effect size both over time (within groups) and between groups [157]. Missing data will be evaluated on a case-by-case basis such that drop-outs will be excluded. Intention-to-treat (all consented subjects) analysis will employ a last observation carried forward approach to evaluate all data collected. Per protocol analysis will evaluate data on subjects who participated in 50% of group exercise sessions (comparing intervention with control subjects), while experimental subjects will have the additional criteria of participating in 50% of health coaching calls.

Sample Size and Power

We previously conducted a simulation for a range of sample sizes and different SD values for precision of the treatment effect estimate. The precision of the estimate is represented by the inverse of the margin of error. Type I error was set at α=0.05 and power at 80%. From our simulation result, a sample size of 35 to 40 was at the elbow point of the curves, indicating the precision of estimates did not proportionally increase with a larger sample size. Therefore, our projected sample size is 80 participants (40 per arm) [158]. With an anticipated drop-out rate of up to 25% [73], we will recruit 107 participants and examine the variance in primary outcomes with precision (low standard error >0.1), while enabling further calculations of effect sizes for planning the phase 3 trial [154]. The large majority of women return to PMCC for follow-up appointments typically

scheduled every 3 to 6 months. Based on data from the PMCC registry, eligibility criteria, and expected participation rates, we anticipate recruiting 8 to 10 participants per month. Study duration is estimated at 30 months.

Interpretation of Results

Interpretation of the effect size and mean difference scores and calculation of the sample size for a larger RCT (fitness outcome) will be based on a minimally important clinical difference (MCD) of 3.5 mL/kg/minute (peak volume of oxygen) between the 2 experimental groups at the 6-month T2 assessment [159,160]. We regard the MCD as a small effect size [159]. With pilot results, we will better estimate small, medium, and large effect sizes for the planned (full-scale) RCT.

Discussion

While current data suggest an important role for physical activity in disease control and the long-term health of cancer survivors, most breast cancer survivors are inactive. This discrepancy must be addressed with physical activity promotion that supports long-term exercise adherence. To date, research has focused on specific physical activity components linked to clinical benefits, but insufficient attention has been paid to factors influencing long-term physical activity maintenance. The current project employs a behavioral support intervention that assists breast cancer survivors in adopting physical activity and maintaining physical activity adherence. While multiple RCTs demonstrate effectiveness in physical activity participation during trial conduct, decreases in physical activity after trial conclusion are an important concern. It is not yet known the degree to which smartphone-enabled health coaching combined with wearable fitness technology can contribute to the lifestyle changes required for breast cancer survivors to maintain healthy physical activity over the longer term.

Our commitment to the devised intervention (combining phone-based health coaching, Fitbit step tracking, health tracking software, face-to-face exercise classes, and fitness testing) accepts the design limitation of being unable to identify which intervention components provide key contributions to significant effects; future studies may be needed to tease out what worked best in further streamlining the intervention. However, we have mitigated limitations by logging all phone counseling calls undertaken (registering time durations per call) and additionally itemizing and quantifying all use of the health tracking software, Connected Wellness (NexJ Health Inc). Furthermore, we track all Fitbit use, including use patterns per time period (day, week, and month). Additionally, use of the Multiple Intervention Satisfaction Survey facilitates each participant in subjectively ranking the intervention components on importance and suggesting deletions of components that have not been significantly helpful. These efforts will enable us to learn about the prioritization of intervention components from each subject's perspective. Another limitation entails not knowing which allocations of staff time (to the intervention) represent a cost savings when compared to other physical activity promotion approaches. Therefore, we will carefully assess staff time, preparing for ascertaining this cost dimension in the future.

This pilot will document the implementation of the methods and intervention, preliminary outcomes, and acceptability of the interventions by qualitative interview. It will assess effect size in primary and multiple secondary outcomes with corresponding confidence intervals for more definitive sample size calculations. Although pilot results will provide a foundation for full-scale RCT planning, we anticipate challenges for which we currently have only partial or potential solutions. For example, we will only have suggestive data for assessing specific intervention components and for selecting the optimal subset for full RCT testing. Furthermore, as a pilot study, we are still refining the ultimate sample size of the planned full-scale RCT. Additionally, while control subjects receive an approximation of current standard care for exercise promotion (at ELLICSR), we cannot fully account for the attentional differences in intervention and control conditions. Nonetheless, this pilot is a distinct step forward in addressing a gap in the promotion of longer term exercise adherence for breast cancer survivors.

Conflicts of Interest

None declared.

References

1. Canadian Cancer Society. Canadian Cancer Statistics: Prostate cancer statistics at a glance. URL: http://www.cancer.ca/en/cancer-information/cancer-type/prostate/statistics/?region=sk [accessed 2017-07-17] [WebCite Cache ID 6s26yQ06p]
2. Altekruse S, Kosary C, Krapcho M, Neyman N, Aminou R, Waldron W, et al. SEER cancer statistics review, 1975-2007. Bethesda: National Cancer Institute URL: https://seer.cancer.gov/archive/csr/1975_2007/ [accessed 2017-07-17] [WebCite Cache ID 6s276DQFN]
3. Lipscombe LL, Chan WW, Yun L, Austin PC, Anderson GM, Rochon PA. Incidence of diabetes among postmenopausal breast cancer survivors. Diabetologia 2013 Mar;56(3):476-483. [doi: 10.1007/s00125-012-2793-9] [Medline: 23238788]
4. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Kosary CL, et al, editors. SEER Cancer Statistics Review, 1975-2014. Bethesda: National Cancer Institute; 2014.
5. Miller KD, Triano LR. Medical issues in cancer survivors—a review. Cancer J 2008;14(6):375-387. [doi: 10.1097/PPO.0b013e31818ee3dc] [Medline: 19060602]
6. Ganz PA. Survivorship: adult cancer survivors. Prim Care 2009 Dec;36(4):721-741. [doi: 10.1016/j.pop.2009.08.001] [Medline: 19913184]

7. Carver JR, Shapiro CL, Ng A, Jacobs L, Schwartz C, Virgo KS, ASCO Cancer Survivorship Expert Panel. American Society of Clinical Oncology clinical evidence review on the ongoing care of adult cancer survivors: cardiac and pulmonary late effects. J Clin Oncol 2007 Sep 01;25(25):3991-4008. [doi: 10.1200/JCO.2007.10.9777] [Medline: 17577017]
8. Stone PC, Minton O. Cancer-related fatigue. Eur J Cancer 2008 May;44(8):1097-1104. [doi: 10.1016/j.ejca.2008.02.037] [Medline: 18381237]
9. Fosså SD, Vassilopoulou-Sellin R, Dahl AA. Long-term physical sequelae after adult-onset cancer. J Cancer Surviv 2008 Mar;2(1):3-11. [doi: 10.1007/s11764-007-0039-5] [Medline: 18648982]
10. Stewart A, Bielajew C, Collins B, Parkinson M, Tomiak E. A meta-analysis of the neuropsychological effects of adjuvant chemotherapy treatment in women treated for breast cancer. Clin Neuropsychol 2006 Feb;20(1):76-89. [doi: 10.1080/138540491005875] [Medline: 16410227]
11. Burton AW, Fanciullo GJ, Beasley RD, Fisch MJ. Chronic pain in the cancer survivor: a new frontier. Pain Med 2007 Mar;8(2):189-198. [doi: 10.1111/j.1526-4637.2006.00220.x] [Medline: 17305690]
12. Camp-Sorrell D. Cardiorespiratory effects in cancer survivors. Cancer Nurs 2006;29(2 Suppl):55-59. [Medline: 16779957]
13. Pelusi J. Sexuality and body image. Research on breast cancer survivors documents altered body image and sexuality. Am J Nurs 2006 Mar;106(3 Suppl):32-38. [Medline: 16481849]
14. Azim HA, de Azambuja E, Colozza M, Bines J, Piccart MJ. Long-term toxic effects of adjuvant chemotherapy in breast cancer. Ann Oncol 2011 Sep;22(9):1939-1947. [doi: 10.1093/annonc/mdq683] [Medline: 21289366]
15. Hewitt M, Greenfield S, Stovall EL. From Cancer Patient to Cancer Survivors: Lost in Transition. Washington: National Academies Press; 2006.
16. Ng AK, Kenney LB, Gilbert ES, Travis LB. Secondary malignancies across the age spectrum. Semin Radiat Oncol 2010 Jan;20(1):67-78 [FREE Full text] [doi: 10.1016/j.semradonc.2009.09.002] [Medline: 19959033]
17. Bernstein JL, Thompson WD, Risch N, Holford TR. Risk factors predicting the incidence of second primary breast cancer among women diagnosed with a first primary breast cancer. Am J Epidemiol 1992 Oct 15;136(8):925-936. [Medline: 1456269]
18. Clarke M, Collins R, Darby S, Davies C, Elphinstone P, Evans V, Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005 Dec 17;366(9503):2087-2106 [FREE Full text] [doi: 10.1016/S0140-6736(05)67887-7] [Medline: 16360786]
19. Buchanan CL, Dorn PL, Fey J, Giron G, Naik A, Mendez J, et al. Locoregional recurrence after mastectomy: incidence and outcomes. J Am Coll Surg 2006 Oct;203(4):469-474. [doi: 10.1016/j.jamcollsurg.2006.06.015] [Medline: 17000389]
20. Demark-Wahnefried W, Clipp EC, Morey MC, Pieper CF, Sloane R, Snyder DC, et al. Lifestyle intervention development study to improve physical function in older adults with cancer: outcomes from Project LEAD. J Clin Oncol 2006 Jul 20;24(21):3465-3473 [FREE Full text] [doi: 10.1200/JCO.2006.05.7224] [Medline: 16849763]
21. Davies C, Godwin J, Gray R, Clarke M, Cutter D, Darby S, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011 Aug 27;378(9793):771-784 [FREE Full text] [doi: 10.1016/S0140-6736(11)60993-8] [Medline: 21802721]
22. Ewertz M, Jensen AB. Late effects of breast cancer treatment and potentials for rehabilitation. Acta Oncol 2011 Feb;50(2):187-193. [doi: 10.3109/0284186X.2010.533190] [Medline: 21231780]
23. Ballard-Barbash R, Friedenreich CM, Courneya KS, Siddiqi SM, McTiernan A, Alfano CM. Physical activity, biomarkers, and disease outcomes in cancer survivors: a systematic review. J Natl Cancer Inst 2012 Jun 06;104(11):815-840 [FREE Full text] [doi: 10.1093/jnci/djs207] [Medline: 22570317]
24. Lemanne D, Cassileth B, Gubili J. The role of physical activity in cancer prevention, treatment, recovery, and survivorship. Oncology (Williston Park) 2013 Jun;27(6):580-585 [FREE Full text] [Medline: 23909073]
25. Kabat GC, Kim M, Caan BJ, Chlebowski RT, Gunter MJ, Ho GYF, et al. Repeated measures of serum glucose and insulin in relation to postmenopausal breast cancer. Int J Cancer 2009 Dec 01;125(11):2704-2710 [FREE Full text] [doi: 10.1002/ijc.24609] [Medline: 19588485]
26. Eliassen AH, Missmer SA, Tworoger SS, Spiegelman D, Barbieri RL, Dowsett M, et al. Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. J Natl Cancer Inst 2006 Oct 04;98(19):1406-1415. [doi: 10.1093/jnci/djj376] [Medline: 17018787]
27. Coussens LM, Werb Z. Inflammation and cancer. Nature 2002;420(6917):860-867 [FREE Full text] [doi: 10.1038/nature01322] [Medline: 12490959]
28. Lynch BM, Dunstan DW, Healy GN, Winkler E, Eakin E, Owen N. Objectively measured physical activity and sedentary time of breast cancer survivors, and associations with adiposity: findings from NHANES (2003-2006). Cancer Causes Control 2010 Feb;21(2):283-288. [doi: 10.1007/s10552-009-9460-6] [Medline: 19882359]
29. Goodwin PJ, Ennis M, Pritchard KI, Trudeau ME, Koo J, Madarnas Y, et al. Fasting insulin and outcome in early-stage breast cancer: results of a prospective cohort study. J Clin Oncol 2002 Jan 01;20(1):42-51. [doi: 10.1200/JCO.2002.20.1.42] [Medline: 11773152]
30. Pollak MN. Insulin resistance, estimated by serum C-peptide level, is associated with reduced event-free survival for postmenopausal women in NCIC CTG MA.14 adjuvant breast cancer trial. J Clin Oncol 2006;24(18):524.

31. Irwin ML, Aiello EJ, McTiernan A, Bernstein L, Gilliland FD, Baumgartner RN, et al. Physical activity, body mass index, and mammographic density in postmenopausal breast cancer survivors. J Clin Oncol 2007 Mar 20;25(9):1061-1066 [FREE Full text] [doi: 10.1200/JCO.2006.07.3965] [Medline: 17261853]
32. World Cancer Research Fund. Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective. Washington: American Institute for Cancer Research; 2007.
33. Rock CL, Doyle C, Demark-Wahnefried W, Meyerhardt J, Courneya KS, Schwartz AL, et al. Nutrition and physical activity guidelines for cancer survivors. CA Cancer J Clin 2012;62(4):243-274 [FREE Full text] [doi: 10.3322/caac.21142] [Medline: 22539238]
34. Brunet J, Meterissian S, Sabiston CM. Physical activity and breast cancer survivorship: evidence-based recommendations. Am J Lifestyle Med 2012;6(6):224-240.
35. Schmitz KH, Courneya KS, Matthews C, Demark-Wahnefried W, Galvão DA, Pinto BM, et al. American College of Sports Medicine roundtable on exercise guidelines for cancer survivors. Med Sci Sports Exerc 2010 Jul;42(7):1409-1426. [doi: 10.1249/MSS.0b013e3181e0c112] [Medline: 20559064]
36. Irwin ML, Crumley D, McTiernan A, Bernstein L, Baumgartner R, Gilliland FD, et al. Physical activity levels before and after a diagnosis of breast carcinoma: the Health, Eating, Activity, and Lifestyle (HEAL) study. Cancer 2003 Apr 01;97(7):1746-1757 [FREE Full text] [doi: 10.1002/cncr.11227] [Medline: 12655532]
37. Irwin ML, Ainsworth BE. Physical activity interventions following cancer diagnosis: methodologic challenges to delivery and assessment. Cancer Invest 2004;22(1):30-50. [Medline: 15069762]
38. Irwin ML, Duggan C, Wang C, Smith AW, McTiernan A, Baumgartner RN, et al. Fasting C-peptide levels and death resulting from all causes and breast cancer: the health, eating, activity, and lifestyle study. J Clin Oncol 2011 Jan 01;29(1):47-53 [FREE Full text] [doi: 10.1200/JCO.2010.28.4752] [Medline: 21115859]
39. Courneya KS, Katzmarzyk PT, Bacon E. Physical activity and obesity in Canadian cancer survivors: population-based estimates from the 2005 Canadian Community Health Survey. Cancer 2008 Jun;112(11):2475-2482 [FREE Full text] [doi: 10.1002/cncr.23455] [Medline: 18428195]
40. Pinto BM, Trunzo JJ, Reiss P, Shiu S. Exercise participation after diagnosis of breast cancer: trends and effects on mood and quality of life. Psychooncology 2002;11(5):389-400. [doi: 10.1002/pon.594] [Medline: 12228872]
41. Pinto BM, Frierson GM, Rabin C, Trunzo JJ, Marcus BH. Home-based physical activity intervention for breast cancer patients. J Clin Oncol 2005 May 20;23(15):3577-3587 [FREE Full text] [doi: 10.1200/JCO.2005.03.080] [Medline: 15908668]
42. Blanchard CM, Courneya KS, Stein K. Cancer survivors' adherence to lifestyle behavior recommendations and associations with health-related quality of life: results from the American Cancer Society's SCS-II. J Clin Oncol 2008 May 1;26(13):2198-2204 [FREE Full text] [doi: 10.1200/JCO.2007.14.6217] [Medline: 18445845]
43. Blanchard CM, Denniston MM, Baker F, Ainsworth SR, Courneya KS, Hann DM, et al. Do adults change their lifestyle behaviors after a cancer diagnosis? Am J Health Behav 2003;27(3):246-256. [Medline: 12751621]
44. Centers for Disease Control and Prevention. A National Action Plan for Cancer Survivorship: Advancing Public Health Strategies. Atlanta: Centers for Disease Control and Prevention; 2004.
45. Harrison S, Hayes SC, Newman B. Level of physical activity and characteristics associated with change following breast cancer diagnosis and treatment. Psychooncology 2009 Apr;18(4):387-394. [doi: 10.1002/pon.1504] [Medline: 19117320]
46. Demark-Wahnefried W, Peterson B, McBride C, Lipkus I, Clipp E. Current health behaviors and readiness to pursue life-style changes among men and women diagnosed with early stage prostate and breast carcinomas. Cancer 2000 Feb 1;88(3):674-684. [Medline: 10649263]
47. Demark-Wahnefried W, Aziz NM, Rowland JH, Pinto BM. Riding the crest of the teachable moment: promoting long-term health after the diagnosis of cancer. J Clin Oncol 2005 Aug 20;23(24):5814-5830 [FREE Full text] [doi: 10.1200/JCO.2005.01.230] [Medline: 16043830]
48. Bellizzi KM, Rowland JH, Jeffery DD, McNeel T. Health behaviors of cancer survivors: examining opportunities for cancer control intervention. J Clin Oncol 2005 Dec 1;23(34):8884-8893. [doi: 10.1200/JCO.2005.02.2343] [Medline: 16314649]
49. Coups EJ, Ostroff JS. A population-based estimate of the prevalence of behavioral risk factors among adult cancer survivors and noncancer controls. Prev Med 2005 Jun;40(6):702-711. [doi: 10.1016/j.ypmed.2004.09.011] [Medline: 15850868]
50. Pinto BM, Eakin E, Maruyama NC. Health behavior changes after a cancer diagnosis: what do we know and where do we go from here? Ann Behav Med 2000;22(1):38-52. [Medline: 10892527]
51. Lemon SC, Zapka JG, Clemow L. Health behavior change among women with recent familial diagnosis of breast cancer. Prev Med 2004 Aug;39(2):253-262. [doi: 10.1016/j.ypmed.2004.03.039] [Medline: 15226033]
52. Blaney JM, Lowe-Strong A, Rankin-Watt J, Campbell A, Gracey JH. Cancer survivors' exercise barriers, facilitators and preferences in the context of fatigue, quality of life and physical activity participation: a questionnaire-survey. Psychooncology 2013 Jan;22(1):186-194. [doi: 10.1002/pon.2072] [Medline: 23296635]
53. Demark-Wahnefried W, Aziz NM, Rowland JH, Pinto BM. Riding the crest of the teachable moment: promoting long-term health after the diagnosis of cancer. J Clin Oncol 2005 Aug 20;23(24):5814-5830 [FREE Full text] [doi: 10.1200/JCO.2005.01.230] [Medline: 16043830]

54. Speck RM, Courneya KS, Mâsse LC, Duval S, Schmitz KH. An update of controlled physical activity trials in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2010 Jun;4(2):87-100. [doi: 10.1007/s11764-009-0110-5] [Medline: 20052559]
55. Schmitz KH, Holtzman J, Courneya KS, Mâsse LC, Duval S, Kane R. Controlled physical activity trials in cancer survivors: a systematic review and meta-analysis. Cancer Epidemiol Biomarkers Prev 2005 Jul;14(7):1588-1595 [FREE Full text] [doi: 10.1158/1055-9965.EPI-04-0703] [Medline: 16030088]
56. Fjeldsoe B, Neuhaus M, Winkler E, Eakin E. Systematic review of maintenance of behavior change following physical activity and dietary interventions. Health Psychol 2011 Jan;30(1):99-109. [doi: 10.1037/a0021974] [Medline: 21299298]
57. Mishra SI, Scherer RW, Snyder C, Geigle PM, Berlanstein DR, Topaloglu O. Exercise interventions on health-related quality of life for people with cancer during active treatment. Cochrane Database Syst Rev 2012:CD008465. [doi: 10.1002/14651858.CD008465.pub2] [Medline: 22895974]
58. Wolin KY, Schwartz AL, Matthews CE, Courneya KS, Schmitz KH. Implementing the exercise guidelines for cancer survivors. J Support Oncol 2012;10(5):171-177 [FREE Full text] [doi: 10.1016/j.suponc.2012.02.001] [Medline: 22579268]
59. Cramp F, James A, Lambert J. The effects of resistance training on quality of life in cancer: a systematic literature review and meta-analysis. Support Care Cancer 2010 Nov;18(11):1367-1376. [doi: 10.1007/s00520-010-0904-z] [Medline: 20502922]
60. Focht BC, Clinton SK, Devor ST, Garver MJ, Lucas AR, Thomas-Ahner JM, et al. Resistance exercise interventions during and following cancer treatment: a systematic review. J Support Oncol 2013 Jun;11(2):45-60. [Medline: 23967493]
61. Galvão DA, Newton RU. Review of exercise intervention studies in cancer patients. J Clin Oncol 2005 Feb 01;23(4):899-909. [doi: 10.1200/JCO.2005.06.085] [Medline: 15681536]
62. Bicego D, Brown K, Ruddick M, Storey D, Wong C, Harris SR. Effects of exercise on quality of life in women living with breast cancer: a systematic review. Breast J 2009;15(1):45-51. [doi: 10.1111/j.1524-4741.2008.00670.x] [Medline: 19120381]
63. Pekmezi DW, Demark-Wahnefried W. Updated evidence in support of diet and exercise interventions in cancer survivors. Acta Oncol 2011 Feb;50(2):167-178 [FREE Full text] [doi: 10.3109/0284186X.2010.529822] [Medline: 21091401]
64. Courneya KS, Segal RJ, Mackey JR, Gelmon K, Reid RD, Friedenreich CM, et al. Effects of aerobic and resistance exercise in breast cancer patients receiving adjuvant chemotherapy: a multicenter randomized controlled trial. J Clin Oncol 2007 Oct 1;25(28):4396-4404 [FREE Full text] [doi: 10.1200/JCO.2006.08.2024] [Medline: 17785708]
65. McNeely ML, Campbell KL, Rowe BH, Klassen TP, Mackey JR, Courneya KS. Effects of exercise on breast cancer patients and survivors: a systematic review and meta-analysis. CMAJ 2006 Jul 04;175(1):34-41 [FREE Full text] [doi: 10.1503/cmaj.051073] [Medline: 16818906]
66. Kirshbaum MN. A review of the benefits of whole body exercise during and after treatment for breast cancer. J Clin Nurs 2007 Jan;16(1):104-121. [doi: 10.1111/j.1365-2702.2006.01638.x] [Medline: 17181672]
67. Sabiston CM, Brunet J, Burke S. Pain, movement, and mind: does physical activity mediate the relationship between pain and mental health among survivors of breast cancer? Clin J Pain 2012 Jul;28(6):489-495. [doi: 10.1097/AJP.0b013e31823853ac] [Medline: 22673481]
68. Sabiston CM, McDonough MH, Crocker PRE. Psychosocial experiences of breast cancer survivors involved in a dragon boat program: exploring links to positive psychological growth. J Sport Exerc Psychol 2007 Aug;29(4):419-438. [Medline: 17968046]
69. Holmes MD, Chen WY, Feskanich D, Kroenke CH, Colditz GA. Physical activity and survival after breast cancer diagnosis. JAMA 2005 May 25;293(20):2479-2486. [doi: 10.1001/jama.293.20.2479] [Medline: 15914748]
70. Stull VB, Snyder DC, Demark-Wahnefried W. Lifestyle interventions in cancer survivors: designing programs that meet the needs of this vulnerable and growing population. J Nutr 2007 Jan;137(1 Suppl):243S-248S [FREE Full text] [Medline: 17182834]
71. Cramp F, Daniel J. Exercise for the management of cancer-related fatigue in adults. Cochrane Database Syst Rev 2008:CD006145. [doi: 10.1002/14651858.CD006145.pub2] [Medline: 18425939]
72. Estabrooks PA, Glasgow RE, Dzewaltowski DA. Physical activity promotion through primary care. JAMA 2003 Jun 11;289(22):2913-2916. [doi: 10.1001/jama.289.22.2913] [Medline: 12799388]
73. White SM, McAuley E, Estabrooks PA, Courneya KS. Translating physical activity interventions for breast cancer survivors into practice: an evaluation of randomized controlled trials. Ann Behav Med 2009 Feb;37(1):10-19 [FREE Full text] [doi: 10.1007/s12160-009-9084-9] [Medline: 19255819]
74. Wigers SH, Stiles TC, Vogel PA. Effects of aerobic exercise versus stress management treatment in fibromyalgia. A 4.5 year prospective study. Scand J Rheumatol 1996;25(2):77-86. [doi: 10.3109/03009749609069212] [Medline: 8614771]
75. Rissanen A, Fogelholm M. Physical activity in the prevention and treatment of other morbid conditions and impairments associated with obesity: current evidence and research issues. Med Sci Sports Exerc 1999 Nov;31(11 Suppl):S635-S645. [Medline: 10593540]
76. Marcus BH, Dubbert PM, Forsyth LH, McKenzie TL, Stone EJ, Dunn AL, et al. Physical activity behavior change: issues in adoption and maintenance. Health Psychol 2000 Jan;19(1S):32-41. [Medline: 10709946]
77. Kahn EB, Ramsey LT, Brownson RC, Heath GW, Howze EH, Powell KE, et al. The effectiveness of interventions to increase physical activity. A systematic review. Am J Prev Med 2002 May;22(4 Suppl):73-107. [Medline: 11985936]

78. van der Bij AK, Laurant MGH, Wensing M. Effectiveness of physical activity interventions for older adults: a review. Am J Prev Med 2002 Feb;22(2):120-133. [Medline: 11818183]
79. Smith BJ. Promotion of physical activity in primary health care: update of the evidence on interventions. J Sci Med Sport 2004 Apr;7(1 Suppl):67-73. [Medline: 15214604]
80. McAuley E, Morris KS, Motl RW, Hu L, Konopack JF, Elavsky S. Long-term follow-up of physical activity behavior in older adults. Health Psychol 2007 May;26(3):375-380. [doi: 10.1037/0278-6133.26.3.375] [Medline: 17500625]
81. Greaney ML, Riebe D, Ewing GC, Rossi JS, Lees FD, Burbank PA, et al. Long-term effects of a stage-based intervention for changing exercise intentions and behavior in older adults. Gerontologist 2008 Jun;48(3):358-367. [Medline: 18591361]
82. Rejeski WJ, Marsh AP, Chmelo E, Prescott AJ, Dobrosielski M, Walkup MP, et al. The Lifestyle Interventions and Independence for Elders Pilot (LIFE-P): 2-year follow-up. J Gerontol A Biol Sci Med Sci 2009 Apr;64(4):462-467 [FREE Full text] [doi: 10.1093/gerona/gln041] [Medline: 19181715]
83. Hall KS, Sloane R, Pieper CF, Peterson MJ, Crowley GM, Cowper PA, et al. Long-term changes in physical activity following a one-year home-based physical activity counseling program in older adults with multiple morbidities. J Aging Res 2010 Dec 26;2011:308407 [FREE Full text] [doi: 10.4061/2011/308407] [Medline: 21234104]
84. Hertogh EM, Vergouwe Y, Schuit AJ, Peeters PHM, Monninkhof EM. Behavioral changes after a 1-yr exercise program and predictors of maintenance. Med Sci Sports Exerc 2010 May;42(5):886-892. [doi: 10.1249/MSS.0b013e3181c4d964] [Medline: 19996989]
85. Spark LC, Reeves MM, Fjeldsoe BS, Eakin EG. Physical activity and/or dietary interventions in breast cancer survivors: a systematic review of the maintenance of outcomes. J Cancer Surviv 2013 Mar;7(1):74-82. [doi: 10.1007/s11764-012-0246-6] [Medline: 23179496]
86. Demark-Wahnefried W, Morey MC, Sloane R, Snyder DC, Miller PE, Hartman TJ, et al. Reach out to enhance wellness home-based diet-exercise intervention promotes reproducible and sustainable long-term improvements in health behaviors, body weight, and physical functioning in older, overweight/obese cancer survivors. J Clin Oncol 2012 Jul 01;30(19):2354-2361 [FREE Full text] [doi: 10.1200/JCO.2011.40.0895] [Medline: 22614994]
87. Short CE, James EL, Stacey F, Plotnikoff RC. A qualitative synthesis of trials promoting physical activity behaviour change among post-treatment breast cancer survivors. J Cancer Surviv 2013 Dec;7(4):570-581 [FREE Full text] [doi: 10.1007/s11764-013-0296-4] [Medline: 23888337]
88. Dacey M, Baltzell A, Zaichkowsky L. Older adults' intrinsic and extrinsic motivation toward physical activity. Am J Health Behav 2008;32(6):570-582. [doi: 10.5555/ajhb.2008.32.6.570] [Medline: 18442337]
89. Mata J, Silva MN, Vieira PN, Carraça EV, Andrade AM, Coutinho SR, et al. Motivational spill-over during weight control: increased self-determination and exercise intrinsic motivation predict eating self-regulation. Health Psychol 2009 Nov;28(6):709-716. [doi: 10.1037/a0016764] [Medline: 19916639]
90. Ory MG, Jordan PJ, Bazzarre T. The Behavior Change Consortium: setting the stage for a new century of health behavior-change research. Health Educ Res 2002 Oct;17(5):500-511. [Medline: 12408195]
91. Pinto BM, Floyd A. Theories underlying health promotion interventions among cancer survivors. Semin Oncol Nurs 2008 Aug;24(3):153-163. [doi: 10.1016/j.soncn.2008.05.003] [Medline: 18687261]
92. Vallance J, Plotnikoff RC, Karvinen KH, Mackey JR, Courneya KS. Understanding physical activity maintenance in breast cancer survivors. Am J Health Behav 2010;34(2):225-236. [Medline: 19814602]
93. Vallance JK, Courneya KS, Plotnikoff RC, Dinu I, Mackey JR. Maintenance of physical activity in breast cancer survivors after a randomized trial. Med Sci Sports Exerc 2008 Jan;40(1):173-180. [doi: 10.1249/mss.0b013e3181586b41] [Medline: 18091007]
94. Harland J, White M, Drinkwater C, Chinn D, Farr L, Howel D. The Newcastle exercise project: a randomised controlled trial of methods to promote physical activity in primary care. BMJ 1999 Sep 25;319(7213):828-832 [FREE Full text] [Medline: 10496829]
95. Elley CR, Kerse N, Arroll B, Robinson E. Effectiveness of counselling patients on physical activity in general practice: cluster randomised controlled trial. BMJ 2003 Apr 12;326(7393):793 [FREE Full text] [doi: 10.1136/bmj.326.7393.793] [Medline: 12689976]
96. Wing RR, Hamman RF, Bray GA, Delahanty L, Edelstein SL, Hill JO, et al. Achieving weight and activity goals among diabetes prevention program lifestyle participants. Obes Res 2004 Sep;12(9):1426-1434 [FREE Full text] [doi: 10.1038/oby.2004.179] [Medline: 15483207]
97. Hillsdon M, Foster C, Thorogood M. Interventions for promoting physical activity. Cochrane Database Syst Rev 2005:CD003180 [FREE Full text] [doi: 10.1002/14651858.CD003180.pub2] [Medline: 15674903]
98. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
99. Bourke L, Homer KE, Thaha MA, Steed L, Rosario DJ, Robb KA, et al. Interventions for promoting habitual exercise in people living with and beyond cancer. Cochrane Database Syst Rev 2013 Jan:CD010192. [doi: 10.1002/14651858.CD010192.pub2] [Medline: 24065550]
100. CRTC communications monitoring report: 5.0 telecommunications. URL: http://www.crtc.gc.ca/eng/publications/reports/PolicyMonitoring/2012/cmr5.htm#n0 [accessed 2017-07-18] [WebCite Cache ID 6s3OdoFJK]

101. Smyth J, Heron K. Health psychology. In: Mehl MR, Connor TS, editors. Handook of Research Methods for Studying Daily Life. New York: Guilford Press; 2011:569.
102. Boulos MNK, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online 2011;10:1-15 [FREE Full text] [doi: 10.1186/1475-925X-10-24] [Medline: 21466669]
103. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
104. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
105. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015;12:9 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]
106. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. J Sci Med Sport 2014 Sep;17(5):496-500. [doi: 10.1016/j.jsams.2013.10.241] [Medline: 24268570]
107. Fulk GD, Combs SA, Danks KA, Nirider CD, Raja B, Reisman DS. Accuracy of 2 activity monitors in detecting steps in people with stroke and traumatic brain injury. Phys Ther 2014 Feb;94(2):222-229 [FREE Full text] [doi: 10.2522/ptj.20120525] [Medline: 24052577]
108. van den Berg MH, Schoones JW, Vliet VTPM. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
109. Hong Y, Dahlke DV, Ory M, Hochhalter A, Reynolds J, Purcell NP, et al. Designing iCanFit: a mobile-enabled Web application to promote physical activity for older cancer survivors. JMIR Res Protoc 2013;2(1):e12 [FREE Full text] [doi: 10.2196/resprot.2440] [Medline: 23612053]
110. Santa MD, Alibhai SMH, 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]
111. Alibhai SMH, Santa MD, Ritvo P, Sabiston C, Krahn M, Tomlinson G, et al. A phase II RCT and economic analysis of three exercise delivery methods in men with prostate cancer on androgen deprivation therapy. BMC Cancer 2015 Apr 25;15:312 [FREE Full text] [doi: 10.1186/s12885-015-1316-8] [Medline: 25908311]
112. Santa MD, Alibhai SMH, Matthew AG, Guglietti CL, Steele J, Trachtenberg J, et al. Exercise in clinical cancer care: a call to action and program development description. Curr Oncol 2012 Jun;19(3):e136-e144 [FREE Full text] [doi: 10.3747/co.19.912] [Medline: 22670103]
113. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
114. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
115. Rollnick S, Miller WR, Butler CC. Motivational Interviewing in Health Care: Helping Patients Change Behavior. New York: Guilford Press; 2007.
116. Brewin CR. Theoretical foundations of cognitive-behavior therapy for anxiety and depression. Annu Rev Psychol 1996;47:33-57. [doi: 10.1146/annurev.psych.47.1.33] [Medline: 8624137]
117. Marlatt GA, Gordon JR. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. New York: Guilford Press; 1985.
118. Parks GA, Anderson BK, Marlatt GA. Relapse prevention therapy. In: Heather N, Peters T, editors. Handbook of Alcohol Dependence and Problems. Sussex: John Wiley & Sons, Ltd; 2001.
119. Eakin EG, Lawler SP, Vandelanotte C, Owen N. Telephone interventions for physical activity and dietary behavior change: a systematic review. Am J Prev Med 2007 May;32(5):419-434. [doi: 10.1016/j.amepre.2007.01.004] [Medline: 17478269]
120. Ang DC, Kaleth AS, Bigatti S, Mazzuca S, Saha C, Hilligoss J, et al. Research to Encourage Exercise for Fibromyalgia (REEF): use of motivational interviewing design and method. Contemp Clin Trials 2011 Jan;32(1):59-68 [FREE Full text] [doi: 10.1016/j.cct.2010.08.014] [Medline: 20828634]
121. Wayne N, Perez DF, Kaplan DM, Ritvo P. Health coaching reduces HbA1c in type 2 diabetic patients from a lower-socioeconomic status community: a randomized controlled trial. J Med Internet Res 2015;17(10):e224 [FREE Full text] [doi: 10.2196/jmir.4871] [Medline: 26441467]
122. Pludwinski S, Ahmad F, Wayne N, Ritvo P. Participant experiences in a smartphone-based health coaching intervention for type 2 diabetes: a qualitative inquiry. J Telemed Telecare 2015 Jul 21:172-178. [doi: 10.1177/1357633X15595178] [Medline: 26199275]
123. Wayne N, Ritvo P. Smartphone-enabled health coach intervention for people with diabetes from a modest socioeconomic strata community: single-arm longitudinal feasibility study. J Med Internet Res 2014;16(6):e149 [FREE Full text] [doi: 10.2196/jmir.3180] [Medline: 24907918]

124. Miller WR, Rollnick S. Ten things that motivational interviewing is not. Behav Cogn Psychother 2009 Mar;37(2):129-140. [doi: 10.1017/S1352465809005128] [Medline: 19364414]
125. Sjöling M, Lundberg K, Englund E, Westman A, Jong MC. Effectiveness of motivational interviewing and physical activity on prescription on leisure exercise time in subjects suffering from mild to moderate hypertension. BMC Res Notes 2011 Sep 12;4:352 [FREE Full text] [doi: 10.1186/1756-0500-4-352] [Medline: 21906404]
126. Burke BL, Arkowitz H, Menchola M. The efficacy of motivational interviewing: a meta-analysis of controlled clinical trials. J Consult Clin Psychol 2003 Oct;71(5):843-861. [doi: 10.1037/0022-006X.71.5.843] [Medline: 14516234]
127. Brodie DA, Inoue A. Motivational interviewing to promote physical activity for people with chronic heart failure. J Adv Nurs 2005 Jun;50(5):518-527. [doi: 10.1111/j.1365-2648.2005.03422.x] [Medline: 15882368]
128. Smith DE, Heckemeyer CM, Kratt PP, Mason DA. Motivational interviewing to improve adherence to a behavioral weight-control program for older obese women with NIDDM. A pilot study. Diabetes Care 1997 Jan;20(1):52-54. [Medline: 9028693]
129. Butterworth S, Linden A, McClay W. Health coaching as an intervention in health management programs. Dis Manag Health Out 2007;15(5):299-307.
130. Bennett JA, Perrin NA, Hanson G, Bennett D, Gaynor W, Flaherty-Robb M, et al. Healthy aging demonstration project: nurse coaching for behavior change in older adults. Res Nurs Health 2005 Jun;28(3):187-197. [doi: 10.1002/nur.20077] [Medline: 15884026]
131. Tse MMY, Vong SKS, Tang SK. Motivational interviewing and exercise programme for community-dwelling older persons with chronic pain: a randomised controlled study. J Clin Nurs 2013 Jul;22(13-14):1843-1856. [doi: 10.1111/j.1365-2702.2012.04317.x] [Medline: 23279630]
132. Olsen JM, Nesbitt BJ. Health coaching to improve healthy lifestyle behaviors: an integrative review. Am J Health Promot 2010;25(1):e1-e12. [doi: 10.4278/ajhp.090313-LIT-101] [Medline: 20809820]
133. Lundahl BW. A meta-analysis of motivational interviewing: twenty-five years of empirical studies. Res Social Work Prac 2010;20(2):137-160.
134. Gotay CC, Moinpour CM, Unger JM, Jiang CS, Coleman D, Martino S, et al. Impact of a peer-delivered telephone intervention for women experiencing a breast cancer recurrence. J Clin Oncol 2007 May 20;25(15):2093-2099 [FREE Full text] [doi: 10.1200/JCO.2006.07.4674] [Medline: 17513815]
135. Badger T, Segrin C, Meek P, Lopez AM, Bonham E, Sieger A. Telephone interpersonal counseling with women with breast cancer: symptom management and quality of life. Oncol Nurs Forum 2005 Mar 05;32(2):273-279. [doi: 10.1188/05.ONF.273-279] [Medline: 15759065]
136. Allen SM, Shah AC, Nezu AM, Nezu CM, Ciambrone D, Hogan J, et al. A problem-solving approach to stress reduction among younger women with breast carcinoma: a randomized controlled trial. Cancer 2002 Jun 15;94(12):3089-3100 [FREE Full text] [doi: 10.1002/cncr.10586] [Medline: 12115339]
137. Marcus AC, Garrett KM, Cella D, Wenzel L, Brady MJ, Fairclough D, et al. Can telephone counseling post-treatment improve psychosocial outcomes among early stage breast cancer survivors? Psychooncology 2010 Sep;19(9):923-932 [FREE Full text] [doi: 10.1002/pon.1653] [Medline: 19941285]
138. Wahab S, Menon U, Szalacha L. Motivational interviewing and colorectal cancer screening: a peek from the inside out. Patient Educ Couns 2008 Aug;72(2):210-217 [FREE Full text] [doi: 10.1016/j.pec.2008.03.023] [Medline: 18467066]
139. Picciano J, Roffman RA, Kalichman SC, Rutledge SE, Berghuis JP. A telephone based brief intervention using motivational enhancement to facilitate HIV risk reduction among MSM: a pilot study. AIDS Behav 2001;5(3):251-262.
140. Walker DD, Roffman RA, Picciano JF, Stephens RS. The check-up: in-person, computerized, and telephone adaptations of motivational enhancement treatment to elicit voluntary participation by the contemplator. Subst Abuse Treat Prev Policy 2007 Jan 08;2:2 [FREE Full text] [doi: 10.1186/1747-597X-2-2] [Medline: 17210075]
141. Carroll KM. Enhancing retention in clinical trials of psychosocial treatments: practical strategies. In: Onken LS, Blaine J, Boren J, editors. Beyond the Therapeutic Alliance: Keeping the Drug-Dependent Individual in Treatment. Bethesda: National Institutes of Health; 1997.
142. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative alternative for developing nursing knowledge. Res Nurs Health 1997 Apr;20(2):169-177. [Medline: 9100747]
143. Heyward V. Advanced Fitness Assessment and Exercise Prescription. Champaign: Human Kinetics; 2002.
144. Heyward V. Advanced Fitness Assessment and Exercise Prescription. 6th Edition. Champaign: Human Kinetics; 2010.
145. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008 Feb;61(2):102-109. [doi: 10.1016/j.jclinepi.2007.03.012] [Medline: 18177782]
146. Liang MH. Evaluating measurement responsiveness. J Rheumatol 1995 Jun;22(6):1191-1192. [Medline: 7674254]
147. Godin G. Godin Leisure-Time Exercise Questionnaire. Med Sci Sport Exer 1997;29(June Supplement):S36-S38.
148. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol 1993 Mar;11(3):570-579. [Medline: 8445433]
149. Spielberger CD. Manual for the State-Trait Anxiety Inventory (Form Y). Palo Alto: Consulting Psychologists Press; 1983.

150. Cella D. The Functional Assessment of Cancer Therapy-Anemia (FACT-An) Scale: a new tool for the assessment of outcomes in cancer anemia and fatigue. Semin Hematol 1997 Jul;34(3 Suppl 2):13-19. [Medline: 9253779]
151. Northouse LL. Mastectomy patients and the fear of cancer recurrence. Cancer Nurs 1981 Jun;4(3):213-220. [Medline: 6909039]
152. Estabrooks PA, Carron AV. The physical activity group environment questionnaire: an instrument for the assessment of cohesion in exercise classes. Group Dynamics 2000;4(3):1089-2699.
153. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singapore 1994 Mar;23(2):129-138. [Medline: 8080219]
154. 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]
155. Boeije H. A purposeful approach to the constant comparative method in the analysis of qualitative interviews. Qual Quant 2002;36(4):391-409.
156. Thompson B. What future quantitative social science research could look like: Confidence intervals for effect sizes. Educational Researcher 2002;31(3):25-32.
157. Hedges L, Olkin I. Statistical Methods for Meta-Analysis. San Diego: Academic Press; 1985.
158. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j.2002.384.doc.x] [Medline: 15189396]
159. Kavanagh T, Mertens DJ, Hamm LF, Beyene J, Kennedy J, Corey P, et al. Prediction of long-term prognosis in 12169 men referred for cardiac rehabilitation. Circulation 2002 Aug 06;106(6):666-671 [FREE Full text] [Medline: 12163425]
160. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]

Abbreviations

ACSM: American College of Sports Medicine
AHT: adjuvant hormone therapy
CBT: cognitive behavioral therapy
CEP: certified exercise physiologist
ELLICSR: Electronic Living Laboratory for Interdisciplinary Cancer Survivorship Research
MCD: minimally important clinical difference
MI: motivational interviewing
MVPA: moderate-to-vigorous physical activity
PMCC: Princess Margaret Cancer Centre
RCT: randomized controlled trial
RKin: registered kinesiologist

Edited by G Eysenbach; submitted 04.09.16; peer-reviewed by D Pekmezi, J Cho, M Cox, D Ehlers; comments to author 11.10.16; revised version received 22.12.16; accepted 25.03.17; published 24.08.17.

Please cite as:
Ritvo P, Obadia M, Santa Mina D, Alibhai S, Sabiston C, Oh P, Campbell K, McCready D, Auger L, Jones JM
Smartphone-Enabled Health Coaching Intervention (iMOVE) to Promote Long-Term Maintenance of Physical Activity in Breast Cancer Survivors: Protocol for a Feasibility Pilot Randomized Controlled Trial
JMIR Res Protoc 2017;6(8):e165
URL: http://www.researchprotocols.org/2017/8/e165/
doi:10.2196/resprot.6615
PMID:28838886

Protocol

Simulation-Based Training of Non-Technical Skills in Colonoscopy: Protocol for a Randomized Controlled Trial

Rishad Khan[1], BSc; Michael A Scaffidi[1], BSc, MEd; Catharine M Walsh[2,3], MEd, MD, PhD; Peter Lin[1]; Ahmed Al-Mazroui[1], MD; Barinder Chana[1], MD; Ruben Kalaichandran[1], BSc; Woojin Lee[1], BSc; Teodor P Grantcharov[4], MD, PhD; Samir C Grover[1], MEd, MD

[1]St. Michael's Hospital, Division of Gastroenterology, University of Toronto, Toronto, ON, Canada

[2]Hospital for Sick Children, Division of Gastroenterology, Hepatology, and Nutrition, Learning Institute, and Research Institute, University of Toronto, Toronto, ON, Canada

[3]The Wilson Centre, University of Toronto, Toronto, ON, Canada

[4]St. Michael's Hospital, Department of General Surgery, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Samir C Grover, MEd, MD
St. Michael's Hospital
Division of Gastroenterology
University of Toronto
30 Bond Street
Toronto, ON, M5B 1W8
Canada
Phone: 1 4168646060
Fax: 1 4168646060
Email: samir.grover@utoronto.ca

Abstract

Background: Non-technical skills (NTS), such as communication and professionalism, contribute to the safe and effective completion of procedures. NTS training has previously been shown to improve surgical performance. Moreover, increases in NTS have been associated with improved clinical endoscopic performance. Despite this evidence, NTS training has not been tested as an intervention in endoscopy.

Objective: The aim of this study is to evaluate the effectiveness of a simulation-based training (SBT) curriculum of NTS on novice endoscopists' performance of clinical colonoscopy.

Methods: Novice endoscopists were randomized to 2 groups. The control group received 4 hours of interactive didactic sessions on colonoscopy theory and 6 hours of SBT. Hours 5 and 6 of the SBT were integrated scenarios, wherein participants interacted with a standardized patient and nurse, while performing a colonoscopy on the virtual reality (VR) simulator. The NTS (intervention) group received the same teaching sessions but the last hour was focused on NTS teaching. The NTS group also reviewed a checklist of tasks relevant to NTS concepts prior to each integrated scenario case and was provided with dedicated feedback on their NTS performance during the integrated scenario practice. All participants were assessed at baseline, immediately after training, and 4 to 6 weeks post-training. The primary outcome measure is colonoscopy-specific performance in the clinical setting.

Results: In total, 42 novice endoscopists completed the study. Data collection and analysis is ongoing. We anticipate completion of all assessments by August 2017. Data analysis, manuscript writing, and subsequent submission for publication is expected to be completed by December 2017.

Conclusions: Results from this study may inform the implementation of NTS training into postgraduate gastrointestinal curricula. NTS curricula may improve attitudes towards patient safety and self-reflection among trainees. Moreover, enhanced NTS may lead to superior clinical performance and outcomes in colonoscopy.

Trial Registration: Clinicaltrial.gov NCT02877420; https://www.clinicaltrials.gov/ct2/show/NCT02877420 (Archived by WebCite at http://www.webcitation.org/6rw94ubXX NCT02877420)

(JMIR Res Protoc 2017;6(8):e153) doi:10.2196/resprot.7690

KEYWORDS

simulation; colonoscopy; non-technical skills; skill acquisition

Introduction

Simulation-based training (SBT) in gastrointestinal (GI) endoscopy improves clinical performance among trainees [1-3]. However, the components of optimal simulation curricula are unclear. In procedural settings like surgery, the teaching of non-technical skills (NTS) has been shown to improve novice surgeons' performance [4]. Despite this evidence, the impact of NTS training on endoscopy performance has not been explored.

NTS, such as teamwork, communication, situational awareness, and decision making, are important factors in healthcare regarding adverse health outcomes. A recent systematic review of critical incidents in intensive care units found that failures in non-technical domains contributed to a large proportion of medical errors [5]. Another systematic review found that deficiencies in NTS were associated with decreased technical skill in surgical settings [6].

Given the importance of these skills, several interventions have been proposed to enhance NTS among physicians. First, didactic training has been shown to improve attitudes and awareness of NTS. In a study among surgical residents, a curriculum featuring didactic teaching of NTS led to improved non-technical specific performance in the operating room, compared to conventional training (ie, daily activities on surgical wards, call schedules, and designated operating room time) [4]. Second, checklists have been shown to improve NTS, such as promoting adherence to procedural protocols, especially during surgical crises [7]. Finally, debriefing and feedback by expert instructors can allow trainees to acquire and consolidate relevant NTS. Among residents, both oral and videotape-assisted feedback can yield superior NTS acquisition compared to no feedback [8].

In endoscopy, improvement of NTS in the simulated setting is associated with superior colonoscopy skills [3]. However, the direct impact of NTS-specific training on clinical performance is unknown. This study aims to evaluate the impact of NTS training on clinical colonoscopy performance among novice endoscopists.

Methods

This single-blind, parallel group, randomized controlled trial (RCT) is being conducted at a tertiary-care academic center. Research ethics approval was granted by the St. Michael's Hospital Research Ethics Board (15-164). Recruitment for the study is complete. All testing and training took place at St. Michael's Hospital (30 Bond Street, Toronto, Ontario, M5B 1W8). Informed written consent was acquired from all participants and patients involved in the study. The study design is summarized in Figure 1.

Participants

A total of 42 postgraduate trainees enrolled in general surgery, adult gastroenterology, and internal medicine programs at the University of Toronto were recruited through purposive sampling. Participants were identified from a list of trainees rotating through the gastroenterology service at St. Michael's Hospital and emailed with recruitment details. Study enrollment took place from June 2015 to June 2016. Participants were excluded if they performed 25 or more real or simulated endoscopic procedures at the time of their participation in the study.

Simulation Devices

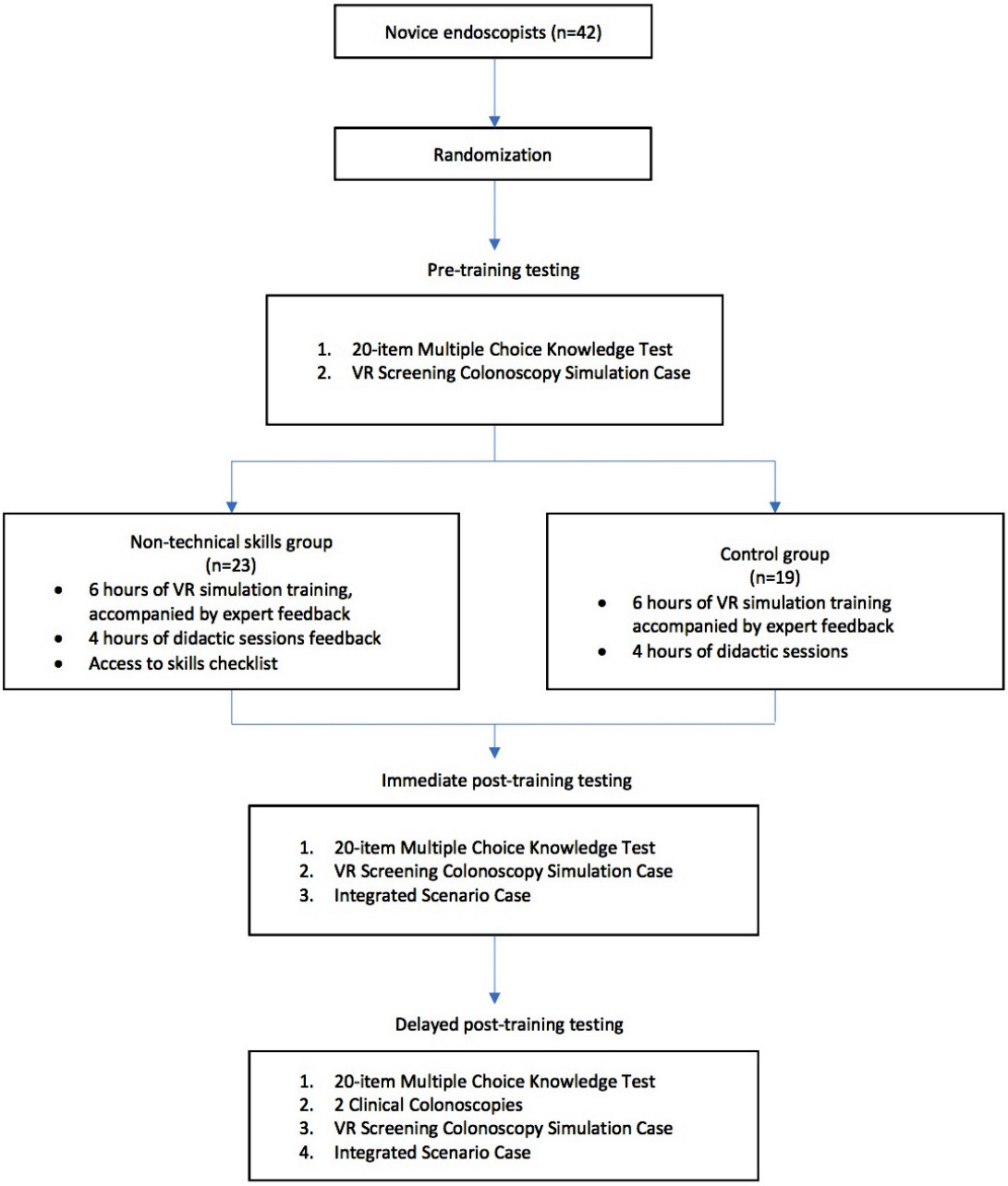
Bench Top, Low-Fidelity Simulator

The low-fidelity simulator is a validated bench-top endoscopy simulator that helps develop general endoscopic skills [9]. The simulator is comprised of a series of vertical wooden barriers with numbered holes conforming to 27 different sequences of varying complexity. An Olympus PCF-180 pediatric videocolonoscope (Olympus Canada) is used to navigate the defined sequences as quickly and accurately as possible, with visual output being displayed on a video monitor.

Virtual Reality, High-Fidelity Simulator

The high-fidelity simulator is the EndoVR endoscopy simulator (CAE Healthcare Canada). It models navigation through a colon, using a specialized endoscope that is inserted into a computer-based module with a screen showing the colonic lumen of a virtual patient. It provides both visual and haptic feedback related to the procedure. The VR simulator has several standardized case-based scenarios of varying complexity for colonoscopy.

Figure 1. Study design.



Experimental Design

Baseline Questionnaire

A written questionnaire is administered to all participants at the start of the training to collect demographic and background information including age, sex, level of training, previous endoscopy experience, and nature of experience (if applicable). Previous experience with team sports and video games is also assessed, as these activities correlate with NTS and baseline endoscopic skill, respectively (Multimedia Appendix 1) [10].

Pre-Test

Participants take part in a pre-test designed to assess their baseline (1) knowledge of colonoscopy (knowledge test); (2) technical skills (VR simulation test); and (3) NTS (VR simulation-based “integrated scenario” test). No feedback is provided at any point during the pre-test.

The knowledge test is a 30-minute test containing 20 multiple choice questions (MCQ) designed to assess participants’ theoretical knowledge of colonoscopy, including indications, sedation, safety, findings, pathology, and follow-up (Multimedia Appendix 2).

Participants' baseline endoscopic technical proficiency is assessed through the completion of a colonoscopy procedure on the VR simulator (EndoVR Colonoscopy Module 3). This scenario simulates a screening colonoscopy, without the need for any type of intervention, such as biopsy. There is a time limit of 30 minutes per procedure. An expert rater assesses performance, but does not provide assistance. Participants are videotaped to obtain performance measures with their faces

hidden to ensure anonymity. Prior to starting the procedure, participants complete a questionnaire to measure their self-efficacy.

Following the simulator-only test, participants complete an integrated scenario format test to assess their baseline endoscopic non-technical proficiency. The integrated scenario requires participants to perform a colonoscopy procedure on the VR simulator while interacting with an endoscopic nurse and a standardized patient (SP) [11]. The simulated procedure mimics the setup of an endoscopic suite with the VR simulator positioned next to a patient table. An SP, who receives instructions regarding their medical role, acts out a scenario on colon cancer screening. Trainees are expected to explain the colonoscopy procedure, its benefits and risks, and obtain procedural consent. Trainees then carry out the procedure on the VR simulator (EndoVR Polypectomy Module #3) while responding to the patient and interacting with the standardized nurse (SN), as appropriate. The SP acts out cues from the VR simulator when the simulator signals that the procedure has exceeded its threshold for discomfort. The performances of all participants are videotaped in a similar manner to the VR simulation test to obtain performance measures. Participants are given a maximum of 45 minutes to complete the procedure. Prior to starting the procedure, participants complete a questionnaire to measure their self-efficacy.

Training Intervention

Participants are subsequently randomized using a randomization algorithm to 1 of 2 groups (control and intervention), following a 1:1 allocation distribution.

The control group receives 4 hours of interactive, small-group, didactic and hands-on sessions on the theory of colonoscopy, led by an expert academic gastroenterologist. The core curriculum was designed based on the American Society for Gastrointestinal Endoscopy (ASGE) colonoscopy curriculum and a detailed endoscopic training textbook and includes teaching on anatomy, pathophysiology, indications, risks and benefits of the procedures, training on specific elements of performance of colonoscopy procedures (eg, one-handed steering), and strategies for loop reduction, terminal ileal cannulation, and adequate visualization of mucosa [12,13]. This curriculum has been shown to be effective when compared to self-regulated learning on the simulator [3]. The sessions are interlaced with 6 hours of expert-assisted instruction on both the low-fidelity simulator (1 hour) and on the high-fidelity VR simulator (5 hours). Six modules of increasing difficulty in colonoscopy and colonoscopy polypectomy are taught with one-on-one feedback from an expert academic endoscopist. The endoscopy instructor demonstrates techniques, answers questions, and provides individualized performance feedback on global performance. The final 2 hours spent on the high-fidelity scenario use the integrated scenario, which features an SP and SN.

The intervention (NTS) group receives the same first 3 hours of interactive, small-group, didactic and hands-on sessions as the control group. The 4th hour of didactic sessions focuses on NTS, which includes a discussion of the major areas of NTS. Participants also watch a video demonstrating an ideal endoscopic procedure (ie, benchmark video) in terms of technical and NTS. These sessions and the video introduce trainees to the Endoscopic Non-Technical Skills (E-NTS) Checklist, which is provided for them to use during the integrated scenario training (Multimedia Appendix 3). This checklist was developed per evidence-based recommendations and targets NTS in endoscopy [14]. The NTS group also receives 6 hours of expert-assisted instruction on both the low-fidelity simulator (1 hour) and on the high-fidelity VR simulator (5 hours). Six modules of increasing difficulty in colonoscopy and colonoscopy polypectomy are taught with one-on-one feedback from an expert academic endoscopist. The endoscopy instructor demonstrates techniques, answers questions, and provides individualized performance feedback on global performance with a focus on NTS. Similar to the control group, the final 2 hours spent on the high-fidelity scenario use the integrated scenario. Terminal feedback dedicated to NTS is given after each integrated scenario by the instructor. Finally, participants in this group have access to the E-NTS Checklist during training in the integrated scenario, as participants can view the checklist prior to and after each case.

Post-Test

A post-test is administered after completion of the training period to compare skill and knowledge acquisition between the control and intervention groups. The immediate post-test is designed to evaluate trainees' (1) knowledge acquisition (knowledge test); (2) technical skills acquisition (VR simulation test); and (3) NTS acquisition (VR simulation-based integrated scenario test). No feedback is provided during the post-test.

Participants' knowledge acquisition is evaluated using the same MCQ test conducted at baseline. Again, trainees have 30 minutes to complete the 20 questions.

Participants' technical skill acquisition is assessed through the completion of the same colonoscopy procedure on the VR simulator (EndoVR Colonoscopy Module 3). The time limit is 30 minutes. Prior to starting the procedure, participants complete a questionnaire to measure their self-efficacy.

Participants' NTS acquisition is assessed using the integrated scenario procedure on the VR simulator (EndoVR Polypectomy Module #3), while also responding to the patient and interacting with the nurse, as appropriate. The time limit is 45 minutes. Prior to starting the procedure, participants complete a questionnaire to measure their self-efficacy.

Delayed Testing

A delayed-test is administered 4 to 6 weeks after completion of the training period to compare retention and transfer of skill between the control and intervention groups. It consists of (1) a knowledge test; (2) a VR simulation test; (3) a VR simulation-based integrated scenario test; and (4) a patient-based transfer test.

Participants' knowledge acquisition is evaluated using the same MCQ test conducted at baseline and during the post-test. Trainees have 30 minutes to complete 20 questions.

Participants' technical skill acquisition is assessed through the completion of the same colonoscopy procedure on the VR

simulator (EndoVR Colonoscopy Module 3). The time limit is 30 minutes. Prior to starting the procedure, participants complete a questionnaire to measure their self-efficacy.

Participants' NTS acquisition is assessed using the integrated scenario procedure on the VR simulator (EndoVR Polypectomy Module #3), while also responding to the patient and interacting with the nurse, as appropriate. The time limit is 45 minutes. Prior to starting the procedure, participants complete a questionnaire to measure their self-efficacy.

Participants' transfer of skills to the clinical setting is assessed using live colonoscopies. Each participant completes 2 colonoscopies on real patients 4 to 6 weeks after the training period. These procedures are videotaped in a manner that anonymizes the participant and patient. Procedures on patients with a history of colonic or pelvic surgery or difficult colonoscopy are excluded. Sedation and monitoring are carried out per standard practice at the endoscopy unit. An experienced attending endoscopist (greater than 1000 completed procedures) provides verbal and/or hands-on assistance as necessary and takes over if the participant cannot complete the procedure or if any concerns regarding patient safety arise.

Outcome Measures

The primary outcome measure is the difference in performance between the control and intervention groups during participants' 2 clinical colonoscopies. Each videotaped clinical colonoscopy will be independently assessed by 2 experienced endoscopists using the Joint Advisory Group for GI Endoscopy Direction Observation of Procedural Skills (JAG DOPS) (Multimedia Appendix 4) [15]. The raters will be blinded to the group assignment. Training on how to use the tool will be provided for raters by the investigators of the study.

Secondary outcome measures include the differences between the control and intervention groups with respect to (1) procedural knowledge, as assessed by the knowledge MCQ tests; (2) NTS performance during the clinical colonoscopies, as assessed by the Modified Objective Structured Assessment of Non-Technical Skills (M-OSANTS) for colonoscopy, which has been previously validated for surgery and modified for endoscopy [4] (Multimedia Appendix 5); (3) clinical performance on clinical colonoscopies, as assessed by the Gastrointestinal Endoscopy Competency Assessment Tool (GiECAT) [16] (Multimedia Appendix 6); (4) technical performance on a VR simulated colonoscopy after training and 4 to 6 weeks after training (immediate and delayed post-training assessments, respectively), as assessed by the JAG DOPS and the GiECAT; (5) technical and non-technical performance during an integrated scenario format test 4 to 6 weeks after training, as assessed by the JAG DOPS, GiECAT, and M-OSANTS; (6) patient comfort during the clinical colonoscopies, as assessed by the Nurse-Assessed Patient Comfort Score (NAPCOMS) [17] (Multimedia Appendix 7); (7) participant self-efficacy, as measured by an adapted scale based on the General Self-Efficacy Scale (GSE) [18,19] (Multimedia Appendix 8); and (8) global performance and communication skills during integrated scenarios as assessed by the Integrated Scenario Global Rating Form (ISGRF) and Integrated Scenario Communication Rating Form (ISCRF) [11,20,21], respectively (Multimedia Appendices 9 and).

Experienced endoscopists will assess participants' colonoscopy-specific skills, technical skills, and NTS during the pre-training, immediate, and delayed post-training simulation-based assessments.

Analysis Plan

Statistical analyses will be performed using SPSS version 20 (SPSS, Inc.). All statistical tests will be considered significant at P less than .05.

Baseline Questionnaire

Patient demographics and baseline variables will be characterized with descriptive statistics, using mean with standard deviation and number with frequency for continuous and categorical variables, respectively.

Clinical Performance

Clinical performance during the live colonoscopies for each group will be determined by comparing the scores from the DOPS, GiECAT, NAPCOMS, and Modified-OSANTS. Specifically, a mixed factor 2 (control curriculum versus intervention curriculum) times 2 (procedure 1 versus procedure 2) analysis of variance (ANOVA) will be used to determine whether there is a difference based on the rating scales. The Tukey honest significant difference (HSD) test will be used as a post-hoc analysis to determine any significant differences.

Technical Performance

Technical performance on the simulator for each group will be determined by comparing the scores from the DOPS and GiECAT. Specifically, a mixed factor 2 (control curriculum versus intervention curriculum) times 3 (pre-test, post-test, delayed test) ANOVA will be used to determine whether there is a difference based on the rating scales. The Tukey HSD test will be used as a post-hoc analysis to determine any significant differences.

Non-Technical Performance

Non-technical performance on the simulator for each group will be determined by comparing the scores from the M-OSANTS, ISGRF, and ISCRF. Specifically, a mixed factor 2 (control curriculum versus intervention curriculum) times 3 (pre-test, post-test, delayed test) ANOVA will be used to determine whether there is a difference based on the rating scales. The Tukey HSD test will be used as a post-hoc analysis to determine any significant differences.

Self-Efficacy

Self-efficacy during the simulated setting for each group will be determined by comparing the scores from the GSE. Specifically, a mixed factor 2 (control curriculum versus intervention curriculum) times 3 (pre-test, post-test, delayed test) ANOVA will be used to determine whether there is a difference based on the rating scales. The Tukey HSD test will be used as a post-hoc analysis to determine any significant differences.

Sample Size Estimation

A power analysis was computed using G*Power version 3.1.9 [22]. Using a previous study that evaluated an NTS training curriculum in surgery, we conducted the analysis using the relevant effect size [4]. Based on an effect size of f=0.65, an alpha of .05 (2-tailed), a beta of .20, 2 groups, and 3 measurements, 16 participants are required to achieve a power of greater than 0.80 using repeated measures ANOVA (between-factors). Furthermore, a previous study comparing a curriculum in endoscopic simulation found that a minimum of 15 participants per group was sufficient to detect a significant difference [3]. To accommodate for a projected 20% dropout and/or non-response, we recruited a total of 42 participants.

Results

A total of 42 participants were recruited, randomized, and completed the study. No participants were lost to follow-up. Through a 1:1 allocation distribution, 19 individuals were randomized into the control group and 23 individuals were randomized into the NTS group. Their demographic information and endoscopic experiences are summarized in Table 1. Two experienced endoscopists were recruited to assess participants' videotaped performances and we anticipate completion of all assessments by August 2017. Data analysis, manuscript writing, and subsequent submission for publication is expected to be completed by December 2017.

Table 1. Baseline demographic characteristics and endoscopy experience of participants (N=42).

Variable		Control group (n=19)	Intervention (NTS[a]) group (n=23)
Age in years, mean (SD)		29.5 (1.8)	28.1 (2.1)
Gender, n (%)			
	Male	11 (58)	16 (70)
	Female	8 (42)	7 (30)
Training program, n (%)			
	Gastroenterology	4 (21)	6 (26)
	General surgery	13 (68)	16 (70)
	Internal medicine	2 (11)	1 (4)
Previous endoscopic experience, mean (SD)			
	Number of independent colonoscopies	0.11 (0.46)	0.30 (0.82)
	Number of assisted colonoscopies	2.63 (4.66)	1.70 (2.50)

[a]NTS: non-technical skills.

Discussion

Principal Findings

The practice of effective NTS is critical in procedural medicine. In addition, deficiencies in NTS are associated with adverse patient outcomes [5,6]. However, there is evidence that NTS training can improve surgical performance in the simulated setting, lead to better attitudes towards patient safety, and promote more self-reflection among trainees [4,23,24].

Previous studies have explored NTS in GI endoscopy. Matharoo et al shared their perspectives on implementing an endoscopy safety checklist to decrease adverse health outcomes [25]. This paper reported on the implications, logistics, and uptake of a safety checklist, and did not present any outcome data. A 2014 report implementing an NTS curriculum found that it improved patient safety knowledge and attitudes among multi-disciplinary endoscopy teams [24]. In another study, an RCT comparing 2 curricula of SBT in endoscopy found that improved NTS in the simulated setting were associated with superior clinical performance [3]. However, that study did not test NTS training as an intervention.

Conclusion

This study aims to evaluate the direct impact of NTS training on clinical colonoscopy performance. Results can inform the potential implementation of NTS into postgraduate gastrointestinal curricula.

Acknowledgments

We would like to thank Roger Chow for his invaluable assistance.

Authors' Contributions

RK, MAS, CMW, BC, TPG, and SCG designed the initial version of the protocol. All authors read and approved the final protocol before submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Baseline demographic and endoscopic experience.

[PDF File (Adobe PDF File), 178KB - resprot_v6i8e153_app1.pdf]

Multimedia Appendix 2

Knowledge test questions.

[PDF File (Adobe PDF File), 769KB - resprot_v6i8e153_app2.pdf]

Multimedia Appendix 3

Endoscopic non-technical skills (E-NTS) checklist.

[PDF File (Adobe PDF File), 89KB - resprot_v6i8e153_app3.pdf]

Multimedia Appendix 4

Joint Advisory Group on GI Endoscopy Direct Observation of Procedural Skills (JAG DOPS) form.

[PDF File (Adobe PDF File), 276KB - resprot_v6i8e153_app4.pdf]

Multimedia Appendix 5

Modified Objective Structured Assessment of Non-Technical Skills (M-OSANTS) form.

[PDF File (Adobe PDF File), 357KB - resprot_v6i8e153_app5.pdf]

Multimedia Appendix 6

Gastrointestinal Endoscopy Competency Assessment Tool (GiECAT) form.

[PDF File (Adobe PDF File), 322KB - resprot_v6i8e153_app6.pdf]

Multimedia Appendix 7

Nurse Assessed Patient Comfort Score (NAPCOMS) form.

[PDF File (Adobe PDF File), 89KB - resprot_v6i8e153_app7.pdf]

Multimedia Appendix 8

General self-efficacy assessment.

[PDF File (Adobe PDF File), 144KB - resprot_v6i8e153_app8.pdf]

Multimedia Appendix 9

Integrated Scenario Global Rating Form (ISGRF).

[PDF File (Adobe PDF File), 347KB - resprot_v6i8e153_app9.pdf]

Multimedia Appendix 10

Integrated Scenario Communication Rating Form (ISCRF).

[PDF File (Adobe PDF File), 271KB - resprot_v6i8e153_app10.pdf]

References

1. Walsh CM, Sherlock ME, Ling SC, Carnahan H. Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy. Cochrane Database Syst Rev 2012 Jun 13(6):CD008237. [doi: 10.1002/14651858.CD008237.pub2] [Medline: 22696375]
2. Singh S, Sedlack RE, Cook DA. Effects of simulation-based training in gastrointestinal endoscopy: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2014 Oct;12(10):1611-23.e4. [doi: 10.1016/j.cgh.2014.01.037] [Medline: 24509241]
3. Grover SC, Garg A, Scaffidi MA, Yu JJ, Plener IS, Yong E, et al. Impact of a simulation training curriculum on technical and nontechnical skills in colonoscopy: a randomized trial. Gastrointest Endosc 2015 Dec;82(6):1072-1079. [doi: 10.1016/j.gie.2015.04.008] [Medline: 26007221]
4. Dedy NJ, Bonrath EM, Ahmed N, Grantcharov TP. Structured training to improve nontechnical performance of junior surgical residents in the operating room: A randomized controlled trial. Ann Surg 2016 Jan;263(1):43-49. [doi: 10.1097/SLA.0000000000001186] [Medline: 25775073]
5. Reader T, Flin R, Lauche K, Cuthbertson BH. Non-technical skills in the intensive care unit. Br J Anaesth 2006 May;96(5):551-559. [doi: 10.1093/bja/ael067] [Medline: 16567346]
6. Hull L, Arora S, Aggarwal R, Darzi A, Vincent C, Sevdalis N. The impact of nontechnical skills on technical performance in surgery: a systematic review. J Am Coll Surg 2012 Feb;214(2):214-230. [doi: 10.1016/j.jamcollsurg.2011.10.016] [Medline: 22200377]
7. Arriaga AF, Bader AM, Wong JM, Lipsitz SR, Berry WR, Ziewacz JE, et al. Simulation-based trial of surgical-crisis checklists. N Engl J Med 2013 Jan 17;368(3):246-253. [doi: 10.1056/NEJMsa1204720] [Medline: 23323901]
8. Savoldelli GL, Naik VN, Park J, Joo HS, Chow R, Hamstra SJ. Value of debriefing during simulated crisis management: oral versus video-assisted oral feedback. Anesthesiology 2006 Aug;105(2):279-285. [Medline: 16871061]
9. Walsh CM, Cooper M, Rabeneck L, Carnahan H. High versus low fidelity simulation training in gastroenterology: expertise discrimination. In: Can J Gastroenterol A.: Canadian Journal of Gastroenterology and Hepatology; 2008 Presented at: Canadian Digestive Diseases Week; 2008 Feb 29-Mar 03; Montreal.
10. van Dongen KW, Verleisdonk EJ, Schijven MP, Broeders IA. Will the Playstation generation become better endoscopic surgeons? Surg Endosc 2011 Jul;25(7):2275-2280 [FREE Full text] [doi: 10.1007/s00464-010-1548-2] [Medline: 21416186]
11. Kneebone R, Nestel D, Yadollahi F, Brown R, Nolan C, Durack J, et al. Assessing procedural skills in context: exploring the feasibility of an Integrated Procedural Performance Instrument (IPPI). Med Educ 2006 Nov;40(11):1105-1114. [doi: 10.1111/j.1365-2929.2006.02612.x] [Medline: 17054620]
12. Training Committee 2010-2011, Sedlack RE, Shami VM, Adler DG, Coyle WJ, DeGregorio B, et al. Colonoscopy core curriculum. Gastrointest Endosc 2012 Sep;76(3):482-490. [doi: 10.1016/j.gie.2012.04.438] [Medline: 22898404]
13. Cotton PB, Williams CB. Practical Gastrointestinal Endoscopy: The Fundamentals, 5th Edition. Hoboken, NJ: Wiley-Blackwell; 2008.
14. Hales B, Terblanche M, Fowler R, Sibbald W. Development of medical checklists for improved quality of patient care. Int J Qual Health Care 2008 Feb;20(1):22-30. [doi: 10.1093/intqhc/mzm062] [Medline: 18073269]
15. Barton JR, Corbett S, van DVCP, English Bowel Cancer Screening Programme, UK Joint Advisory Group for Gastrointestinal Endoscopy. The validity and reliability of a direct observation of procedural skills assessment tool: assessing colonoscopic skills of senior endoscopists. Gastrointest Endosc 2012 Mar;75(3):591-597. [doi: 10.1016/j.gie.2011.09.053] [Medline: 22227035]
16. Walsh CM, Ling SC, Khanna N, Cooper MA, Grover SC, May G, et al. Gastrointestinal endoscopy competency assessment tool: development of a procedure-specific assessment tool for colonoscopy. Gastrointest Endosc 2014 May;79(5):798-807.e5. [doi: 10.1016/j.gie.2013.10.035] [Medline: 24321390]
17. Rostom A, Ross ED, Dubé C, Rutter MD, Lee T, Valori R, et al. Development and validation of a nurse-assessed patient comfort score for colonoscopy. Gastrointest Endosc 2013 Feb;77(2):255-261. [doi: 10.1016/j.gie.2012.10.003] [Medline: 23317691]
18. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Measures in Health Psychology: A User's Portfolio. Windsor: NFER-Nelson; 1995:35-37.
19. Pena G, Altree M, Field J, Thomas MJ, Hewett P, Babidge W, et al. Surgeons' and trainees' perceived self-efficacy in operating theatre non-technical skills. Br J Surg 2015 May;102(6):708-715. [doi: 10.1002/bjs.9787] [Medline: 25790065]
20. LeBlanc VR, Tabak D, Kneebone R, Nestel D, MacRae H, Moulton C. Psychometric properties of an integrated assessment of technical and communication skills. Am J Surg 2009 Jan;197(1):96-101. [doi: 10.1016/j.amjsurg.2008.08.011] [Medline: 19101250]
21. Hodges B, McIlroy JH. Analytic global OSCE ratings are sensitive to level of training. Med Educ 2003 Nov;37(11):1012-1016. [Medline: 14629415]
22. 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. [Medline: 17695343]
23. Flin R, Yule S, Paterson-Brown S, Maran N, Rowley D, Youngson G. Teaching surgeons about non-technical skills. Surgeon 2007 Apr;5(2):86-89. [Medline: 17450689]

24. Matharoo M, Haycock A, Sevdalis N, Thomas-Gibson S. Endoscopic non-technical skills team training: the next step in quality assurance of endoscopy training. World J Gastroenterol 2014 Dec 14;20(46):17507-17515 [FREE Full text] [doi: 10.3748/wjg.v20.i46.17507] [Medline: 25516665]
25. Matharoo M, Sevdalis N, Thillai M, Bouri S, Marjot T, Haycock A, et al. The endoscopy safety checklist: a longitudinal study of factors affecting compliance in a tertiary referral centre within the United Kingdom. BMJ Qual Improv Rep 2015;4(1) [FREE Full text] [doi: 10.1136/bmjquality.u206344.w2567] [Medline: 26734331]

Abbreviations

ANOVA: analysis of variance
E-NTS: endoscopic non-technical skills
GI: gastrointestinal
GiECAT: Gastrointestinal Endoscopy Competency Assessment Tool
GSE: General Self-Efficacy Scale
HSD: honest significant difference
ISCRF: Integrated Scenario Communication Rating Form
ISGRF: Integrated Scenario Global Rating Form (ISGRF)
JAG DOPS: Joint Advisory Group for GI Endoscopy Direction Observation of Procedural Skills
M-OSANTS: Modified Objective Structured Assessment of Nontechnical Skills
MCQ: multiple choice questions
NAPCOMS: Nurse-Assessed Patient Comfort Score
NTS: non-technical skills
RCT: randomized controlled trial
SBT: simulation-based training
SN: standardized nurse
SP: standardized patient
VR: virtual reality

Edited by G Eysenbach; submitted 13.03.17; peer-reviewed by D Bergeron; accepted 30.05.17; published 04.08.17.

Please cite as:
Khan R, Scaffidi MA, Walsh CM, Lin P, Al-Mazroui A, Chana B, Kalaichandran R, Lee W, Grantcharov TP, Grover SC
Simulation-Based Training of Non-Technical Skills in Colonoscopy: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2017;6(8):e153
URL: http://www.researchprotocols.org/2017/8/e153/
doi:10.2196/resprot.7690
PMID:28778849

Protocol

Are Routine Radiographs Needed the Day After Open Reduction and Internal Fixation Surgery for Distal Radius and Ankle Fractures: Study Protocol for a Prospective, Open Label, Randomized Controlled Trial

Florian Oehme[1*], MD; Annika Rühle[1*], MD; Julia Mühlhäusser[1], MD; Lana Fourie[1], MD; Björn-Christian Link[2], MD; Reto Babst[2], Prof. Dr. med.; Frank JP Beeres[2], MD PhD

[1]Lucerne Cantonal Hospital, Surgery Department, Lucerne, Switzerland

[2]Lucerne Cantonal Hospital, Orthopaedic and Trauma Surgery, Lucerne, Switzerland

[*]these authors contributed equally

Corresponding Author:
Florian Oehme, MD
Lucerne Cantonal Hospital
Surgery Department
Spitalstrasse
Lucerne, 6003
Switzerland
Phone: 41 41 205 1577
Fax: 41 41 205 2484
Email: florianoehme85@gmail.com

Abstract

Background: Distal radius and ankle fractures are one of the most common operatively treated fractures. To date, there is no consensus concerning the need for a standard postoperative radiograph. This leads to undesirable practice variations. A standardized radiograph in the department of radiology would theoretically be more reproducible and operator independent than an intraoperatively obtained fluoroscopic image. However, if adequate intraoperative radiographs have been obtained, it is questionable if these postoperative radiographs are necessary and will lead to changes in the treatment strategy. If standard postoperative radiographs are no longer required, this would lead to a reduction in radiation exposure and health care costs. The hypothesis is that routine standardized postoperative radiographs do not influence the quality of care for patients operated on for either a distal radius or an ankle fracture if adequate intraoperative standardized radiographs have been obtained.

Objective: The primary aim of this study is to evaluate if there is a need for routine postoperative radiographs after an osteosynthesis of a distal radius or ankle fracture.

Methods: In a prospective, randomized controlled, open label trial based on a noninferiority design, we will enroll 332 patients. Patients will be randomized either in the control or the intervention group. The control group will be treated according to our current, standard protocol in which all patients receive a standard anterior-posterior and lateral radiograph on the first postoperative day. Patients randomized to the intervention group will be treated without a standard postoperative radiograph. All patients (N=332) will have a routine clinical and radiographic control after 6 weeks in the outpatient clinic. Primary outcome is a change in treatment plan, defined as either additional imaging or a reoperation based on the postoperative imaging. Secondary outcome measures include a 36-Item Short Form Survey, Patient-Rated Wrist Hand Evaluation, Foot and Ankle Outcome Score, Visual Analogue Scale, and the range of motion. Those questionnaires will be filled out at the 6-week outpatient control.

Results: The trial was started in August 2016, and 104 patients have been enrolled up to this point.

Conclusions: Our findings will be reported in peer-reviewed publications and may lead to a strong reduction in radiation exposure and health care costs. A preliminary, conservative estimation suggests a yearly cost saving of CHF 1.3 million in Switzerland.

(JMIR Res Protoc 2017;6(8):e159) doi:10.2196/resprot.7698

KEYWORDS

wrist fracture; distal radius fracture; ankle fracture; postoperative radiograph; functional outcome; cost reduction; radiation exposure

Introduction

In a society that is increasingly demanding about the functional outcome after treatment of fractures, the indications for operative treatment have increased over the past decade. Wrist fractures specifically are more often being treated operatively. In addition, due to an increase in the number of patients with obesity, the numbers of patients treated for ankle fractures are also increasing [1,2].

Many hospitals still use routine postoperative radiographs, despite the fact that intraoperative images have been obtained for reduction and implant control. In addition, these images are stored and available in the Picture Archiving and Communication System (PACS). These postoperative radiographs lead to a significant increase in radiation exposure and costs. It remains questionable if these standard postoperative radiographs are justified since the quality of intraoperative C-arm images has improved over the last decade.

Rationale

Intraoperative radiograph documentation, if standardized and adequately performed, has the potential to assess the quality of reduction and fixation [3,4]. Consequently, the additional value of routine postoperative radiographs should be questioned [5].

Given the trend towards cost-effective medicine [6] and keeping cumulative radiation exposure in mind, it is not surprising that these standardized postoperative radiographs are under debate [8]. Despite the above, routine postoperative radiographs are still being performed in many hospitals as a standard of care. Many surgeons and radiologists argue that radiographs obtained in the department of radiology are more standardized and less biased. We therefore evaluated the frequency of changes in treatment plan due to standardized postoperative radiographs in a retrospective trial. Changes in treatment plan were defined as a deviation from the standard postoperative protocol. This included additional imaging or a revision operation.

We found in 7.2% of patients a change in the treatment plan following the evaluation of the standardized postoperative radiograph for patients operated on for distal radius or ankle fractures in 2014. These numbers are high in percentage and suggest that standardized postoperative radiographs do add to the quality of care for the patient.

On the other hand, this retrospective study had several methodological drawbacks, especially as the intraoperative radiographs performed were not standardized. Therefore, we decided to evaluate the need for postoperative radiographs using a prospective randomized control trial. We would expect a significant reduction in radiation exposure and costs of around CHF 1.3 million (estimated for employed patients in Switzerland in 2015 [9]).

Hypothesis

Our hypothesis is that routine standardized postoperative radiographs do not influence the quality of care for patients operated for a distal radius or ankle fracture if adequate intraoperative standardized radiographs have been obtained.

This is the first prospective randomized trial evaluating the additional value of postoperative routine radiographs in operative fracture care of distal radius and ankle fractures.

Human research ethics approval has been obtained from the Ethikkomission Nordwest- und Zentralschweiz (EKNZ). The commission accepted this trial on April 4, 2016 (EKNZ BASEC 2016-00114).

Methods

Based on a noninferiority design, we planned a prospective, randomized controlled, open label trial. Multimedia Appendix 1 shows the study protocol in accordance with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) checklist [10].

Study Population

The study includes patients operated for a distal radius or ankle fracture at a Level 1 trauma center in Switzerland. Patients presenting with these fractures at the emergency department will be eligible for study inclusion if they fulfill all of the following criteria: indication for an operation of a distal radius fracture, according to the Arbeitsgemeinschaft für Osteosynthesefragen (AO) Classification types 23-A-C [4], or an ankle (according to the AO Classification Types 44-A-C) fracture, 18 years of age or older, sufficient understanding and writing of German language, and signed informed consent. The informed consent forms will be stored by the study nurse during the trial.

Patients are excluded if they meet at least one of the following exclusion criteria: not willing/able to sign the informed consent, indication for postoperative computed tomography (CT), pathological fractures, open fractures (>grade I according to Anderson and Gustilo [11]), patients not able to attend the 6 weeks outpatient control, or missed intraoperative standardized radiographs.

Preliminary

Essential pre-study preparations are needed to ensure standardization of intraoperative C-arm handling, limb positioning, and reproducibility of the collected data.

Defining Period

Standard intraoperative radiographs for patients operated for radius fractures are defined as:

1. anterior-posterior and posterior-anterior radiograph of the wrist with a free visualization of the distal radio-ulnar space (defined as no overlap of distal radius and distal ulnar)

2. 20° radial tilted lateral view in which the os pisiform needs to have an overlap with the scaphoid and the ventral cortex of the lunate
3. tangential view [12,13]

Sufficient standard intraoperative radiographs are shown in Figures 1,2, and 3.

Standard intraoperative radiographs for patients operated for ankle fractures are defined as:

1. anterior-posterior view (mortise view) with cleared space between trochlea tali and the tibia/fibula (achieved by inward rotation of 20°)
2. true lateral view with an overlap of the medial and lateral domes of the talus and the fibula forming the posterior one-third of the tibia

Sufficient standard intraoperative radiographs are shown in Figures 4 and 5.

Figure 1. Intraoperative C-arm controlled radiograph (anterior-posterior) in supination with cleared radio-ulnar space.

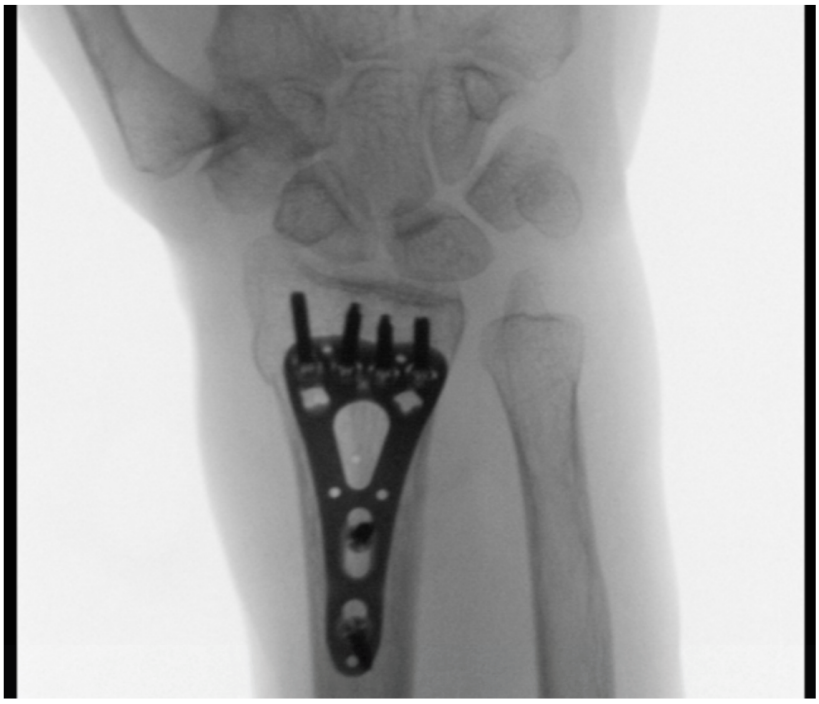


Figure 2. Lateral view with free os pisiforme on the volar level and intra-articular view.

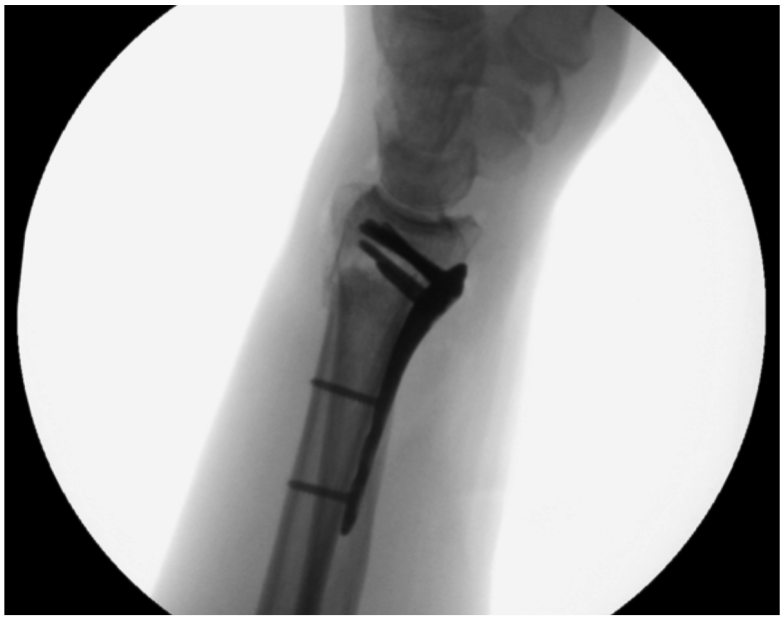


Figure 3. Intraoperative tangential view performed to evaluate the length of the screws.

Figure 4. Intraoperative C-arm control: the ankle is shown in anterior-posterior view, free mortise view is achieved with 90° dorsal extension and 20° rotation inside.

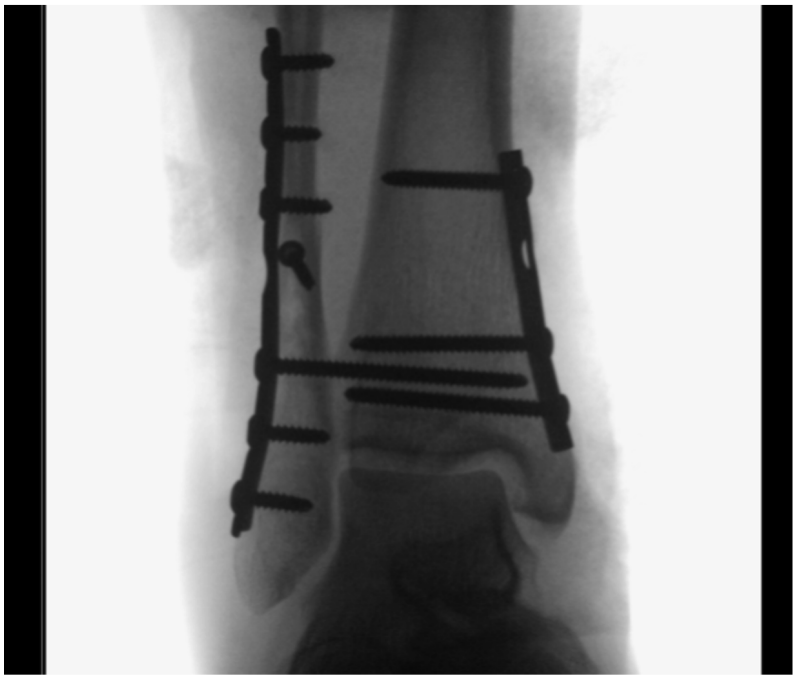
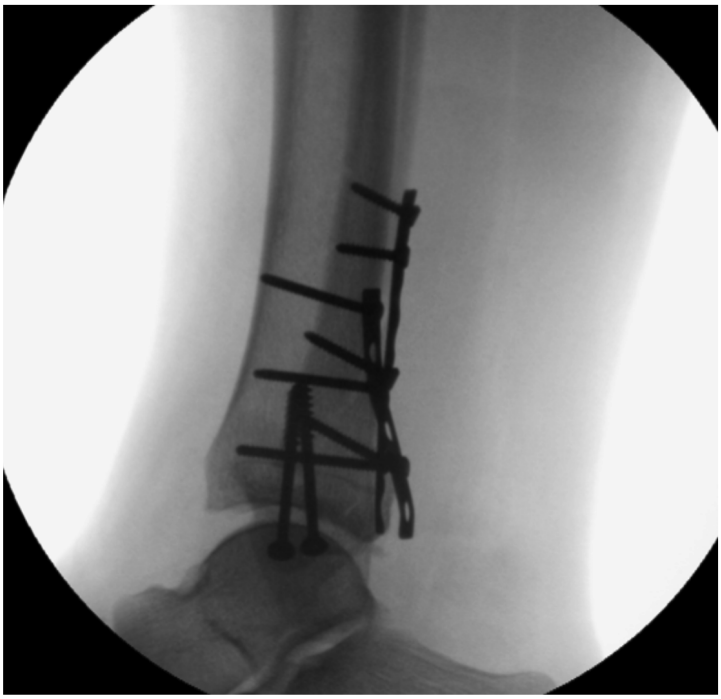


Figure 5. Intraoperative C-arm control: the ankle is shown in a lateral view, to show the trochlea tali for intra-articular view; the fibula forms the posterior one-third of the tibia.



Implementation Period

An educational period teaching standard intraoperative radiographs will be initiated before the inclusion period starts. All trauma surgeons treating the above mentioned fractures will participate in a special lecture about standardization of intraoperative radiographs. Additionally, a tailored personal instruction lecture by the senior surgeon will be initiated.

In a last educational step, personal hand-outs, providing information about appropriate standardization, are given to each trauma surgeon as well as a special lecture involving the residents and interns who participate in orthopedic trauma surgery. They are informed about the necessity of a signed informed consent as the surgeon at the emergency department is responsible for the inclusion.

Inclusion Period

After sufficient information is provided to all trauma surgeons, inclusion starts. Intraoperative radiographs are obtained by the ZIEHM SOLO (Ziehm Imaging GmbH). These radiographs are saved separately in our PACS. The next day, these radiographs are evaluated by the senior surgeon of the department of orthopedics and trauma surgery and a senior radiologist. If the quality of the intraoperative radiographs is standardized and good, the patient remains in the study arm to which they were assigned.

Randomization

Prior to the operation, the patients are randomized to the control or the intervention group. The randomization will be carried out using a special in-hospital randomization tool. Thereafter, the patient will be included in one of the following two groups.

Control Group (Group I)

Patients belonging to the control group will be treated according to our current protocol, including a standardized postoperative radiograph. Regular postoperative radiographs are performed in the department of radiology and contain an anterior-posterior and lateral tilted view of the wrist. In patients with an ankle fracture, a lateral view and a mortise view are obtained. Generally, patients operated on for a distal radius fracture will have a functional after treatment and will be restricted in weight bearing for the first 6 weeks postoperative. Operatively treated ankle fractures will generally receive a functional after treatment with weight bearing restricted to 15 kg for 6 weeks. All patients will have a clinical and radiological outpatient control after 6 weeks.

Intervention Group (Group II)

Patients belonging to the intervention group will have a modified postoperative protocol. After interdisciplinary consensus (between the head of the orthopedics department and trauma surgery and radiology) that the quality of the intraoperative radiographs are performed in a standardized fashion with good quality and that the reposition is within the predefined standards, patients will be mobilized according to our routine postoperative protocol and will be reviewed in our outpatient department after 6 weeks (clinically and radiologically).

If the quality of the intraoperative radiographs or the reduction is not acceptable, additional imaging (plain radiographs or CT) or a revision operation is performed.

Outcome

As we present a noninferiority approach, the purpose of this analysis is to show comparable results in the intervention group in terms of functional outcome and safety for the patients.

Primary Outcome

We defined the primary outcome as any re-intervention or necessity for additional imaging due to insufficient intraoperative imaging or achieved reduction.

Secondary Outcome

Secondary outcome is measured using different specific functional outcome scores as well as nonspecific outcome tools.

Specific Tools

The following tools used correspond to the specific fracture treated.

Patient-Rated Wrist Hand Evaluation

Patients operated for distal radius fractures will fill out the Patient-Rated Wrist Hand Evaluation. This 15-item questionnaire includes both a section concerning functional outcome as well as a section that analyzes the associated pain. In both sections, the items (range from 0 reflecting no pain/disability to 100) are added up and the sum is divided by two.

Range of Motion

Additionally the range of motion (ROM) will be measured during the outpatient clinical visits. The ROM of the wrist will be tested by measuring the palmar flexion, dorsal extension, pronation, supination, and radial and ulnar deviation.

Foot and Ankle Outcome Score

Patients operated for ankle fractures have to complete the Foot and Ankle Outcome Score (FAOS Score). This score was developed to assess the patient's opinion about nonspecific ankle problems. The FAOS consists of 5 subscales that investigate pain, function in daily living, function in sport and recreation, foot and ankle-related quality of life, and other symptoms. Standardized answer options are given (% Likert boxes) and each question gets a score from 0-4. A normalized score (100 indicating no symptoms and 0 indicating extreme symptoms) is calculated for each subscale.

Nonspecific Tool

Health status measurement will be performed using the Short-Form 36-Item Health Survey (SF-36) in all patients. The SF-36 is a patient-reported survey of the patient's health [7]. The questionnaire includes eight different sections (ie, vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning, mental health). Each section has a scoring range between 0 and 100. The score is proportional to the outcome with the best possible result of 100.

The visual analogue scale is used on all our patients at the 6-week control. This scale ranges from 0-10 where 0 equals no pain and 10 the worst imaginable pain.

Additionally, a medical questionnaire answered by the trauma surgeon in the outpatient clinic investigates whether any signs of a complex regional pain syndrome or complications are present.

Statistical Analysis

Statistical analysis will be carried out using SPSS, IBM Version 21. In order to answer the primary research questions, the number of re-interventions will be compared between both groups. Additionally, we will use regression *models* to try to identify risk factors for these re-interventions such as timing and duration of surgery and the experience of the surgeon.

In order to address the secondary research questions, the evaluated specific and nonspecific outcome scores will be compared between the two groups using means and standard deviation.

Sample Size Calculation

We determined the sample size for both the radius and ankle fracture group after a retrospective evaluation concerning the rate of postoperative treatment plan changes. We found a rate of 7.2% changes in treatment plan after obtaining standard postoperative radiographs in 2014. Based on these findings and using a power of 0.80 and alpha failure of .05, we performed a sample size calculation.

Based on a noninferiority approach, we calculated 158 patients needed in each treatment arm. Corresponding to the standard protocol following operative fracture care, we have a standard first outpatient control after 6 weeks. Based on a separate evaluation, concerning the frequency of patients being lost to follow-up, 4.6% of our patients in 2014 did not obtain their 6-week control. Therefore we added 5% loss of follow-up to our power calculation, resulting in 158 x 1.05=166 patients in each treatment arm.

Ethical Considerations

The study design is in accordance with the Declaration of Helsinki [14] and with Swiss laws like the human research act (HFG) and the human research regulation (HFV).

This study was approved by the medical research ethics committee of Basel. The EKNZ accepted this trial on April 4, 2016 (EKNZ BASEC 2016-00114).

All forms handed out to the patient and the information that will be obtained using the above mentioned questionnaires are approved by the EKNZ. Essential changes in the course of the trial will be reported immediately and need to be approved by the ethics committee. Information as well as results will not be presented to the EKNZ on a regular basis. However, for data verification, authorized representatives of the project manager or the ethics committee do have access to the medical data relevant to the project, including the medical history of participants at any time.

Serious adverse events must be reported immediately and if potential life-threatening complications occur, the trial will be stopped until safety is proven by the ethics committee. Patients participating in this clinical trial are covered by a special hospital insurance. This insurance is free for patients and covers any damage or potential damage as well as death caused by the study.

Results

The trial was started in August 2016 and 104 patients have been enrolled up to this point. See Multimedia Appendix 2 for the schedule of enrollment.

Discussion

Principal Considerations

To date, standard postoperative radiographs are often obtained after operatively treated fractures despite the fact that the value of these standardized postoperative radiographs remains under debate. Considering a trend towards cost-effective medicine and keeping the cumulative radiation exposure in mind, this discussion is not surprising. Initial studies in 1996 showed a lack of additional information gained by the postoperative radiograph [15,16] compared to obtained intraoperative radiographs.

A prospective randomized trial is needed to investigate whether a change to a protocol without standard postoperative radiographs is justifiable or not. Therefore, our study is important to proving that standard postoperative radiographs are not needed and that the outcome of patients without a standard postoperative radiograph is comparable with those treated according to the current protocol (ie, postoperative radiographs are routinely performed).

Strengths and Limitations

This randomized trial will provide prospective data for a common health care problem and the appropriate aftercare. Due to its randomized design, this trial will provide high-quality evidence with clearly predefined objective results.

Outpatient control, for all patients 6 weeks postoperatively, performed by a specialized (orthopedic) trauma surgeon ensures comparable and objective analysis in terms of early surgical outcome.

As a secondary outcome measurement, functional analysis using validated questionnaires will provide additional information. If this trial demonstrates comparability in terms of quality in aftercare between both groups, this could be the basis for a change in standard postoperative process with savings in radiation exposure and health care costs.

Limitations of this study are the single-center design and the lack of blinding during the study period.

Conclusion

With this prospective randomized trial, we will provide data on the necessity for postoperative standardized radiographs after open reduction and internal fixation of distal radius and ankle fractures. This could lead to a change in the standard postoperative protocol after operative fracture care and will help reduce radiation exposure during the postoperative course.

Authors' Contributions

FO and FB drafted the manuscript and designed the study. FB will act as trial coordinator. LF, BCL, AR, and JM will participate in the organization, control, and follow-up of the patients. Statistical analysis will be carried out by FO and FB as well as RB. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

SPIRIT checklist.

[PDF File (Adobe PDF File), 60KB - resprot_v6i8e159_app1.pdf]

Multimedia Appendix 2

Time schedule of enrollment and assessment for patients participating in this trial.

[PDF File (Adobe PDF File), 77KB - resprot_v6i8e159_app2.pdf]

References

1. Chaudhry S, Egol KA. Ankle injuries and fractures in the obese patient. Orthop Clin North Am 2011 Jan;42(1):45-53, vi. [doi: 10.1016/j.ocl.2010.07.003] [Medline: 21095434]
2. King CM, Hamilton GA, Cobb M, Carpenter D, Ford LA. Association between ankle fractures and obesity. J Foot Ankle Surg 2012;51(5):543-547. [doi: 10.1053/j.jfas.2012.05.016] [Medline: 22789485]
3. Orthopaedic Trauma Association Committee for Coding and Classification. Fracture and dislocation compendium. J Orthop Trauma 1996;10 Suppl 1:v-ix, 1. [Medline: 8814583]
4. Marsh JL, Slongo TF, Agel J, Broderick JS, Creevey W, DeCoster TA, et al. Fracture and dislocation classification compendium - 2007: Orthopaedic Trauma Association classification, database and outcomes committee. J Orthop Trauma 2007;21(10 Suppl):S1-133. [Medline: 18277234]

5. Tufescu T. Working toward reducing postoperative fracture radiographs: a survey of Canadian surgeons. Can J Surg 2016 Feb;59(1):26-28 [FREE Full text] [Medline: 26812405]
6. Tufescu T. The cost of screening radiographs after stable fracture fixation. Can J Surg 2017 Feb;60(1):53-56 [FREE Full text] [Medline: 28234590]
7. Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998 Nov;51(11):1055-1068. [Medline: 9817123]
8. Weil NL, Termaat MF, Rubinstein SM, El Moumni MM, Zuidema WP, Derksen RJ, et al. WARRIOR-trial - is routine radiography following the 2-week initial follow-up in trauma patients with wrist and ankle fractures necessary: study protocol for a randomized controlled trial. Trials 2015 Feb 27;16(1):66 [FREE Full text] [doi: 10.1186/s13063-015-0600-x] [Medline: 25872504]
9. Koordinationsgruppe für die Statistik der Unfallversicherung. Sammelstelle fuer die Statistik der Unfallversicherung UVG (SSUV). 2015 Jun 11. Unfallstatistik UVG 2015 URL: https://www.unfallstatistik.ch/d/publik/unfstat/pdf/Ts15.pdf [accessed 2017-08-09] [WebCite Cache ID 6sacXW71I]
10. 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 Feb 05;158(3):200-207 [FREE Full text] [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
11. Gustilo RB, Anderson JT. Prevention of infection in the treatment of one thousand and twenty-five open fractures of long bones: retrospective and prospective analyses. J Bone Joint Surg Am 1976 Jun;58(4):453-458. [Medline: 773941]
12. Ozer K, Wolf JM, Watkins B, Hak DJ. Comparison of 4 fluoroscopic views for dorsal cortex screw penetration after volar plating of the distal radius. J Hand Surg Am 2012 May;37(5):963-967. [doi: 10.1016/j.jhsa.2012.02.026] [Medline: 22480500]
13. Brunner A, Siebert C, Stieger C, Kastius A, Link B, Babst R. The dorsal tangential X-ray view to determine dorsal screw penetration during volar plating of distal radius fractures. J Hand Surg Am 2015 Jan;40(1):27-33. [doi: 10.1016/j.jhsa.2014.10.021] [Medline: 25457447]
14. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. J Postgrad Med 2002;48(3):206-208 [FREE Full text] [Medline: 12432198]
15. Pattison RM, Calzada S, Koka SR, James SE. Postoperative radiographs or thermal prints after internal fixation of fractures? A study of DHS fixation of hip fractures. Ann R Coll Surg Engl 1996 Nov;78(6):509-511 [FREE Full text] [Medline: 8943634]
16. Chakravarthy J, Mangat K, Qureshi A, Porter K. Postoperative radiographs following hip fracture surgery. Do they influence patient management? Int J Clin Pract 2007 Mar;61(3):421-424. [doi: 10.1111/j.1742-1241.2006.01279.x] [Medline: 17313609]

Abbreviations

AO: Arbeitsgemeinschaft für Osteosynthesefragen
CT: computed tomography
EKNZ: Ethikkomission Nordwest- und Zentralschweiz
FAOS: Foot and Ankle Outcome Score
PACS: Picture Archiving and Communication System
ROM: range of motion
SF-36: Short-Form 36-Item Health Survey

Edited by G Eysenbach; submitted 15.03.17; peer-reviewed by N Dewan; comments to author 30.05.17; revised version received 30.06.17; accepted 30.06.17; published 16.08.17.

Please cite as:
Oehme F, Rühle A, Mühlhäusser J, Fourie L, Link BC, Babst R, Beeres FJP
Are Routine Radiographs Needed the Day After Open Reduction and Internal Fixation Surgery for Distal Radius and Ankle Fractures: Study Protocol for a Prospective, Open Label, Randomized Controlled Trial
JMIR Res Protoc 2017;6(8):e159
URL: http://www.researchprotocols.org/2017/8/e159/
doi:10.2196/resprot.7698
PMID:28814379

Protocol

The Spanish Version of the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Adolescents (UP-A) Adapted as a School-Based Anxiety and Depression Prevention Program: Study Protocol for a Cluster Randomized Controlled Trial

Julia García-Escalera[1], MS; Rosa M Valiente[1], PhD; Paloma Chorot[1], PhD; Jill Ehrenreich-May[2], PhD; Sarah M Kennedy[2], PhD; Bonifacio Sandín[1], PhD

[1]Faculty of Psychology, Universidad Nacional de Educación a Distancia, Madrid, Spain

[2]Department of Psychology, University of Miami, Coral Gables, FL, United States

Corresponding Author:
Paloma Chorot, PhD
Faculty of Psychology
Universidad Nacional de Educación a Distancia
Juan del Rosal, 10
Madrid, 28040
Spain
Phone: 34 913987940
Fax: 34 913987748
Email: pchorot@psi.uned.es

Abstract

Background: Anxiety and depression are common, impairing conditions that evidence high comorbidity rates in adolescence. The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Adolescents (UP-A) is one of the few existing resources aimed at applying transdiagnostic treatment principles to target core dysfunctions associated with both anxiety and depression within a single protocol. To our knowledge, this is the first study examining the efficacy of the UP-A adapted as a universal preventive intervention program.

Objective: The primary aim of this study is to examine whether the Spanish version of the UP-A is more effective than a waitlist (WL) control group in reducing and preventing symptoms of anxiety and depression when employed as a universal, classroom-based preventive intervention. The secondary aim is to investigate changes in a broad range of secondary outcome measures, including negative and positive affect, anxiety sensitivity, emotional avoidance, top problems ratings, school grades, depression and anxiety-related interference, self-esteem, life satisfaction, quality of life, conduct problems, hyperactivity/inattention symptoms, peer problems, prosocial behavior, school adjustment, and discipline problems. Other aims are to assess a range of possible predictors of intervention effects and to examine the feasibility and the acceptability of implementing UP-A in a prevention group format and in a school setting.

Methods: A cluster, randomized, WL, controlled trial design with classroom as the unit of randomization was used in this study. Five classes including a total of 152 adolescents were randomized to the experimental or WL control groups. Participants in the experimental group received 9 55-minute sessions delivered by advanced doctoral and masters students in clinical psychology. The WL control group will receive the intervention once the 3-month follow-up assessment is completed.

Results: We have recruited participants to the cluster randomized controlled trial (RCT) and have conducted the intervention with the experimental group. We expect the WL control group to complete the intervention in July 2017. Data analysis will take place during the second semester of 2017.

Conclusions: We expect the experimental group to outperform the WL control group at post-intervention and 3-month follow-up. We also expect the WL control group to show improvements in primary and secondary outcome measures after receiving the intervention. Results will have implications for researchers, families, and education providers.

Trial Registration: Clinicaltrials.gov NCT03123991; https://clinicaltrials.gov/ct2/show/NCT03123991 (Archived by WebCite at http://www.webcitation.org/6qp7GIzcR)

(JMIR Res Protoc 2017;6(8):e149) doi:10.2196/resprot.7934

KEYWORDS

universal prevention; transdiagnostic; cluster; randomized controlled trial; unified protocol; adolescents; anxiety; depression; emotional disorders; school intervention

Introduction

Background

Depression and anxiety disorders are highly prevalent conditions in children and adolescents and are associated with significant impairment in everyday life [1,2]. For instance, a national study conducted in the United States including adolescents aged 13 to 18 years found that anxiety disorders were the most common condition (31.9%) and that mood disorders were also highly prevalent (14.3%) [2]. In Spain, after conduct disorders, depressive disorders are the most prevalent psychological disorders in children and adolescents (14.6%), followed closely by anxiety disorders (13.3%) [1].

The costs of supporting children and adolescents with psychiatric disorders can be far higher than for their peers without these problems, and these disorders usually lead to continued financial burden and functional impairment into adulthood (eg, increased use of public sector services, lower-quality peer and family relationships, and reduced participation in the labor market) [3]. As an example, a Dutch cost-of-illness study focusing on 118 anxious 8 to 18-year old youth concluded that "the societal costs of families with a clinically anxious child who seek treatment amount to more than 20 million Euros a year in the Dutch population, and were 21 times higher than in families of the general population" [4]. Beyond the significant emotional and financial costs of childhood emotional disorders, depression and anxiety disorders are associated with increased risk for a range of difficulties throughout adolescence and into adulthood. Specifically, adolescent anxiety and depression are associated with an increased risk of educational under-achievement, substance dependence, chronic anxiety and/or depression, as well as impairment in a range of important psychosocial domains including interpersonal functioning, quality of life, and physical well-being [5,6].

Despite this chronic and concerning level of impairment, only about 20% of youth with anxiety or depressive disorders receive mental health services [7]. Given the sequelae of negative outcomes associated with childhood emotional disorders, preventative interventions are all the more important for their potential to reduce and even eliminate the distress, long-term impairment, and significant societal burden these disorders cause. Moreover, delivering universal prevention programs for anxiety and depression in a school setting may be an ideal method for making preventative interventions more widely accessible to those with a high risk of developing emotional disorders.

Prevention programs may be universal, selected, or indicated [8]. A universal approach to anxiety and depression prevention in schools could be especially beneficial because such an approach "seeks to target a large number of children regardless of their risk status over a short period of time, helps to reduce difficulties in screening for inclusion, potentially reduces the incidence of anxiety disorders by intervening before the onset of these disorders, and serves to reduce stigmatization" [9]. Preventive interventions have often been developed separately for anxiety and for depression and have typically considered these disorder classes as independent constructs [10]. However, anxiety and depressive disorders are frequently comorbid in children and adolescents, with studies reporting that about 25% to 50% of depressed youth have comorbid anxiety disorders and about 10% to 15% of anxious youth have comorbid depression [11]. In addition, both clinical manifestations share vulnerability and maintenance factors such as neuroticism or high negative affect, maladaptive action tendencies (eg, avoidance or withdrawal), and high levels of distress or discomfort in response to emotional experiences [12,13].

These findings suggest that a transdiagnostic approach to the prevention of adolescent anxiety and depression may be useful. Transdiagnostic cognitive behavioral therapy (T-CBT) has been defined as a type of cognitive behavioral therapy (CBT) that can be applied to a number of different disorders that share commonalities in cognitive, behavioral, and/or emotional dysregulation [14], without tailoring the protocol to specific diagnoses [15]. Fairburn, Cooper, and Shafran [16] initially developed and evaluated a transdiagnostic approach to CBT for eating disorders. Research on the development and efficacy of transdiagnostic interventions was advanced by Barlow and colleagues, who provided a transdiagnostic theoretical model and developed a unified (transdiagnostic) treatment for emotional disorders among adults: The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders (UP) [17]. Drawing from research with the UP in adult samples, Ehrenreich-May and colleagues developed the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Adolescents (UP-A) and the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Children (UP-C) [18]. Preliminary data suggest that both protocols are effective in reducing principal and overall emotional disorder severity [19-21] and, overall, T-CBT for children and adolescents has demonstrated promising results according to a recent meta-analysis [21].

Applying a transdiagnostic approach to prevention that focuses on core dysfunctions across emotional disorders may reap a larger benefit than intervening with risk factors specific to only one disorder or domain type (eg, anxiety versus depression, etc) [10]. However, transdiagnostic, theory-driven CBT protocols for preventing anxiety and depression are still scarce, currently with only the following 2 well-consolidated protocols: (1) EMOTION: "Coping Kids" Managing Anxiety and Depression [22]; and (2) Super Skills for Life [23]. Studies applying these T-CBT-based prevention protocols have shown promising results [9,24]. In addition, a transdiagnostic, school-based

preventive intervention for adolescents with elevated symptoms of social anxiety and/or depression and elevated peer victimization was developed and initially evaluated [25]. However, the EMOTION and the Super Skills for Life protocols were developed for young children [22,23], and T-CBT preventive interventions for anxiety and depression in adolescents are scarce, to say the least. An initial version of the UP-C was developed as a universal preventive intervention in a summer camp setting [26], providing useful preliminary data on the feasibility and efficacy of adapting and delivering the UP-C and UP-A as preventive interventions, but larger studies are needed, particularly in adolescents for whom T-CBT preventive interventions are lacking.

Study Aims

The current study proposes to extend transdiagnostic prevention research for adolescents by establishing initial pre- to post-prevention and follow-up outcomes associated with the use of the Spanish version of the UP-A, adapted as a school-based anxiety and depression preventive intervention.

The primary aim of this study is to examine whether the UP-A, a T-CBT intervention adapted as a school-based preventive intervention and delivered by doctoral and masters students in clinical psychology, is more effective than a waitlist (WL) control condition consisting of the usual academic curriculum in (1) reducing symptoms of anxiety and depression; and (2) preventing clinically elevated levels of anxiety and depression from developing. The second aim is to investigate post-intervention and follow-up changes in a broad range of secondary outcome measures, including negative and positive affect, anxiety sensitivity, emotional avoidance, depression and anxiety-related interference, top problems ratings, school grades, self-esteem, life satisfaction, quality of life, school adjustment, discipline problems, conduct problems, hyperactivity/inattention symptoms, peer problems, and prosocial behavior. The third aim is to assess whether any benefits of the intervention are predicted by age, gender, number of sessions attended, participant-rated behaviors indicating level of engagement and effort during sessions, understanding of basic program concepts, adherence to home practice, practice of specific strategies outside session, and interest in psychology. The fourth aim is to examine if the effects at post-intervention are maintained at the 3-month follow-up. The fifth aim of this study is to assess the feasibility (eg, obtaining consent, assessment completion, group attendance) and the acceptability of implementing UP-A in a prevention group format and in a school setting. Finally, the sixth aim is to calculate the intracluster correlation coefficient (ICC) and measures of variability (ie, standard deviation, interquartile range) for all primary and secondary outcomes in order to determine sample sizes of future full-scale, cluster randomized controlled trials (RCTs) applying the UP-A adapted as a preventive intervention.

We predict that the UP-A preventive intervention group will exhibit greater improvement on all primary and secondary outcome measures at post-intervention and follow-up compared to the WL control group. We also expect that WL participants will experience significant improvements on all outcome measures after receiving the intervention. Greater improvement in the primary and secondary outcome variables is expected in those with higher levels of home practice, more frequent practice of a greater number of strategies outside session, higher levels of engagement and effort during sessions, more sessions attended, greater knowledge of core program contents after the intervention, and greater interest in psychology. In relation to age and gender as predictors, earlier studies have shown contradictory results [27] and therefore no hypothesis was forwarded in this regard. We also hypothesize that results will support the feasibility of school-based implementation of UP-A in a prevention group format, as evidenced by a number of participants achieving treatment completer status (attending at least 7 out of the 9 sessions), as well as the acceptability of the intervention, as evidenced by participants' self-rated satisfaction at post-intervention.

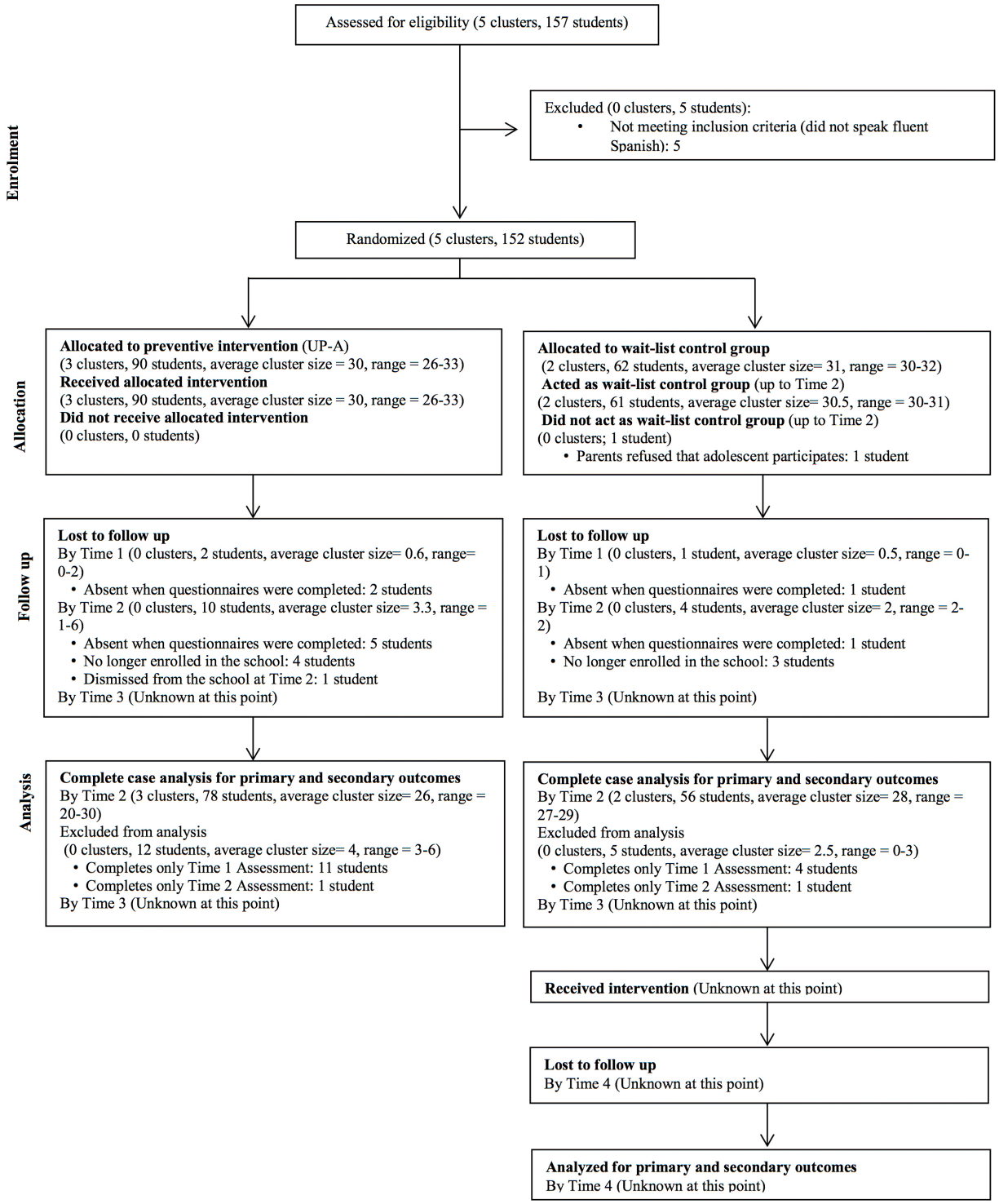
Methods

The study design reported is in line with the Consolidated Standards of Reporting Trials (CONSORT) statement: extension to cluster randomized trials [28]. The study was granted ethical approval from the Research Ethics Committee of Universidad Nacional de Educación a Distancia, Madrid, Spain. All parents or guardians as well as adolescent participants provided written informed consent. The study is registered in Clinicaltrials.gov (NCT03123991).

Study Design

This study will be implemented as a 2-arm, cluster RCT [28], with an intervention condition (UP-A adapted as a preventive intervention program) and a 3-month WL control condition. Measurements will be taken on the following occasions: (1) Time 1 (1 week before the experimental group starts the intervention); (2) Time 2 (1 week after the experimental group finishes the intervention); (3) Time 3 (3 months after the experimental group finishes the intervention and 1 week before the WL control group starts the intervention); and (4) Time 4 (1 week after the WL control group finishes the intervention). On all occasions, except for Time 4, both groups will complete the outcome measures at about the same time. For Time 4, only the WL control group will complete the outcome measures. This represents a 3 (time) by 2 (group) repeated measures design. The flow diagram of the study is shown in Figure 1.

Figure 1. CONSORT flow diagram.



Participants

An urban secondary school in the city of Madrid (Spain) that was previously known to the authors expressed interest in being involved in research and agreed to participate. Students enrolled in levels 3º ESO (usually aged 14 to 15 years old) and 4º ESO (usually aged 15 to 16 years old) were targeted, given that early to mid-adolescence is a key time for the emergence of common mental health disorders [29]. Participants came from 3 classes at level 3º ESO and 2 classes at level 4º ESO (each class having around 30 students each). No incentives are being given to the adolescents or the school for participating in this project.

Parents or guardians, as well as adolescents, were provided with information about the study and were required to provide signed consent before Time 1 assessments. Participants and families were informed that (1) responses to the questionnaires will be keep confidential; (2) they will be able to withdraw from the study at any time; (3) participants receiving the preventive intervention will be taught specific skills that will help them to better cope with situations that provoke anxiety/depression both now and in the future; and (4) participants assigned to the WL

control condition will receive the preventive intervention after the WL period. Inclusion criteria for participants were providing written, informed consent (both the adolescent and at least 1 parent or legal guardian) and being able to understand, write, and read Spanish. Spanish proficiency was determined based on teacher report. Due to the universal prevention goal of this study, there were no other exclusion criteria. Classes of students were randomly assigned either to the preventive intervention group or to the WL control group (see Figure 1). Time 1, Time 2, Time 3, and Time 4 assessments, as well as intervention sessions, are being conducted at the school during school hours.

Intervention

Participants will receive the Spanish version of the UP-A [18], modified for delivery as a 9-session, school-based universal preventive intervention. Core modules of the UP-A include (1) Building and Keeping Motivation; (2) Getting to Know Your Emotions and Behaviors; (3) Emotion-Focused Behavioral Experiments; (4) Awareness of Physical Sensations; (5) Being Flexible in Your Thinking; (6) Awareness of Emotional Experiences; (7) Situational Emotion Exposures; and (8) Keeping it Going-Maintaining Your Gains [18] (Figure 1).

The UP-A is a downward extension of the UP [17], modified for adolescents (ages 12 to 17) presenting with any primary emotional disorder, including anxiety, depressive, obsessive-compulsive, and stress-related disorders and problem areas. The UP-A distills common evidence-based techniques that cut across disorder-specific treatment manuals for youth emotional disorders (eg, psychoeducation, non-judgmental awareness, cognitive reappraisal, exposure, behavioral activation, etc) [30], drawing from CBTs, motivational enhancement, and third-wave behavioral therapies (eg, acceptance and mindfulness-based techniques). As with the UP, the UP-A emphasizes 5 core treatment principles: (1) increasing present-focused awareness of emotions; (2) enhancing cognitive flexibility; (3) identifying and preventing emotional avoidance and maladaptive emotion-driven behaviors; (4) increasing emotional awareness and acceptance of uncomfortable emotion-related physiological sensations; and (5) facilitating exposure to both interoceptive and situational triggers of emotional experiences [17,31].

The preventive intervention applied in this study consists of 9 weekly lessons, the length of which corresponds to a school's typical class period (55 minutes in the school in our study). It is delivered in a group format to entire classes of adolescents as part of the school curriculum. Specifically, the intervention sessions are carried out during school hours designated for "Tutorías," which are 1-hour weekly sessions that, in the Spanish Education System, are meant to target issues occurring within the school context, such as providing professional development, providing academic support, assisting in solving problems between students or between students and teachers, etc. The WL control group will receive their normal class schedule without any planned socioemotional focus, followed by the intervention after the Time 3 assessment is completed. A detailed description of the content of each UP-A session is shown in Table 1.

The students in the preventive intervention group are encouraged to practice skills learned in sessions by completing structured home learning assignments outside of formal session time. Completed home learning assignments are discussed at the beginning of each session, with the exception of the first. All intervention sessions are delivered by author JGE, an advanced doctoral student in clinical psychology, and by an advanced master's student in clinical psychology.

Researchers attempt to contact students who miss one of the weekly sessions and provide them with the opportunity to make up the content in the following days. During this makeup session, students are given a content summary and any missed home learning assignments to facilitate preparation for the next session. If a student is unable to be contacted, he or she is able participate in the following session after a brief, individual content review meeting with one of the group leaders delivering the sessions.

Implementation of the Program

Prior to implementing the UP-A program, JGE received training on the UP-A protocol by its developer, JEM at University of Miami (Coral Gables, US). Likewise, RMV, PC, and BS received specific training on the UP-A through a course on the UP-A and UP-C protocols at Universidad Nacional de Educación a Distancia (Madrid, Spain) given by JEM. The UP-A was translated into Spanish by JGE, and its translation and adaptation were supervised by BS, PC, and RMV. The translation process was also supervised by JEM. Adherence to the protocol for the current study is self-monitored by the group leaders, who complete a checklist at the end of each session indicating whether each skill within the session was presented.

Primary Outcome Measures

Questionnaires will be completed at the Time 1, Time 2, Time 3, and Time 4 assessments (Table 2). Top Problems Assessment (TPA) is not included in this table because it was completed at time-points during the 1st, 5th, and 9th sessions (see below).

Table 1. The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Adolescents (UP-A) preventive intervention session descriptions.

Session	UP-A[a] module	Module title	Main content
1	Module 1	Building and Keeping Motivation	Introduce confidentiality and group rules; obtain 3 top problems, severity ratings, and a SMART[b] goal for each problem; complete emotion identification skills activity if sufficient time remains.
2	Module 2	Getting to Know Your Emotions and Behaviors	Psychoeducation about emotions and their function; introduce emotional behaviors, the 3 parts of an emotional experience, and the "Before, During, and After" form for tracking emotional experiences outside of sessions.
3	Module 3	Emotion-Focused Behavioral Experiments	Psychoeducation about cycle of avoidance, opposite action, and behavioral experiments; reflect on current use of free time and come up with a list of enjoyed activities; introduce weekly activity tracker for ongoing behavioral activation.
4	Module 4	Awareness of Physical Sensations	Psychoeducation about body sensations, their relationship to intense emotions and their harmlessness; introduce the concept of "fight or flight response" and review cycle of avoidance; conduct sensational exposures with the group.
5	Module 5	Being Flexible in Your Thinking	Introduce the concept of "thinking traps" (ie, cognitive distortions) and teach common thinking traps; introduce the concept of automatic and alternative thoughts as well as detective thinking skills; Re-rate top problems obtained in session 1.
6	Module 5	Being Flexible in Your Thinking	Review thinking traps and detective thinking skills; introduce and ensure understanding of problem solving skills; conduct examples using problem solving skills with group members; review skills learnt so far in the program.
7	Module 6	Awareness of Emotional Experiences	Introduce the rationale for present-moment awareness and practice this skill in session using non-emotional stimuli (eg, focus on breathing); introduce rationale for non-judgmental awareness; do an individual mini-test assessing skills taught in the program so far.
8	Module 7	Situational Emotion Exposures	Review cycle of avoidance, reinforcement, and maintenance of learned behavior; provide psychoeducation about emotion exposures; create emotional behaviors forms to identify relevant exposures; if time permits, conduct a group exposure activity; assign exposure homework.
9	Module 8	Keeping it Going—Maintaining Your Gains	Review exposure homework and plan future exposures if necessary; re-rate top problems and revisit SMART goals; review skills that have been most useful for each group member and make an individualized post-program plan to practice skills.

[a]UP-A: Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Adolescents.

[b]SMART: specific, measurable, attainable, relevant, and time-bound.

Table 2. Outcome measures used in the study.

Questionnaire	Assessment								Items, n
	Time 1		Time 2		Time 3		Time 4		
	E[a]	WL[b]	E	WL	E	WL	E	WL	
Demographics[c]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	9/20
RCADS-30[d]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	30
CDN[e]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	14
EAN[f]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	10
SDQ[g]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	25
EIDAN[h]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	11
PANASN[i]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	20
CASI[j]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	18
EASI-A[k]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	17
SWLS-C[l]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	5
KIDSCREEN-10[m]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	10
SES[n]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	10
EBAE-10[o]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	10
IG[p]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	10
Satisfaction Q[q]	No	No	Yes	No	No	No	No	Yes	10
Curriculum Q[r]	No	No	Yes	No	Yes	No	No	Yes	3
Strategies Q[s]	No	No	Yes	No	Yes	No	No	Yes	8
Discipline Q[t]	No	No	Yes	No	No	No	No	Yes	6
Other questions[u]	No	No	Yes	No	No	No	No	Yes	4
Psychology Q[v]	No	No	Yes	Yes	Yes	Yes	No	Yes	3
PID-5-Brief[w]	No	No	Yes	Yes	No	No	No	No	25

[a]E: experimental group.

[b]WL: waitlist control group.

[c]The demographics questionnaire has 20 items at Time 1 and 9 items at Time 2, Time 3, and Time 4.

[d]RCADS-30: Revised Child Anxiety and Depression Scale-30.

[e]CDN: Depression Questionnaire for Children (Cuestionario de Depresión para Niños).

[f]EAN: Anxiety Scale for Children (Escala de Ansiedad para Niños).

[g]SDQ: Strengths and Difficulties Questionnaire.

[h]EIDAN: Depression and Anxiety Interference Scale for Children (Escala de Interferencia de la Depresión y Ansiedad para Niños).

[i]PANASN: Positive and Negative Affect Schedule for Children and Adolescents (Escalas PANAS de Afecto Positivo y Negativo para Niños y Adolescentes).

[j]CASI: Childhood Anxiety Sensitivity Index.

[k]EASI-A: Emotional Avoidance Strategy Inventory for Adolescents.

[l]SWLS-C: Satisfaction with Life Scale for Children.

[m]KIDSCREEN-10: Kidscreen-10 Quality of Life Scale.

[n]SES: Self-Esteem Scale.

[o]EBAE-10: School Adjustment Brief Scale (Escala Breve de Ajuste Escolar).

[p]IG: General Indiscipline Scale (Escala de Indisciplina General).

[q]Satisfaction Q: satisfaction with the program questionnaire.

[r]Curriculum Q: curriculum knowledge questionnaire.

[s]Strategies Q: strategies practiced outside of session questionnaire.

[t]Discipline Q: discipline problems during sessions questionnaire.

[u]Other questions: other end of program questions.

[v]Psychology Q: assistance to therapy and interest for psychology questions.

[w]PID-5-Brief: Personality Inventory for Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5)-Brief Form.

Revised Child Anxiety and Depression Scale-30

The Revised Child Anxiety and Depression Scale-30 (RCADS-30) [32] is a widely-used questionnaire measuring self-reported anxiety and depressive symptoms in children and adolescents. The scale is comprised of the following subscales derived from the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) criteria: (1) social phobia, (2) generalized anxiety disorder, (3) obsessive-compulsive disorder, (4) panic disorder, (5) separation anxiety disorder, and (6) major depressive disorder. The 6 subscales are summed to create a Total Anxiety and Depression score. Each item is scored from 0 ("Never") to 3 ("Always"), with higher scores representing more severe symptoms. The RCADS-30 has demonstrated good psychometric properties with normative and clinical populations [33,34]. In the current sample, the 6 individual subscales have demonstrated adequate to good internal consistency both at Time 1 (alpha .70 to .85) and at Time 2 (alpha .72 to .86). The Total Anxiety and Depression scale has demonstrated excellent internal consistency at Time 1 (alpha = .92) and at Time 2 (alpha = .93) assessments.

Depression Questionnaire for Children

Depression Questionnaire for Children (Cuestionario de Depresión para Niños; CDN) [35] is a 16-item, self-report questionnaire designed to assess symptoms of DSM-IV major depressive disorder and dysthymic disorder in children and adolescents. In this study, the 2 items targeting suicidal ideation were not included as per the request of school personnel. Participants rate each item on a 3-point scale from 0 ("Never"), 1 ("Sometimes"), to 2 ("Very often") to indicate the frequency with which they experience depression symptoms. The sum of all items provides an overall score, with higher scores indicating greater depression symptoms. The CDN has demonstrated adequate psychometric properties [32]. At Time 1 and Time 2 assessments, the CDN overall scale has demonstrated good internal consistency in this sample with alpha values of .87 and .89, respectively.

Anxiety Scale for Children

Anxiety Scale for Children (Escala de Ansiedad para Niños; EAN) [35] is a 10-item questionnaire that assesses anxiety symptoms during the past few weeks in children and adolescents. Participants are instructed to indicate how frequently they have experienced general anxiety symptoms on a 4-point, Likert-type scale, ranging from 0 ("Never or almost never") to 3 ("A lot of the times or almost always"). The sum of all items provides an overall score, with higher scores indicating more elevated anxiety symptoms. The EAN has shown good psychometric properties [36]. In the current sample, the EAN scale has demonstrated excellent internal consistency at Time 1 (alpha = .91) and Time 2 (alpha = .94) assessments.

Secondary Outcome Measures

The following questionnaires will be completed at Time 1, Time 2, Time 3, and Time 4 assessments (see Table 2).

Top Problems Assessment

The adolescent version of TPA provides a means of collecting information about changes in the severity of problems identified by the adolescent to be of greatest concern. Group leaders first provide instructions and examples about the types of problems that can be targeted within the program and adolescents are then asked to generate 3 top problems of their own. Adolescents also generate a corresponding specific, measurable, attainable, relevant, and time-bound (SMART) goal for each problem. During the 1st, 5th, and 9th sessions, participants are asked to rate the severity of each problem on a 0 to 8 scale (with higher ratings indicating greater problem severity). Top problems are a central tool for progress monitoring in the UP-A and UP-C protocols. They were adapted by Ehrenreich-May from original work by Weisz and colleagues [37]. Top problems have been shown to demonstrate good psychometric properties [37].

Strengths and Difficulties Questionnaire

The Spanish version of the Strengths and Difficulties Questionnaire (SDQ) by García and colleagues was used [38,39]. The scale provides scores for 5 subscales including emotional symptoms, conduct problems, symptoms of hyperactivity/ inattention, peer problems, and prosocial behavior. Each subscale contains 5 items scored on a 3-point, Likert-type scale, ranging from 0 ("Not true") to 2 ("Certainly true"). A Total Difficulties Score is obtained by summing the 5 items on all scales, with the exception of the prosocial behavior scale. The SDQ is widely used and both the original and Spanish versions have good psychometric properties [40,41]. The Total Difficulties Scale has demonstrated adequate to good internal consistency in this sample at Time 1 (alpha = .73) and Time 2 (alpha = .81) assessments.

Depression and Anxiety Interference Scale for Children

Students complete a Depression and Anxiety Interference Scale for Children (Escala de Interferencia de la Depresión y Ansiedad para Niños; EIDAN) [42] developed for this investigation. This 11-item questionnaire assesses the degree to which feeling worried, nervous, or sad interferes with various domains of the adolescent's life (school, peer/family/general functioning), employing a 4-point, Likert-type scale ranging from 0 ("Nothing or almost nothing") to 3 ("A lot"). Higher scores indicate greater levels of interference. In the current sample, the EIDAN scale has demonstrated good internal consistency at Time 1 (alpha = .80) and Time 2 (alpha = .87) assessments.

Positive and Negative Affect Schedule for Children and Adolescents

The Positive and Negative Affect Schedule for Children and Adolescents (Escalas PANAS de Afecto Positivo y Negativo para Niños y Adolescentes; PANASN) [43] questionnaire is an age-downward version of the Positive and Negative Affect Schedule [44] for individuals aged 7 to 17 years. The scale provides scores for 2 subscales of 10 items each measuring positive and negative affect. Participants are asked to rate items according to how they usually feel from 1 ("Never or almost never"), 2 ("Sometimes"), to 3 ("A lot of the time"). The PANASN has demonstrated good psychometric properties [43,45]. In the current sample, the Positive Affect Scale demonstrated adequate and good psychometric properties at Time 1 (alpha = .77) and Time 2 (alpha =.81) assessments. Similarly, the Negative Affect Scale demonstrated adequate internal consistency at Time 1 assessment (alpha = .79) and good internal consistency at Time 2 assessment (alpha = .81).

Childhood Anxiety Sensitivity Index

The Spanish version of the Childhood Anxiety Sensitivity Index (CASI) [46,47] was used for this study. The CASI is an 18-item self-report questionnaire measuring anxiety sensitivity in children or distress reactions to symptoms of anxiety (eg, "It scares me when my heart beats fast"). Participants rate the frequency with which they experience each item using a 3-point, Likert-type scale from 1 ("Never"), 2 ("Sometimes"), to 3 ("A lot of the time"). The Spanish adaptation of the CASI used in the present study has demonstrated good psychometric properties and positive correlations with constructs related to anxiety sensitivity [48]. In the current sample, the CASI has demonstrated excellent internal consistency at Time 1 (alpha = .90) and Time 2 (alpha = .90) assessments.

Emotional Avoidance Strategy Inventory for Adolescents

Emotional Avoidance Strategy Inventory for Adolescents (EASI-A) [49] is a 17-item self-report questionnaire in which respondents are instructed to indicate the degree to which each statement is true using a 5-point, Likert-type scale ranging from 0 ("Not at all true of me") to 4 ("Extremely true of me"). The EASI-A was translated and adapted to Spanish for this study [50]. Our Spanish adaptation also uses a 5-point Likert-type scale, but with options indicating frequency rather than degree from 0 ("Never or almost never"), 1 ("Seldom"), 2 ("Sometimes"), 3 ("A lot of times"), to 4 ("Always or almost always"). The EASI-A has demonstrated good psychometric properties and positive correlations with anxiety and depression symptoms [49]. The EASI scale has demonstrated good internal consistency at Time 1 and Time 2 assessments with alpha values of .87. and .86, respectively, in the current sample.

Satisfaction with Life Scale for Children

The Satisfaction with Life Scale for Children (SWLS-C) [51] measure is an age-downward version of the measure of life satisfaction developed by Diener and colleagues [52]. It is a 5-item, self-report instrument in which respondents are asked to indicate the degree to which each statement is true of their life using a 4-point, Likert-type scale ranging from 1 ("Not at all") to 4 ("A lot or completely"). In the current sample, the SWLS scale has demonstrated adequate good consistency at Time 1 (alpha = .81) and Time 2 (alpha = .89) assessments.

Kidscreen-10

The official Spanish version developed by the KIDSCREEN Group was used [53]. KIDSCREEN-10 is a widely-used brief questionnaire assessing generic health-related quality of life in children and adolescents that was adapted from the longer KIDSCREEN-27 [53]. The 10 items on the KIDSCREEN-10 index assess affective symptoms of depressed mood, difficulty concentrating, decreased energy, impaired school functioning, and impaired relations with peers and family. It uses a 5-point, Likert-type scale ranging from 1 ("Not at all") to 5 ("Extremely"). KIDSCREEN-10 has been shown to possess good psychometric properties [54]. In the current sample, the KIDSCREEN-10 scale has demonstrated good internal consistency at Time 1 (alpha =.82) and Time 2 (alpha = .85) assessments.

Self-Esteem Scale

A Spanish adaptation of the Self-Esteem Scale (SES) [55] was used in this study [56]. SES is a 10-item self-report instrument that measures global self-esteem. Respondents are asked to indicate how much they agree with each statement on a 4-point, Likert-type scale from 1 ("Strongly disagree") to 4 ("Strongly agree"). The scale has been shown to possess good psychometric properties in adolescent samples [57]. In the current sample, the SES scale has demonstrated good internal consistency at Time 1 (alpha = .83) and excellent internal consistency at Time 2 (alpha = .90) assessments.

School Adjustment Brief Scale

School Adjustment Brief Scale (Escala Breve de Ajuste escolar; EBAE-10) [58] is a 10-item questionnaire used to measure adaptive functioning in the area of school performance. Respondents are instructed to respond to questions about their grades, their relationships with teachers and peers, and their expectations regarding their educational future. They are asked to indicate how much they agree with each statement on a 6-point, Likert-type scale ranging from 1 ("Completely agree") to 6 ("Completely disagree"). This scale has demonstrated adequate psychometric properties [59]. In the current sample, the EBAE scale has demonstrated acceptable internal consistency at Time 1 (alpha=.76) and Time 2 (alpha=.72) assessments.

General Indiscipline Scale

The General Indiscipline Scale (Escala de Indisciplina General; IG), adapted from Martín and colleagues' questionnaire [60], is an 11-item questionnaire that assesses problematic behaviors of students in the classroom. Respondents are asked to indicate the frequency of each behavior stated in the items on a 4-circle bull's-eye, with each circle closer to the bulls-eye's center representing more frequent demonstration of the behavior. For this study, we adapted the scale such that adolescents were asked to rate the frequency with which they demonstrated each behavior on a 4-point, Likert-type scale with options being 0 ("Never or almost never"), 1 ("Only sometimes"), 2 ("Quite a few times"), or 3 ("A lot of the time"). The IG has demonstrated adequate properties [61]. In the current sample, this scale has

demonstrated adequate internal consistency at Time 1 (alpha = .75) and good internal consistency at Time 2 (alpha= .80) assessments.

Socio-Demographic Information Questionnaire

All participants are asked to provide demographic information including their age, gender, school grades, socio-economic status, and place of birth.

Other Secondary Outcome Measures

The Personality Inventory for DSM-5-Brief Form

The Spanish version of the Personality Inventory for Diagnostic and Statistical Manual of Mental Disorders, 5th edition-Brief Form (PID-5-BF) [62,63] was used. It assesses 5 personality trait domains (negative affect, detachment, antagonism, disinhibition, and psychoticism), with each trait domain consisting of 5 items to be rated on a 4-point scale from 0 (“Very false or often false”), 1 (“Sometimes or somewhat false”), 2 (“Sometimes or somewhat true”), to 3 (“Very true or often true”). All items are summed to derive an overall score, with higher scores indicating greater overall personality dysfunction. Similarly, higher scores in each trait domain indicate greater dysfunction in that personality trait domain. In the current sample, the PID-5-BF overall scale has demonstrated excellent internal consistency (alpha = .90).

Therapy and Psychology Questions

At Time 2, Time 3, and Time 4 assessments, students are also asked whether they have attended therapy in the last 3 months (and if yes, for how many sessions), as well as a question regarding their interest in psychology, rated on a 4-point, Likert-type scale from 1 (“None”) to 4 (“A lot”).

Measures Completed at Post-Intervention Only

The following questionnaires are to be completed by students in the preventive intervention group at Time 2 and by students in the WL control group at Time 4 (Table 2).

Satisfaction with the Program Questionnaire

We used 6 of the 7 questions from Rapee and colleagues’ Satisfaction Questionnaire [64]. The 6 questions assess enjoyment of the program, amount learnt in the program, the effectiveness of the program in improving general life coping skills, likelihood of recommending the program to others, and the ability to cope with emotions before and after the program. We added one other question “Did this program help you to learn more about emotions and how they work?” All items are assessed on a 10-point scale from 1 (“Least or none”) to 10 (“A lot”), with the exception of questions related to recommending the program to others and whether this program increased knowledge about emotions and how they work. These latter 2 questions are assessed using a dichotomous, 2-point scale with 1 being “Yes” and 2 being “No.” Rapee and colleagues’ Satisfaction Questionnaire has showed adequate psychometric properties [24,64].

Discipline Problems During Sessions Questionnaire

Students are asked 6 questions about how often they demonstrated certain behaviors during the sessions using a 4-point, Likert-type scale from 1 (“In no or almost no sessions of the program“), 2 (“Only in some sessions of the program”), 3 (“In quite a lot of the sessions of the program”), to 4 (“In all or almost all sessions of the program”). Specifically, the items were (1) “I have talked to my classmates when I should not have”; (2) “I have paid attention to what the girls delivering the program were saying”; (3) “I have done things from other subjects during program sessions”; (4) “I have been reprimanded for my behavior”; (5) “I have taken the program seriously”; and (6) “I have tried to do my best when doing the in-class activities of the program.” In the current sample, this questionnaire has demonstrated only fair internal consistency at Time 2 assessment (alpha = .68).

Other End of Program Questions

Adolescents are also asked the following other questions related to the program: (1) “What did you like best about the program?”; (2) “What did you like worst?”; (3) “Being honest with yourself, are you going to make efforts in the future to apply the strategies that you learned in this program in your daily life?”; (4) “When you missed a session, did you read the summary of the session that was given to you?”; and (5) “When you missed a session, did you do the homework that was given to you?” Questions 1 and 2 are open-choice questions and question 3 is to be answered using a dichotomous scale with 1 being “Probably yes” and 2 being “Probably not.” Questions 4 and 5 are to be rated on a 3-point scale from 1 (”Most of the time yes”), 2 (“Most of the times no”), to 3 (“I did not miss any sessions”).

Measures Completed at Post-Intervention and Follow-Up

The following questionnaires are to be completed by students in the preventive intervention group at Time 2 and Time 3 and by students in the WL control group at Time 4 (Table 2).

Curriculum Knowledge Questionnaire

A questionnaire was created on the basis of the program curriculum to assess participants’ knowledge of core information presented in the program. Specifically, there are 2 open-choice questions (“What are the three parts of an emotion?” and “What can you do when you are feeling sad or down to feel better?”) and one multiple choice question (“What is a thinking trap?”) with 3 answer choices (“It is a kind of unpleasant emotion”/“It is what happens when someone tries to trick us into thinking what they want”/“It is a thought that makes us feel unpleasant emotions”).

Strategies Practiced Outside of Session Questionnaire

We adapted the questionnaire for our study based on the format used in a previous study by Johnson and colleagues [65]. Students are asked “During the 9-week course, how often did you practice each of the following techniques outside of the lessons?”, are supplied with a list of techniques learned during the course, and are asked to rate how much they practiced each technique. Adolescents rate each item on a 5-point, Likert-type scale from 1 (“Never”), 2 (“Once or twice in total”), 3 (“Greater than twice in total but less than once a week”), 4 (“Once or twice each week”), to 5 (“Three times or more each week”), with higher scores indicating more frequent practice of strategies learned in the program. Specifically, participants are asked about

their use of the following strategies: (1) "Identify the three parts of the emotion you are feeling (what you think, what you feel in your body and what you do)"; (2) "Plan for how long are you going to do school work and what pleasant activities are you going to do"; (3) "When you are sad or worried, do something that you like or value even if you do not feel like it"; (4) "Realize that you are falling into a thinking trap (eg, thinking the worst, ignoring the positive, etc) and try to change the thought to an alternative one that makes you feel better"; (5) "When you have a problem, think about all the possible solutions, then think about the good and bad things about each solution and, lastly, choose one of the solutions to try"; (6) "Try to focus in the present moment"; (7) "Meditate, that is, sit and try to focus in your breathing for a few minutes"; and (8) "Expose yourself little by little to those things that scare you or make you nervous because you know it is the only way to overcome your fears." At Time 3, the question is re-worded to: "Since the end of the program at school, how often have you used the following strategies?" In the current sample, this questionnaire has demonstrated adequate internal consistency at Time 2 assessment (alpha = .82).

Sample Size

One significant impact of the adoption of a cluster design is the comparatively large sample size requirement since, in contrast to individually randomized trials where inter-individual variation is the only source of variability, cluster studies involve both variation among individuals and variation among clusters. As a result, cluster studies must recruit a larger number of individuals in order to achieve power equivalent to that of an individually randomized trial [28]. The magnitude of this within-cluster dependence, which ultimately influences the eventual trial size, is quantified by the intracluster correlation coefficient (ICC) [66]. The ICC accounts for the extent to which responses of adolescents attending the same class (that is, sharing the same classroom, classmates, teacher, etc) are more likely to be similar compared with adolescents from a different class.

Power analyses were conducted using G*Power Version 3 software [67]. Sample size required to detect a Cohen *d* effect of .30 was estimated, based upon effect sizes reported in meta-analyses of school-based anxiety and depression preventive interventions conducted all over the world [68,69], including in Spain [70]. Calculations showed that with a power level of .80 and a significance level of alpha = .05, a total sample of $N_{Non\text{-}cluster}$ of 74 was required to detect a significant effect. However, potential loss of power due to data clustering had to be considered in the sample size calculation. The impact of the ICC on the planned trial size depends on the so-called design effect, which can be calculated as 1+(*m*-1) ICC, with *m* referring to the number of participants recruited per cluster [66]. For this study, the ICC is unknown although, on the basis of prior school-based prevention research, the expected ICC for anxiety and depression-related outcomes is approximately 0.02 [27]. The anticipated average class (cluster) size for this study is 27 students meeting inclusion criteria, resulting in a design effect of 1.52. The design effect is then multiplied by the previously calculated $N_{Non\text{-}cluster}$ [66], resulting in an estimated sample size for this trial of 112, considering the design effect. In addition, we estimated a dropout rate of 10% based on previous studies [65,71], resulting in an estimated total sample size of 123 (at least 62 students in each group). The total number of clusters required can be calculated dividing the estimated total sample size by the estimated number of participants recruited per cluster (123/27), resulting in 4.56 clusters [66]. Therefore, to achieve a sample of this size, a total of 157 students and 5 clusters were recruited.

As the flow diagram shows (Figure 1), the final number of participants matches these a priori computations closely: 90 students (10 lost at Time 2) were allocated to the preventive intervention group and 62 students (4 lost at Time 2) were assigned to the WL control group. In addition, this study will allow for calculation of the ICC and effect sizes for all primary and secondary outcomes, which could be very important in assisting future researchers in establishing the feasibility and ideal sample size for future full-scale cluster RCTs applying the UP-A as a preventive intervention.

Randomization

Sequence Generation and Allocation Concealment

Each participating class (cluster) was randomly allocated 1:1 to the preventive intervention or WL control condition. We used a balanced design, resulting in about the same number of classes in each of the preventive intervention and WL control groups (Figure 1). No matching, blocking, or stratification took place. Cluster randomization was undertaken for the ecological validity of providing the intervention at the class level. The randomization was conducted by a researcher not involved in the current project by using a computer random number generator. Random assignment occurred before Time 1 measurements took place because the Research Ethics Committee that provided ethical approval for this study requested that the Informed Consent forms signed by parents/guardians and participants state whether the student was going to be in the experimental or the WL control group.

Implementation and Blinding

The adolescents complete all questionnaires using Qualtrics Survey software in a designated classroom, and a research assistant is available to provide assistance if necessary and to ensure independent responding. The research assistant is blind to the allocated treatment group at time of completing questionnaires to reduce risk of bias. Blinding of participants at the cluster or individual level is not possible for the ethical reasons explained above. Regardless, blinding participants at the cluster or the individual level after baseline would have been impossible due to the nature of the experimental intervention, which requires active participation from the preventive intervention group compared to no involvement or participation from the WL control group in the first phase of the study. An attention-control intervention would have been ideal but was beyond the scope of this study.

Results

Analysis and presentation of data will be in accordance with CONSORT guidelines and, in particular, the extension to cluster randomized trials [28]. Statistical significance will be considered as a *P* value of less than .05, and statistical analysis will be mainly carried out using IBM Statistical Package for the Social Sciences, Version 24.0 (IBM SPSS). Data will be analyzed taking the clustering of students within classes into account.

To check for possible differences at pre-intervention between intervention and WL control groups, a chi-square test for nonparametric variables and a 2-tailed *t* test for continuous variables will be performed at the student-level and class-level on sociodemographic variables as well as on primary and secondary outcome measures. Complete case analyses (excluding participants with missing data), intent-to-treat analyses (including all randomized individuals), and completer status analyses (ie, participants that achieved completer status, that is, participants that attended to 7 or more sessions), will be conducted according to the initial allocation of classes to either preventive intervention or WL control groups. Results will be reported at cluster and individual levels, including information about the estimated effect size and its precision, as well as ICCs for each primary and secondary outcome.

In relation to missing data, we will follow the guidelines for analyzing and reporting cluster RCTs with missing data established by Fiero and et al [72] and Díaz and et al [73]. Specifically, we will (1) attempt to follow up on all randomized clusters and individuals in order to limit the extent of missing data; (2) report the number of clusters and individuals lost to follow-up, as well as numbers of missing values for each variable of interest, which will be important when deciding what analysis to use; (3) collect and report information about reasons for losses to follow-up and other missing values, which may help when deciding the plausibility of the missing-at-random assumption; (4) justify the choice of principal analysis, the missing data mechanism assumed, and the plausibility of these assumptions; (5) perform and report sensitivity analyses to explore the robustness of the trial results to departures from the missing data assumption made in the primary analysis; and (6) follow the CONSORT extension for cluster trials [28] to ensure comprehensive reporting and transparency of methods used. To investigate the potential effects of missing data, the baseline characteristics of the adolescents will be compared for those with and without missing data.

The effectiveness of the intervention will be assessed using complex samples procedures in SPSS as well as multi-level (or hierarchical) modeling that includes both fixed (intervention effects) and random (students in classrooms) effects. Experimental and WL control groups will be compared at 3 points in time: Time 1 (pre-treatment), Time 2 (post-treatment for intervention group), and Time 3 (3-month follow-up for intervention group). In addition, within-participant analyses will be conducted only with the WL control group at Time 4. ICCs will be calculated for all primary and secondary outcomes to compare the variation due to school class and the total variance.

Secondary analyses will include (1) repeating the primary analysis adjusting for any variables exhibiting significant imbalance at baseline to assess whether this influences the findings; (2) examining intervention changes for adolescents who scored above the clinical cut-off for anxiety and/or depression at Time 1; (3) investigating potential predictors of intervention effects including all participants (both intervention and WL control groups) after the WL receives UP-A; (4) comparing observed and expected attrition rates, as well as observed and expected ICCs; and (5) analyzing answers to checklists completed by group leaders.

At the time of submitting this manuscript, all participants have been recruited and we have conducted Time 1 and Time 2 assessments with the experimental and WL control groups. The experimental group has also completed the preventive intervention. We expect the WL control group to complete the preventive intervention and Time 4 assessments in July 2017.

Discussion

Principal Findings

This paper describes the study protocol of a cluster RCT testing the first adaptation of the UP-A as a school-based anxiety and depression preventive intervention. The primary outcomes are change in anxiety and depression symptoms. Secondary outcomes include change in participant-identified top problems ratings, conduct problems, hyperactivity/inattention symptoms, peer problems, prosocial behavior, school grades, depression and anxiety-related interference, positive and negative affect, anxiety sensitivity, emotional avoidance, life satisfaction, quality of life, self-esteem, school adjustment, and discipline problems. This study will also improve our understanding of factors that might influence the effectiveness of UP-A through the preliminary examination of a number of potential predictors of treatment outcomes, including age, gender, number of sessions attended, engagement and effort during sessions, understanding of basic program concepts, adherence to home practice, and practice of specific strategies outside of sessions.

This study focuses on a school-based, universal prevention program for anxiety and depression, which has the advantage of being easily integrated into a school curriculum and thereby limiting the extent to which participants are segregated from peers and at risk for possible stigmatization due to their participation [74]. It is surprising that, despite the high co-occurrence between anxiety and depression and their shared risk and vulnerability factors, preventive interventions have typically focused on either depression or on anxiety alone [10]. For instance, of the 30 total studies included in a meta-analysis conducted by Ahlen and colleagues, 13 studies focused primarily on depression prevention, 10 studies focused primarily on anxiety prevention, and only 7 studies focused on prevention of both disorders [27]. A transdiagnostic approach to preventing anxiety and depression may enhance the efficacy, generalizability, and cost-effectiveness of prevention programs

[10], as well as prevent the development of related emotional disorders.

The location of this research in Spain is particularly significant, as there is a lack of anxiety and depression preventive interventions for youth in the country [70]. A very recent systematic review and meta-analysis on school-based depression and anxiety prevention programs for young people [75] only found 2 studies that took place in Spain, one of them applying a selective preventive intervention and the other applying a universal one. In addition, the current organization of public healthcare in Spain is such that citizens lack direct access to psychological care services due to lengthy waiting lists and a limited number of psychologists [76], although there are some recent initiatives targeting adult populations that aim to implement evidence-based psychological treatments and increase the number of clinical psychologists in primary care [76,77]. Given the high need for mental health services in Spain relative to the number of existing providers, and given the high individual, familial, and societal burden of untreated or undertreated mental health concerns, the development of new and effective preventive intervention programs in the country is crucial.

Strengths

The greatest strength of this study is that it is the first study to evaluate the effectiveness of the UP-A adapted as a school-based, universal anxiety and depression preventive intervention. In addition, it includes random assignment of clusters (classes) to intervention or WL conditions, implements an evidence-based protocol (the UP-A), uses highly reliable assessment measures, and includes a brief but reasonable window for follow-up after prevention programming (3 months). Another strength of this study is that we are not only examining the effectiveness of the UP-A program in reducing anxiety and depressive symptoms, but we are also examining its impact on a range of other variables. These variables include positive outcomes (quality of life, quality of peer relationships, prosocial behavior, self-esteem, satisfaction with life), which are especially important to measure when working at the universal prevention level, as the majority of participants tend to exhibit non-clinical level concerns [71]. We have also included a measure of personality traits, as the presence of certain personality traits has been identified as a potentially useful means of identifying individuals at risk for developing emotional disorders [78]. Furthermore, this study will not only evaluate the effectiveness of the UP-A program for the total sample of participating students, but it will also conduct preliminary analyses to investigate whether the program is more beneficial for certain groups of adolescents than others (eg, those with higher anxiety and depression at pre-treatment) and will examine predictors of treatment efficacy.

Limitations

This study has certain limitations due to its preliminary nature. The major limitation of this study is its moderate sample size (152 adolescents across 5 classes were randomized). In addition, the adolescents were recruited from only one school, which may limit generalizability of findings. Furthermore, due to financial and personnel constraints, this study only incorporated adolescent self-report measures. Although this methodology facilitates assessment of a large cohort of adolescents in a relatively short time frame [9], it also has limitations (eg, relies on subjective perceptions), and future studies should consider using information from multiple sources (interviews with the students, parents, and teacher reports, etc). In addition, participants were not blinded to allocation when measures were obtained due to constraints established by the Ethical Research Committee, which specified that informed consent forms must state group allocation. Lastly, using the same group leaders to administer the intervention for all classes increases consistency but, at the same time, may limit generalizability of findings [65].

Conclusions

The current trial will provide insight into the implementation and effectiveness of the UP-A as a universal, school-based anxiety and depression prevention program, adding knowledge to the research base on transdiagnostic prevention programs more generally. This initial cluster RCT was also designed to provide information regarding the feasibility of potential future full-scale trials in order to optimize the intervention and design approach.

Acknowledgments

This work is supported by the Spanish Ministry of Economy, Industry and Competitiveness (grant #PSI2013-44480-P) to authors BS, PC and RMV and the Spanish Ministry of Education, Culture and Sport (grant #FPU13/03315) to author JGE. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors thank the Secondary School that participated in the study (IES Jaime Vera, Madrid, Spain).

Conflicts of Interest

None declared.

References

1. Fernández-Máximo A, Martínez-Arias R, Rodríguez-Sutil C. Prevalencia de trastornos psicológicos en niños y adolescentes, su relación con la edad y el género. Psicothema 2000;12(4):525-532 [FREE Full text]

2. Merikangas KR, He J, Burstein M, Swanson SA, Avenevoli S, Cui L, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2010 Oct;49(10):980-989 [FREE Full text] [doi: 10.1016/j.jaac.2010.05.017] [Medline: 20855043]
3. Beecham J. Annual research review: child and adolescent mental health interventions: a review of progress in economic studies across different disorders. J Child Psychol Psychiatry 2014 Jun;55(6):714-732 [FREE Full text] [doi: 10.1111/jcpp.12216] [Medline: 24580503]
4. Bodden DH, Dirksen CD, Bögels SM. Societal burden of clinically anxious youth referred for treatment: a cost-of-illness study. J Abnorm Child Psychol 2008 May;36(4):487-497 [FREE Full text] [doi: 10.1007/s10802-007-9194-4] [Medline: 18214667]
5. Lewinsohn PM, Solomon A, Seeley JR, Zeiss A. Clinical implications of "subthreshold" depressive symptoms. J Abnorm Psychol 2000 May;109(2):345-351. [Medline: 10895574]
6. Woodward LJ, Fergusson DM. Life course outcomes of young people with anxiety disorders in adolescence. J Am Acad Child Adolesc Psychiatry 2001 Sep;40(9):1086-1093. [doi: 10.1097/00004583-200109000-00018] [Medline: 11556633]
7. Essau CA. Frequency and patterns of mental health services utilization among adolescents with anxiety and depressive disorders. Depress Anxiety 2005;22(3):130-137. [doi: 10.1002/da.20115] [Medline: 16175563]
8. Mrazek PJ, Haggerty RJ, editors. Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research. Washington, DC: National Academy Press; 1994.
9. Essau CA, Conradt J, Sasagawa S, Ollendick TH. Prevention of anxiety symptoms in children: results from a universal school-based trial. Behav Ther 2012 Jun;43(2):450-464. [doi: 10.1016/j.beth.2011.08.003] [Medline: 22440079]
10. Dozois DJA, Seeds PM, Collins KA. Transdiagnostic approaches to the prevention of depression and anxiety. J Cogn Psychother 2009;23(1):44-59 [FREE Full text]
11. Axelson DA, Birmaher B. Relation between anxiety and depressive disorders in childhood and adolescence. Depress Anxiety 2001;14(2):67-78. [Medline: 11668659]
12. Joiner TE, Catanzaro SJ, Laurent J, Sandín B, Blalock JA. Modelo tripartito sobre el afecto positivo y negativo, la depresión y la ansiedad: Evidencia basada en la estructura de los síntomas y en diferencias sexuales. Rev Psicopatol Psicol Clin 1996;1(1):27-34 [FREE Full text]
13. Ehrenreich-May J, Bilek EL, Queen AH, Hernandez-Rodríguez J. A unified protocol for the group treatment of childhood anxiety and depression. Rev Psicopatol Psicol Clin 2012;17(3):219-236 [FREE Full text]
14. Sandín B, Chorot P, Valiente RM. Transdiagnóstico: Nueva frontera en psicología clínica. Rev Psicopatol Psicol Clin 2012;17(3):185-203 [FREE Full text] [doi: 10.5944/rppc.vol.17.num.3.2012.11839]
15. McEvoy PM, Nathan P, Norton PJ. Efficacy of transdiagnostic treatments: a review of published outcome studies and future research directions. J Cogn Psychother 2009 Feb 01;23(1):20-33 [FREE Full text] [doi: 10.1891/0889-8391.23.1.20]
16. Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. Behav Res Ther 2003 May;41(5):509-528. [Medline: 12711261]
17. Barlow DH, Farchione TJ, Fairholme CP, Ellard KK, Boisseau CL, Allen L, et al. Unified Protocol for Transdiagnostic Treatment of Emotional Disorders: Therapist Guide (Treatments That Work). New York, NY: Oxford University Press, USA; 2010.
18. Ehrenreich-May J, Bilek E, Buzzella B, Kennedy SM, Mash JA, Bennett SM, et al. The unified protocols for transdiagnostic treatment of emotional disorders in children and adolescents: Therapist guide [In Press]. New York, NY: Oxford University Press; 2017.
19. Bilek EL, Ehrenreich-May J. An open trial investigation of a transdiagnostic group treatment for children with anxiety and depressive symptoms. Behav Ther 2012 Dec;43(4):887-897. [doi: 10.1016/j.beth.2012.04.007] [Medline: 23046789]
20. Ehrenreich-May J, Rosenfield D, Queen AH, Kennedy SM, Remmes CS, Barlow DH. An initial waitlist-controlled trial of the unified protocol for the treatment of emotional disorders in adolescents. J Anxiety Disord 2017 Mar;46:46-55. [doi: 10.1016/j.janxdis.2016.10.006] [Medline: 27771133]
21. García-Escalera J, Chorot P, Valiente RM, Reales JM, Sandín B. Efficacy of transdiagnostic cognitive-behavioral therapy for anxiety and depression in adults, children and adolescents: A meta-analysis. Rev Psicopatol Psicol Clin 2016 10;21(3):147-175 [FREE Full text] [doi: 10.5944/rppc.vol.21.num.3.2016.17811]
22. Kendall P, Stark K, Martinsen K, Rodriguez K, Arora P. Group Leader Manual for Emotion: "coping Kids" Managing Anxiety and Depression. Ardmore, PA: Workbook Publishing; 2013.
23. Essau CA, Ollendick TH. The Super Skills for Life Programme. London: University of Roehampton; 2013.
24. Martinsen K, Kendall P, Stark K, Neumer S. Prevention of anxiety and depression in children: acceptability and feasibility of the transdiagnostic EMOTION program. Cogn Behav Pract 2016 Feb;23(1):1-13 [FREE Full text] [doi: 10.1016/j.cbpra.2014.06.005]
25. La Greca AM, Ehrenreich-May J, Mufson L, Chan S. Preventing adolescent social anxiety and depression and reducing peer victimization: intervention development and open trial. Child Youth Care Forum 2016 Dec;45(6):905-926. [doi: 10.1007/s10566-016-9363-0] [Medline: 27857509]

26. Ehrenreich-May J, Bilek EL. Universal prevention of anxiety and depression in a recreational camp setting: an initial open trial. Child Youth Care Forum 2011 Dec;40(6):435-455 [FREE Full text] [doi: 10.1007/s10566-011-9148-4] [Medline: 22977297]
27. Ahlen J, Lenhard F, Ghaderi A. Universal prevention for anxiety and depressive symptoms in children: a meta-analysis of randomized and cluster-randomized trials. J Prim Prev 2015 Dec;36(6):387-403. [doi: 10.1007/s10935-015-0405-4] [Medline: 26502085]
28. Campbell MK, Elbourne DR, Altman DG, CONSORT Group. CONSORT statement: extension to cluster randomised trials. BMJ 2004 Mar 20;328(7441):702-708 [FREE Full text] [doi: 10.1136/bmj.328.7441.702] [Medline: 15031246]
29. Calkins SD. Psychobiological models of adolescent risk: implications for prevention and intervention. Dev Psychobiol 2010 Apr;52(3):213-215. [doi: 10.1002/dev.20435] [Medline: 20213753]
30. Queen AH, Barlow DH, Ehrenreich-May J. The trajectories of adolescent anxiety and depressive symptoms over the course of a transdiagnostic treatment. J Anxiety Disord 2014 Aug;28(6):511-521. [doi: 10.1016/j.janxdis.2014.05.007] [Medline: 24960439]
31. Girio-Herrera E, Ehrenreich-May J. Using flexible clinical processes in the unified protocol for the treatment of emotional disorders in adolescence. Psychotherapy (Chic) 2014 Mar;51(1):117-122. [doi: 10.1037/a0032517] [Medline: 24635001]
32. Sandín B, Chorot P, Valiente RM, Chorpita BF. Development of a 30-item version of the Revised Child Anxiety and Depression Scale. Rev Psicopatol Psicol Clin 2010;15(3):165-178 [FREE Full text] [doi: 10.5944/rppc.vol.15.num.3.2010.4095]
33. Piqueras JA, Martín-Vivar M, Sandin B, San Luis C, Pineda D. The Revised Child Anxiety and Depression Scale: a systematic review and reliability generalization meta-analysis. J Affect Disord 2017 Aug 15;218:153-169. [doi: 10.1016/j.jad.2017.04.022] [Medline: 28475961]
34. Chorot P, Valiente RM, Magaz AM, Santed MA, Sandin B. Perceived parental child rearing and attachment as predictors of anxiety and depressive disorder symptoms in children: the mediational role of attachment. Psychiatry Res 2017 Jul;253:287-295. [doi: 10.1016/j.psychres.2017.04.015] [Medline: 28411577]
35. Sandín B, Chorot P, Valiente RM. TCC de los trastornos de ansiedad: Innovaciones en niños y adolescentes. Madrid: Klinik; 2016.
36. Chorot P, Magaz A, Valiente RM. Estilos de crianza y vulnerabilidad psicopatológica [Parenting styles and vulnerability to psycholopathology]. Presented at: Symposium Nacional de Psicología Clínica y de la Salud en Niños y Adolescentes; 2013; Elche (Spain).
37. Weisz JR, Chorpita BF, Frye A, Ng MY, Lau N, Bearman SK, et al. Youth top problems: Using idiographic, consumer-guided assessment to identify treatment needs and to track change during psychotherapy. J Consult Clin Psychol 2011 Jun;79(3):369-380. [doi: 10.1037/a0023307] [Medline: 21500888]
38. Goodman R. The Strengths and Difficulties Questionnaire: a research note. J Child Psychol Psychiatry 1997 Jul;38(5):581-586. [Medline: 9255702]
39. García P, Goodman R, Mazaria J, Torres A, Rodríguez-Sacristán J, Hervás A. El Cuestionario de Capacidades y Dificultades. Rev Psiquiatr Infanto-Juv 2000;1:12-17.
40. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. J Am Acad Child Adolesc Psychiatry 2001 Nov;40(11):1337-1345. [doi: 10.1097/00004583-200111000-00015] [Medline: 11699809]
41. Gómez-Beneyto M, Nolasco A, Moncho J, Pereyra-Zamora P, Tamayo-Fonseca N, Munarriz M, et al. Psychometric behaviour of the strengths and difficulties questionnaire (SDQ) in the Spanish national health survey 2006. BMC Psychiatry 2013 Mar 22;13:95 [FREE Full text] [doi: 10.1186/1471-244X-13-95] [Medline: 23522343]
42. García-Escalera J, Sandín B, Chorot P, Valiente RM. Escala de Interferencia de la Depresión y Ansiedad para Niños (EIDAN). Madrid, Spain: Universidad Nacional de Educación a Distancia [Unpublished Questionnaire]; 2016.
43. Sandín B. Escalas PANAS de afecto positivo y negativo para niños y adolescentes (PANASN). Rev Psicopatol Psicol Clin 2003;8(2):173-182 [FREE Full text]
44. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [Medline: 3397865]
45. Tortella-Feliu M, Balle M, Sesé A. Relationships between negative affectivity, emotion regulation, anxiety, and depressive symptoms in adolescents as examined through structural equation modeling. J Anxiety Disord 2010 Oct;24(7):686-693. [doi: 10.1016/j.janxdis.2010.04.012] [Medline: 20472383]
46. Silverman WK, Fleisig W, Rabian B, Peterson RA. Childhood Anxiety Sensitivity Index. J Clin Child Adolesc Psychol 1991;20(2):162-168 [FREE Full text]
47. Sandín B. Ansiedad, miedos y fobias en niños y adolescentes. Madrid: Dykinson; 1997.
48. Sandín B, Chorot P, Santed M, Valiente RM. Análisis factorial confirmatorio del Indice de Sensibilidad a la Ansiedad para Niños. Psicothema 2002;14(2):333-339 [FREE Full text]
49. Kennedy SM, Ehrenreich-May J. Assessment of emotional avoidance in adolescents: psychometric properties of a new multidimensional measure. J Psychopathol Behav Assess 2016 Nov 7;39(2):279-290 [FREE Full text] [doi: 10.1007/s10862-016-9581-7]

50. García-Escalera J, Chorot P, Valiente RM, Sandín B, Tonarely N, Ehrenreich-May J. Spanish version of the Emotional Avoidance Strategy Inventory for Adolescents (EASI). Madrid, Spain: Universidad Nacional de Educación a Distancia [Unpublished questionnaire]; 2016.
51. Sandín B, Chorot P, Valiente RM. Escala de Satisfacción con la Vida para Niños y Adolescentes (SWLS-C). Madrid, Spain: Universidad Nacional de Educación a Distancia [Unpublished questionnaire]; 2015.
52. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
53. KIDSCREEN Group. The KIDSCREEN Questionnaires: Quality of Life Questionnaires for Children and Adolescents. Lengerich: Pabst Science Publishers; 2006.
54. Erhart M, Ottova V, Gaspar T, Jericek H, Schnohr C, Alikasifoglu M, et al. Measuring mental health and well-being of school-children in 15 European countries using the KIDSCREEN-10 Index. Int J Public Health 2009;54(2):160-166. [doi: 10.1007/s00038-009-5407-7] [Medline: 19652910]
55. Rosenberg M. Society and the Adolescent Self-Image. Princeton: Princeton University Press; 1965.
56. Sandín B, Valiente RM, Chorot P. Instrumentos para la evaluación del estrés psicosocial. In: Sandín B, editor. El Estrés psicosocial: Conceptos y consecuencias clínicas. Madrid: Klinik; 2008:267-346.
57. Molina J, Chorot P, Valiente RM, Sandín B. Miedo a la evaluación negativa, autoestima y presión psicológica: Efectos sobre el rendimiento deportivo en adolescentes. Cuad Psicol Deporte 2014;14(3):57-66 [FREE Full text]
58. Moral JC, Sánchez-Sosa JC, Villarreal ME. Desarrollo de una escala multidimensional breve de ajuste escolar. REMA 2010;15(1):1-11 [FREE Full text]
59. Sánchez-Sosa JC, Villarreal-González ME, Ávila Guerrero M, Jiménez A, Musitu G. Contextos de socialización y consumo de drogas ilegales en adolescentes escolarizados. Psychosocial Intervention 2014 Apr;23(1):69-78 [FREE Full text] [doi: 10.5093/in2014a7]
60. Martín MB, Pros RC, Busquets CG, Muntada MC. La indisciplina social e instruccional en adolescentes y universitarios: la influencia del género. Rev Electron Investig Psicoeduc Psigopedag 2012;10(27):693-712 [FREE Full text]
61. Gotzens Busquets C, Cladellas Pros R, Clariana Muntada M, Badia Martín M. Indisciplina instruccional y convencional: su predicción en el rendimiento académico. Rev Colomb Psicol 2015 Oct 07;24(2):317-330 [FREE Full text] [doi: 10.15446/rcp.v24n2.44148]
62. American Psychiatric Association. Online Assessment Measures. 2013. The Personality Inventory for DSM-5-Brief Form (PID-5-BF) Child Age 11 to 17 URL: https://www.psychiatry.org/psychiatrists/practice/dsm/educational-resources/assessment-measures [accessed 2016-09-02] [WebCite Cache ID 6qqmwiDKp]
63. Sandín B, Valiente RM, Chorot P, García-Escalera J. Spanish Version of the Personality Inventory for DSM-5 Brief Form (PID-5-BF) Child Age 11-17. Madrid, Spain: Universidad Nacional de Educación a Distancia [Unpublished questionnaire]; 2014.
64. Rapee RM, Wignall A, Sheffield J, Kowalenko N, Davis A, McLoone J, et al. Adolescents' reactions to universal and indicated prevention programs for depression: perceived stigma and consumer satisfaction. Prev Sci 2006 Jun;7(2):167-177. [doi: 10.1007/s11121-006-0035-4] [Medline: 16791521]
65. Johnson C, Burke C, Brinkman S, Wade T. Effectiveness of a school-based mindfulness program for transdiagnostic prevention in young adolescents. Behav Res Ther 2016 Jun;81:1-11. [doi: 10.1016/j.brat.2016.03.002] [Medline: 27054828]
66. Gao F, Earnest A, Matchar DB, Campbell MJ, Machin D. Sample size calculations for the design of cluster randomized trials: a summary of methodology. Contemp Clin Trials 2015 May;42:41-50. [doi: 10.1016/j.cct.2015.02.011] [Medline: 25766887]
67. Faul F, Erdfelder E, Lang A, Buchner AG. 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. [Medline: 17695343]
68. Calear AL, Christensen H. Systematic review of school-based prevention and early intervention programs for depression. J Adolesc 2010 Jun;33(3):429-438. [doi: 10.1016/j.adolescence.2009.07.004] [Medline: 19647310]
69. Corrieri S, Heider D, Conrad I, Blume A, König HH, Riedel-Heller SG. School-based prevention programs for depression and anxiety in adolescence: a systematic review. Health Promot Int 2014 Sep;29(3):427-441. [doi: 10.1093/heapro/dat001] [Medline: 23376883]
70. Sánchez-Hernández Ó, Méndez FX, Garber J. Prevención de la depresión en niños y adolescentes: revisión y reflexión. Rev Psicopatol Psicol Clin 2014;19(1):63-76 [FREE Full text]
71. Gallegos J, Linan-Thompson S, Stark K, Ruvalcaba N. Preventing childhood anxiety and depression: testing the effectiveness of a schoolbased program in Mexico. Psicología Educativa 2013 Jun;19(1):37-44 [FREE Full text] [doi: 10.5093/ed2013a6]
72. Fiero MH, Huang S, Oren E, Bell ML. Statistical analysis and handling of missing data in cluster randomized trials: a systematic review. Trials 2016 Feb 09;17:72 [FREE Full text] [doi: 10.1186/s13063-016-1201-z] [Medline: 26862034]
73. Díaz-Ordaz K, Kenward MG, Cohen A, Coleman CL, Eldridge S. Are missing data adequately handled in cluster randomised trials? A systematic review and guidelines. Clin Trials 2014 Oct;11(5):590-600. [doi: 10.1177/1740774514537136] [Medline: 24902924]

74. Nehmy TJ, Wade TD. Reduction in the prospective incidence of adolescent psychopathology: a review of school-based prevention approaches. Mental Health & Prevention 2014 Dec;2(3-4):66-79 [FREE Full text] [doi: 10.1016/j.mhp.2014.11.002]
75. Werner-Seidler A, Perry Y, Calear AL, Newby JM, Christensen H. School-based depression and anxiety prevention programs for young people: a systematic review and meta-analysis. Clin Psychol Rev 2017 Feb;51:30-47 [FREE Full text] [doi: 10.1016/j.cpr.2016.10.005] [Medline: 27821267]
76. Cano-Vindel A, Muñoz-Navarro R, Wood CM, Limonero JT, Medrano LA, Ruiz-Rodríguez P, et al. Transdiagnostic cognitive behavioral therapy versus treatment as usual in adult patients with emotional disorders in the primary care setting (PsicAP Study): protocol for a randomized controlled trial. JMIR Res Protoc 2016 Dec 23;5(4):e246 [FREE Full text] [doi: 10.2196/resprot.6351] [Medline: 28011446]
77. Montero-Marín J, Araya R, Pérez-Yus MC, Mayoral F, Gili M, Botella C, et al. An internet-based intervention for depression in primary care in Spain: a randomized controlled trial. J Med Internet Res 2016 Aug 26;18(8):e231 [FREE Full text] [doi: 10.2196/jmir.5695] [Medline: 27565118]
78. Klein DN, Kotov R, Bufferd SJ. Personality and depression: explanatory models and review of the evidence. Annu Rev Clin Psychol 2011;7:269-295 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032210-104540] [Medline: 21166535]

Abbreviations

CASI: Childhood Anxiety Sensitivity Index
CBT: cognitive behavioral therapy
CDN: Depression Questionnaire for Children (Cuestionario de Depresión para Niños)
CONSORT: Consolidated Standards of Reporting Trials
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th edition
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th edition
EAN: Anxiety Scale for Children (Escala de Ansiedad para Niños)
EASI-A: Emotional Avoidance Strategy Inventory for Adolescents
EBAE-10: School Adjustment Brief Scale (Escala Breve de Ajuste Escolar)
EIDAN: Depression and Anxiety Interference Scale for Children (Escala de Interferencia de la Depresión y Ansiedad para Niños)
ICC: intracluster correlation coefficient
IG: General Indiscipline Scale (Escala de Indisciplina General)
KIDSCREEN-10: Kidscreen-10 Quality of Life Scale
PANASN: Positive and Negative Affect Schedule for Children and Adolescents (Escalas PANAS de Afecto Positivo y Negativo para Niños y Adolescentes)
PID-5-BF: The Personality Inventory for DSM-5-Brief Form
RCT: randomized controlled trial
RCADS-30: Revised Child Anxiety and Depression Scale–30
SDQ: Strengths and Difficulties Questionnaire
SES: Self-Esteem Scale
SMART: specific, measurable, attainable, relevant, and time-bound
SPSS: Statistical Package for the Social Sciences
SWLS-C: Satisfaction with Life Scale for Children
T-CBT: transdiagnostic cognitive behavioral therapy
TPA: Top Problems Assessment
UP: Unified Protocol for Transdiagnostic Treatment of Emotional Disorders
UP-A: Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Adolescents
UP-C: Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Children
WL: waitlist

Edited by G Eysenbach; submitted 27.04.17; peer-reviewed by T Ollendick, X Méndez; comments to author 17.05.17; revised version received 30.05.17; accepted 31.05.17; published 21.08.17.

Please cite as:
García-Escalera J, Valiente RM, Chorot P, Ehrenreich-May J, Kennedy SM, Sandín B
The Spanish Version of the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders in Adolescents (UP-A) Adapted as a School-Based Anxiety and Depression Prevention Program: Study Protocol for a Cluster Randomized Controlled Trial
JMIR Res Protoc 2017;6(8):e149
URL: http://www.researchprotocols.org/2017/8/e149/
doi:10.2196/resprot.7934
PMID:28827212

Protocol

Ivermectin Treatment in Patients With Onchocerciasis-Associated Epilepsy: Protocol of a Randomized Clinical Trial

Robert Colebunders[1], MD, PhD; Michel Mandro[2], MD; Deby Mukendi[3], MD; Housseini Dolo[1], MD; Patrick Suykerbuyk[1], PhD; Marieke Van Oijen[1], MD, PhD

[1]University of Antwerp, Global Health Institute, Antwerp, Belgium

[2]Provincial Ministry of Health, Bunia, The Democratic Republic Of The Congo

[3]Centre Neuro Psycho Pathologique, University of Kinshasa, Kinshasa, The Democratic Republic Of The Congo

Corresponding Author:
Robert Colebunders, MD, PhD
University of Antwerp
Global Health Institute
Doornstraat 331
Antwerp,
Belgium
Phone: 32 486 92 01 49
Fax: 32 32 651 836
Email: robert.colebunders@uantwerpen.be

Abstract

Background: Many studies have reported an association between epilepsy, nodding syndrome (NS), and onchocerciasis (river blindness). A high prevalence of epilepsy has been noted particularly in onchocerciasis hyperendemic areas where onchocerciasis is not or insufficiently controlled with mass ivermectin distribution. There is evidence that increasing the coverage of ivermectin reduces the incidence of epilepsy, and anecdotal evidence suggests a reduction in seizure frequency in onchocerciasis-associated epilepsy (OAE) patients who receive ivermectin. Finding an alternative treatment for epilepsy in these patients will have major consequences.

Objective: The goal of the study is to assess whether ivermectin treatment decreases the frequency of seizures and leads to seizure freedom in OAE patients, including patients with NS. If we are able to demonstrate such an effect, this would strengthen the argument that onchocerciasis is causing epilepsy and therefore we should increase our efforts to eliminate onchocerciasis.

Methods: We will conduct a randomized clinical trial in the Democratic Republic of Congo to compare seizure freedom in onchocerciasis-infested epilepsy patients who receive immediate ivermectin treatment with delayed (after 4 months) ivermectin treatment. All participants will simultaneously receive antiepilepsy drugs (AEDs) according to local guidelines for epilepsy treatment. The primary endpoint is seizure freedom defined as no seizures during the 4 month of follow-up. Secondary endpoint is significant (>50%) seizure reduction compared to baseline seizure frequency. Reduction of seizures will be compared between ivermectin and nonivermectin arms.

Results: Start of enrollment is planned for August 2017, and we expect to have enrolled all 110 participants by December 2017. Results are expected in June 2018.

Conclusions: If ivermectin treatment in addition to AEDs is able to lead to seizure freedom or significantly reduces seizure frequency in OAE patients, this will have major consequences for epilepsy treatment in onchocerciasis-endemic regions. Ivermectin is donated for free and in non Loa-Loa–endemic regions has negligible side effects. Reducing the burden of epilepsy will have a major impact on quality of life and socioeconomic status of families with affected members in Africa.

Trial Registration: ClinicalTrials.gov NCT03052998; https://clinicaltrials.gov/ct2/show/NCT03052998 (Archived by WebCite at http://www.webcitation.org/6roFVQSG0)

(JMIR Res Protoc 2017;6(8):e137) doi:10.2196/resprot.7186

KEYWORDS

onchocerciasis; epilepsy; nodding syndrome; ivermectin; randomized clinical trial; Democratic Republic of the Congo

Introduction

Onchocerciasis-Associated Epilepsy

Many studies have reported an association between epilepsy, nodding syndrome (NS), and onchocerciasis [1-4]. A meta-analysis of African population-based surveys showed a variation in epilepsy prevalence consistent with onchocerciasis prevalence, with epilepsy prevalence being increased, on average, by 0.4% for each 10% increase in onchocerciasis prevalence [2]. NS is an epileptic disorder occurring in children in onchocerciasis (river blindness) endemic regions, initially only observed in South Sudan, Uganda, and Tanzania [5]. NS occurs in previously healthy children, aged mainly from 5 to 18 years and is characterized by head-nodding, an atonic epileptic seizure [5,6]. Individuals may also develop other types of seizures and stunted growth. Many affected children not only suffer from recurrent seizures but also from severe cognitive, behavioral, and psychiatric problems [7,8]. NS should be considered part of a spectrum of onchocerciasis-associated epilepsy (OAE) disorders [9]. We recently suggested that these epileptic disorders share etiological factors related to *Onchocerca volvulus* infection and therefore considered ivermectin, used to treat onchocerciasis, as a treatment option for OAE [10].

Ivermectin Treatment May Prevent Onchocerciasis-Associated Epilepsy

Ivermectin is effective against many types of parasites and is recommended every 6 months for treatment of onchocerciasis. Medication is taken orally and generally well tolerated [11]. Accumulating evidence is suggesting that ivermectin may prevent OAE. After the introduction of biannual mass treatment with ivermectin in 2012, no new NS cases have been observed in Northern Uganda [12]. In a recent age-matched case control study in Titule, Bas Uélé, in the Democratic Republic of Congo (DRC), 16/18 (89%) of patients with epilepsy had not taken ivermectin the year before they developed epilepsy, compared to 7/18 (39%) controls (P=.002) [13]. We recently visited several villages in the Mbam valley (an onchocerciasis-endemic area in Cameroon) where the prevalence of epilepsy in 1992 ranged between 4% and 8% and most epilepsy cases occurred in people aged younger than 20 years [3]. In 2015, after more than 15 years of administering ivermectin annually, nearly all epilepsy cases were in people over the age of 20 years, also suggesting a protective effect of ivermectin on the incidence of epilepsy (R Colebunders, personal communication). A similar age shift of epilepsy cases to older age groups after the introduction of ivermectin was reported in an onchocerciasis-endemic region in Nigeria [14]. With ivermectin distribution only once a year, it may take many years with high treatment coverage before a significant effect on the incidence of OAE is observed. Indeed, several months after the administration of one dose of ivermectin, microfilariae reappear and therefore create a risk for the infected individual to develop epilepsy. For this reason, in order to decrease the incidence of OAE more rapidly, at least biannual intake of ivermectin may be required.

Pathophysiological Mechanism of Onchocerciasis-Associated Epilepsy

The pathophysiological mechanism of OAE remains unclear. In recent studies, *O volvulus* DNA was never isolated from cerebrospinal fluid (CSF) in patients with NS [15]. However, several of the patients enrolled in these studies had taken ivermectin in the past. In 1938, before the use of ivermectin, microfilariae in CSF were described by Casis Sacre [16] in Mexican patients with onchocerciasis with NS and Nakalanga-like clinical manifestations. Dead and live microfilariae were also found in 1959 by Mazotti [17] in CSF of patients with onchocerciasis treated with diethylcarbamazepine. In 1976, Duke et al [18] noted the presence of small numbers of *O volvulus* microfilariae in the CSF (<2 mf/mL) in 5 of 8 untreated heavily infected (>100 mf/mg skin) onchocerciasis patients. During diethylcarbamazine treatment, in 10 out of 11 heavily infected patients presenting with an ocular form of onchocerciasis, the numbers of *O volvulus* microfilariae in the CSF increased even up to 8 to 31 mf/mL [18]. Patients with *O volvulus* infection receiving diethylcarbamazine were reported to develop optic atrophy, probably because of a Mazotti reaction caused by dead microfilariae present in the optic nerve [19]. Perhaps this nerve could be an entry point of the microfilariae to the brain. Another explanation for the association between onchocerciasis and epilepsy could be the occurrence of an autoimmune response related to the *O volvulus* infection [20] leading to seizures.

Ivermectin acts on the chloride-dependent channels of both glutamate and γ-aminobutyric acid, interrupting neurotransmission in invertebrates. However, in humans, several mechanisms of brain protection exist, including P-glycoprotein, present on the apical face of endothelial cells of the blood-brain barrier and coded by the MDR1 gene [21]. Therefore in humans ivermectin generally does not enter the brain [21] and would not be able to kill microfilariae inside the brain. Ivermectin could, however, reduce the microfilariae load outside the brain and either reduce the risk that additional microfilariae penetrate the brain or reduce the neurotoxic immunological response caused by the microfilarae.

One dose of ivermectin eliminates microfilarae very rapidly [22]. A mathematical model predicted that microfilaridermia would be reduced by half 24 hours after the intake of ivermectin [23]. Therefore, if the microfilariae load plays an important role in causing OAE, it may be that ivermectin also has a rapid effect on the frequency of seizures.

Ivermectin Treatment May Decrease Seizure Frequency in Patients With Onchocerciasis-Associated Epilepsy

In a study in Kabarole district in Uganda in 1992, 34/91 (37%) patients reported some decrease in either the frequency or severity of seizures after 1 dose of ivermectin (150 μg/kg) [24]. After being treated with ivermectin, 13/91 (14%) individuals had no seizures for 3.7 months (on average). Seizures were unchanged in 51/91 (56%), and worsened in 6/91 (7%) [24]. In a recent trial in the DRC comparing moxidectin (an anti- *O volvulus* experimental drug with a longer half-life than ivermectin) with ivermectin, 6 (80%) out of 7 *O*

volvulus-infested patients with epilepsy became seizure free after treatment with moxidectin or ivermectin (the randomization code has not been broken yet). In one person, seizure frequency was significantly reduced over the 18-month follow-up period. In this person, microfilariae remained present in skin snips, although at a lower level than before the onchocerciasis treatment. In all subjects who became seizure free, the skin snips too became microfilaria free [M Mandro, unpublished].

To assess whether ivermectin treatment may reduce the frequency of seizures and lead to seizure freedom, we intend to conduct a short proof-of-concept randomized clinical trial to compare immediate ivermectin treatment with delayed (after 4 months) ivermectin treatment in onchocerciasis-infested persons with epilepsy. The primary outcome is seizure freedom at 4 months. Reducing the burden of epilepsy will have a major impact on quality of life and socioeconomic status of families with affected members in Africa. If we are able to demonstrate an effect of ivermectin on the frequency of seizures, this would be an extra argument that onchocerciasis is causing epilepsy and that therefore we should increase our efforts to eliminate onchocerciasis.

Methods

Study Design

This is a proof-of-concept randomized treatment trial consisting of 2 treatment arms, immediate (arm A) and delayed (4 months later) ivermectin treatment (arm B). A computer-based, preplanned age and frequency of seizures stratified randomization list will be used. The trial will not be placebo-controlled as this will be costly and we do not expect this to greatly influence reporting of seizures, our primary outcome. Epilepsy patients in both arms will additionally receive antiepileptic drugs (AEDs) following local guidelines of epilepsy treatment in DRC. Study investigators collecting and analyzing data will be blinded for treatment.

Study Population, Setting, and Enrollment

The trial will take place in selected villages in the Logo health zone, an onchocerciasis-endemic region in the Ituri province in the DRC, in areas where so far mass ivermectin administration has not been implemented but where the national onchocerciasis program is planning to start community-directed distribution of ivermectin in 2017. Pilot studies by our group in this area show prevalence estimates of OAE of approximately 5% [25].

Before starting, the village chief and community health workers will be informed of the purpose and specifics of our study. When permission is obtained, our medical team will visit the village and set up a mobile clinic. Patients who may fulfill eligibility criteria are selected and examined for a screening visit after informed consent is obtained from patient and/or caretaker in the local language (Alur). Patients who meet the enrollment criteria (see Textbox 1) are invited to participate in the treatment trial. Detailed information about the trial is given in local language, and enrollment informed consent is obtained.

Our study population consists of epilepsy patients aged 5 years and older with onchocerciasis and without other obvious risk factors for epilepsy.

Trial Registration

The trial is registered at ClinicalTrials.gov [NCT03052998] and will be registered at the World Health Organization International Registry Network.

Screening of Epilepsy Patients for Onchocerca volvulus Infection

After informed consent is obtained, trial candidates with epilepsy will be tested for the presence of antibodies to the parasite antigen OV16 with the SD BIOLINE Onchocerciasis IgG4 Rapid Test (Standard Diagnostics Inc). Also, a skin snip will be taken from the left and right iliac crests with a Holtz corneoscleral punch (2 mm). One sterilized punch will be used per subject. Each snip will be weighed on an analytical balance and incubated for 24 hours in isotonic saline in a well of a flat-bottomed microtiter plate. The microfilariae that emerge will be counted using an inverted microscope. The number of microfilariae in each well and the weight of the associated skin snip before incubation will be recorded. The mean of the skin microfilarial (mf) density (mf count/weight) of snip across all 4 skin snips will be calculated and recorded as mf/mg. Skin biopsies will then be stored in 90% ethanol to be tested for *O volvulus* by an in-house polymerase chain reaction method (supplemental methods).

Diagnosis of Onchocerca volvulus Infection

Diagnosis of onchocerciasis infection will be made when microfilariae are detected in skin snip and/or antibodies to the parasite antigen OV16 are detected.

Diagnosis of Epilepsy

To identify eligible epilepsy patients with generalized tonic clonic seizures, we will use a 10-item epilepsy questionnaire previously used in international epilepsy prevalence studies [26,27]. A person identified on the basis of this questionnaire as possibly a person with epilepsy will be examined by a neurologist to make an accurate diagnosis according to definitions proposed by the International League Against Epilepsy [28]. A case of epilepsy will be defined as a patient who has had (1) at least 2 times, unprovoked and without fever, lost consciousness with convulsions with a minimal time difference of 24 hours between the 2 events or (2) 1 unprovoked seizure and a probability of future seizures similar to the general recurrence risk after 2 unprovoked seizures. Detailed questions including the age at seizure onset, seizure frequency, and family history of seizures are part of the baseline questionnaire. The eligibility criteria are listed in Textbox 1.

Textbox 1. Selection criteria.

Inclusion criteria:

- Age 5 years and older
- Signed informed consent form
- Normal neurological development until onset of epilepsy
- Onset of epilepsy between ages of 5 and 18 years
- Seizure frequency of ≥2 seizures per month
- Presence of microfilariae in skin snip and/or antibodies against OV16
- Generalized (primarily or secondarily) tonic clonic seizures are present

Exclusion criteria:

- Ivermectin intake in the last 9 months
- Pregnancy or breastfeeding
- Known or suspected allergy to ivermectin
- *Loa Loa* microfilariae in blood[a]
- Epilepsy with known cause (eg, severe head trauma, perinatal asphyxia, patients with a history of cerebral malaria, meningitis, or encephalitis)
- Concomitant acute illness or chronic medication use
- Chronic alcohol or substance use
- Use of antiepilepsy drugs in the past 2 weeks

[a]Treatment with ivermectin in persons with *Loa Loa* co-infection may cause fatal encephalopathy.

Blood samples will be tested for *Taenia solium* antibodies and antigen, but in the absence of a point-of-care test, results will only become available after screening. Therefore these test results are not part of the eligibility criteria. However, they will be taken into account in analyzing the results.

Ivermectin Treatment Strategy

Following recommendations, 1 dose of 150 µg/kg ivermectin (Mectizan) will be administered orally, and treatment will be directly observed. Ivermectin is generally well tolerated. Common side effects of ivermectin include fever, itching, skin rash, edema, myalgia, and headache.

Anti-Epileptic Drug Treatment Strategy

We designed a standard epilepsy treatment protocol for both treatment arms. To summarize, we will start with phenobarbital 100 mg once a day which may be increased to 150 to 200 mg per day after 2 months if there is insufficient seizure reduction (less than 50% reduction of seizure frequency). If there are contraindications for use of phenobarbital (intellectual or behavioral disorders) or persistent side effects, carbamazepine will be prescribed (in adults initial dose 100 to 200 mg per day, maintenance dose 400 to 1400 mg; in children initial dose is 5 mg/kg per day, maintenance dose 10 to 30 mg/kg per day). In the case of side effects related to carbamazepine, we will start with sodium valproate (in adults initial dose 400 mg per day, maintenance dose 400 to 2000 mg per day; in children initial dose 15 to 20 mg/kg per day and maintenance dose 15 to 30 mg/kg per day). In the case of carbamazepine and sodium valproate, we will prescribe initial dose in all patients and will increase this to the lowest maintenance dose at 2-week visit. In case of dose-determined side effects, dose reduction is permitted. Dose may be increased at the 2-month visit if there is insufficient seizure control (less than 50% reduction in seizure frequency). Individual treatment decisions will be made by the team physician who has received specific training in epilepsy treatment and can consult with the team neurologist. Patient and family will be informed of the following regarding epilepsy treatment [29]:

- Delay in onset of effect and the time course of treatment
- Potential side effects and the risk of these symptoms (seek help promptly if these are distressing)
- Risk of abrupt discontinuation/withdrawal symptoms on missing doses
- Need for regular follow-up

Compliance

We will perform indirect (pill count) and direct (AED blood levels) methods to check for compliance to AED treatment. We will train community volunteers to assist the research team and local health team with the follow-up of the trial participants and for compliance monitoring at the home of the participants. The center visit at week 2 is scheduled to check for side effects in order to minimize withdrawal from AED treatment. Although noncompliant patients are expected to be equally distributed among treatment arms because of randomization and therefore not influence outcome, it is important to put effort into minimizing noncompliance, especially since we intend to continue treatment with AEDs past the duration of the trial regardless of the results. Community volunteers will also be trained to become community directed distributors of ivermectin after completion of the trial.

Endpoints

Primary endpoint is proportion of patients who have achieved seizure freedom after 4 months. Seizure freedom will be defined as no seizures the last month of the trial (month 4). Secondary endpoints are proportion of patients at month 4 with more than 50% reduction in seizure frequency compared to reported seizure frequency at randomization and microfilarial load measured in skin snip. The seizure frequency data will be collected starting from day 1 using a seizure diary. Reduction of seizures will be compared between ivermectin and nonivermectin arms.

Baseline and Follow-Up Procedures

At baseline, information will be collected on seizure semiology, frequency, risk factors, treatment history, and ivermectin treatment in the past. A full physical and neurological examination will be performed together with serological testing and skin snip examination. Weight and height measurements will be carried out and the participant body mass index will be calculated at baseline and follow-up visits. Trial participants will be instructed on how to fill out a seizure calendar and record intake of AEDs. A center visit is scheduled 2 weeks after randomization to assess potential side effects of AEDs. Side effects will be recorded using a structured questionnaire inquiring about known side effects of phenobarbital, carbamazepine, or sodium valproate. To assess seizure frequency, center visits are scheduled after 2 weeks, 1, 2, 3, and 4 months. To minimize loss to follow-up, we use Global Positioning System coordinates to locate study participants. During these visits, neurological and physical examination will be repeated, adverse events will be evaluated, and we will count AED pills for indirect measurement of AED compliance. At the fourth visit, skin snip examination and *O volvulus* serology will be repeated. Additionally, AED blood levels will be measured to directly assess AED compliance (Table 1).

Table 1. Overview of study baseline and follow-up assessments.

Overview procedures	Trial stage						
	t=–1	t=0	t=2w	t=m1	t=m2	t=m3	t=m4
Check of inclusion and exclusion criteria	X						
Patient characteristics questionnaire	X						
Physical exam and neurological assesment	X		X		X	X	X
Seizure diary review	X		X	X	X	X	X
Administering AED[a]/ivermectin		X	X	X	X	X	X
AED pill count	X		X	X	X	X	X
Blood exam, skin snip, urine test	X						X
Adverse event questionnaire		X	X	X	X	X	X

[a]AED: antiepilepsy drug.

If a participant is not able to visit the study center, a home visit will be performed to monitor AED use and seizure frequency.

Sample Size Calculation

We expect that 4 months of treatment with AEDs (phenobarbital, carbamazepine, or sodium valproate) will lead to seizure freedom in 50% of the patients (experience of R Idro in Uganda). In a clinical trial performed in Rethy (Ituri) comparing the safety and parasitological efficacy of moxidectin versus ivermectin treatment in persons with *O volvulus* infection not receiving antiepileptic treatment, 6 (80%) of 7 trial participants with epilepsy were seizure free at 4 months.

Null hypothesis: The probability of seizure freedom at 4 months for immediate ivermectin treatment is equal to the probability of seizure freedom at 4 months for delayed ivermectin treatment. If we expect that seizure freedom at 4 months will be obtained in 50% of the participants with phenobarbital alone and that with additional ivermectin treatment 80% of patients will achieve seizure freedom at 4 months, about 104 subjects (52 per group) are needed to achieve the power of 90% to reject the null hypothesis at the 5% significance level. Considering that 5% of the patients will be lost to follow-up, 110 patients will be enrolled in the trial.

All comparative analyses will be based on the intention-to-treat principle: all randomized patients will be included in the analysis according to the result of the randomization.

For the primary endpoint, the null hypothesis will be tested by comparing the observed proportion of responses in arm A with the corresponding proportion in arm B at the 1-sided 5% significance level by using the Cochran-Mantel-Haenszel test for comparison of 2 independent proportions. The Cochran-Mantel-Haenszel test will be performed controlling for baseline frequency of seizures. The same test will be used for the secondary endpoint. Change versus baseline in skin micofilarial load at month 4 will be analyzed by means of a *t* test. Frequencies of seizures will be compared between participants with and without positive microfilaria skin snips at month 4.

Patients lost to follow-up will be regarded as nonresponders. Similarly, patients for whom the AED regimen had to be adapted because of an increasing number of seizures will be considered as nonresponders. AED treatment changes because of side effects or possible interactions with other drugs will not be considered as treatment failure.

Data Handling and Record Keeping

All relevant clinical information will be collected on tablets. The identity and information of trial participants is kept confidential. Data will be entered in a Web-based electronic database compliant with Good Clinical Practice as defined by the International Conference on Harmonisation that is access-controlled and anonymized.

Monitoring, Oversight, and Reporting

The trial sponsor is the University of Antwerp. The study team will undergo Good Clinical and Laboratory Practice protocol training and training in special procedures. An independent experienced clinical trial monitor will monitor the trial and report to the sponsor. The monitoring will include checking the consent procedure, clinical event reporting, compliance with protocol standard operating procedures, and treatment adherence. Data queries will be handled according to a quality management plan. A Data Safety Monitoring Board will be established to review safety but not for efficacy, as early-stopping for efficacy is not considered. All adverse study drug reactions, serious adverse events, and deaths will be reported to the sponsor.

Ethics and Dissemination

The trial will be conducted in accordance with applicable laws and regulations including but not limited to the International Conference on Harmonisation guideline for Good Clinical Practice and ethical principles that have their origins in the Declaration of Helsinki. The purpose and nature of the investigation were explained to participants or parents/guardians, including risks and benefits of each procedure. If a subject wishes to participate in the study, the subject and (if applicable) parent/legal guardian must participate in the informed consent process and an informed consent/assent form must be signed or thumb-printed and dated by the subject (and parent/legal guardian, if applicable) or by his or her legally authorized representative, a literate witness (for illiterate subjects), and by the principal investigator or designee before any protocol-required procedures were performed. Ethical approval will be obtained from the ethical committee of the School of Public Health of the University of Kinshasa in the DRC and the ethical committee of the University of Antwerp.

Study results will be discussed with all stakeholders involved. The scientific community will be reached through publication in peer-reviewed open source international scientific journals and presentations at national and international scientific symposia. We will share both the overview of the research as well as the source data, inviting others to analyze and comment on the data and create their own analysis.

In collaboration with the nongovernmental organization Malteser International, a decentralized program to treat all persons with epilepsy in the villages where trial participants are recruited will be set up. Local health care workers have already been trained by 2 neurologists, and AEDs have been ordered. Malteser international has promised a sustainable provision of AEDs after the trial. Monitoring of the AED treatment program for at least 18 months is also planned.

Results

We expect to start inclusion of patients in August 2017, and enrollment to be completed in December 2017. First results of the trial are expected by June 2018.

Discussion

The burden of OAE in African communities is currently being investigated but is expected to be high. It has been estimated that at least 100,000 people may be affected by OAE. Current treatment with locally available AEDs is a challenge because of factors such as availability, noncompliance because of side effects, need for regular follow-up visits, daily administration, and misconceptions about the origin of seizures that prevent patients from seeking medical attention. If ivermectin treatment, potentially in addition to AEDs, is able to lead to seizure freedom or significantly reduces seizure frequency in OAE patients, this will have major consequences for epilepsy treatment in onchocerciasis-endemic regions. Ivermectin is donated for free through the Mectizan donation program [30] and is administered annually or biannually to populations at risk for onchocerciasis. Ivermectin, however, does not kill the adult worm; it only decreases its fertility [31]. After 3 to 6 months of embryostasis, the production of microfilariae resumes [32]. Therefore its effect on the frequency of seizures may also decrease over time. If this 4-month proof-of-concept trial shows a beneficial effect of ivermectin, a trial comparing a dosing of ivermectin every 3 months compared with an annual or biannual dosing needs to be considered. In a previous study, a dosing every 3 months of ivermectin caused a 30% decrease in microfilariae production [33] and was not associated with more side effects than annual dosing of ivermectin [34]. Ultimately, we aim to design a treatment strategy for OAE that can be sustained by the community. Reducing the burden of epilepsy will have a major impact on quality of life and socioeconomic status of families with affected members in Africa.

Acknowledgments

We thank Tony Vangeneugden for statistical advice, Malteser International for organizing the treatment for persons with epilepsy in the Logo health zone, and Deby Mukendi and Eric De Smet for organizing the epilepsy training of local health care workers. The trial is funded by a European Research Council grant (671055) to Robert Colebunders.

Conflicts of Interest

None declared.

References

1. Pion SDS, Boussinesq M. Significant association between epilepsy and presence of onchocercal nodules: case-control study in Cameroon. Am J Trop Med Hyg 2012 Mar;86(3):557 [FREE Full text] [doi: 10.4269/ajtmh.2012.11-0603a] [Medline: 22403337]
2. Pion SDS, Kaiser C, Boutros-Toni F, Cournil A, Taylor MM, Meredith SEO, et al. Epilepsy in onchocerciasis endemic areas: systematic review and meta-analysis of population-based surveys. PLoS Negl Trop Dis 2009 Jun 16;3(6):e461 [FREE Full text] [doi: 10.1371/journal.pntd.0000461] [Medline: 19529767]
3. Boussinesq M, Pion SDS, Demanga N, Kamgno J. Relationship between onchocerciasis and epilepsy: a matched case-control study in the Mbam Valley, Republic of Cameroon. Trans R Soc Trop Med Hyg 2002;96(5):537-541. [Medline: 12474484]
4. Tumwine JK, Vandemaele K, Chungong S, Richer M, Anker M, Ayana Y, et al. Clinical and epidemiologic characteristics of nodding syndrome in Mundri County, southern Sudan. Afr Health Sci 2012 Sep;12(3):242-248. [Medline: 23382736]
5. Dowell S, Sejvar J, Riek L, Vandemaele K, Lamunu M, Kuesel A. Nodding syndrome. Emerg Infect Dis 2013;19(9):1374-1384. [Medline: 23965548]
6. Sejvar J, Kakooza A, Foltz J, Makumbi I, Atai-Omoruto A, Malimbo M. Clinical, neurological, and electrophysiological features of nodding syndrome in Kitgum, Uganda: an observational case series. Lancet Neurol 2013;12(2):166-174. [Medline: 23305742]
7. Musisi S, Akena D, Nakimuli-Mpungu E, Abbo C, Okello J. Neuropsychiatric perspectives on nodding syndrome in northern Uganda: a case series study and a review of the literature. Afr Health Sci 2013;13(2):205-218. [Medline: 24235916]
8. Idro R, Opoka R, Aanyu H, Kakooza-Mwesige A, Piloya-Were T, Namusoke H. Nodding syndrome in Ugandan children—clinical features, brain imaging and complications: a case series. BMJ Open 2013;3(5). [Medline: 23645924]
9. Wamala J, Malimbo M, Tepage F, Lukwago L, Okot C, Cannon R. Nodding syndrome may be only the ears of the hippo. PLoS Negl Trop Dis 2015;9(8):e0003880. [Medline: 26270048]
10. Colebunders R, Hendy A, van Oijen M. Nodding syndrome in onchocerciasis endemic areas. Trends Parasitol 2016;32(8):581-583.
11. González P, González FA, Ueno K. Ivermectin in human medicine: an overview of the current status of its clinical applications. Curr Pharm Biotechnol 2012 May;13(6):1103-1109. [Medline: 22039800]
12. Colebunders R, Post R, O'Neill S, Haesaert G, Opar B, Lakwo T. Nodding syndrome since 2012: recent progress, challenges and recommendations for future research. Trop Med Int Health 2015;20(2):194-200. [Medline: 25348848]
13. Colebunders R, Mandro M, Mokili J, Mucinya G, Mambandu G, Pfarr K. Risk factors for epilepsy in Bas-Uele Province, Democratic Republic of the Congo: a case-control study. Int J Infect Dis 2016;49:1-8. [Medline: 27210267]
14. Dozie I, Onwuliri C, Nwoke B, Chukwuocha U, Chikwendu C, Okoro I. Onchocerciasis and epilepsy in parts of the Imo river basin, Nigeria: a preliminary report. Public Health 2006;120(5):448-450. [Medline: 16524608]
15. Winkler AS, Friedrich K, Velicheti S, Dharsee J, König R, Nassri A, et al. MRI findings in people with epilepsy and nodding syndrome in an area endemic for onchocerciasis: an observational study. Afr Health Sci 2013 Jun;13(2):529-540 [FREE Full text] [doi: 10.4314/ahs.v13i2.51] [Medline: 24235964]
16. Casis SG. [El sindrome epileptico y sus reaciones con onchocercosis]. Boletin de Salubridad e Higiene 1938;1:11-31.
17. Mazotti L. [Presencia de microfilarias de onchocerca volvulus en el liquido cephaloraquideo de enfermos tratados con hetrazan]. Revista del Instituto de Salubridad y Efermedades Tropicales 1959;19:1-5.
18. Duke BO, Vincelette J, Moore PJ. Microfilariae in the cerebrospinal fluid, and neurological complications, during treatment of onchocerciasis with diethylcarbamazine. Tropenmed Parasitol 1976 Jun;27(2):123-132. [Medline: 941247]
19. Taylor HR, Greene BM. Ocular changes with oral and transepidermal diethylcarbamazine therapy of onchocerciasis. Br J Ophthalmol 1981 Jul;65(7):494-502 [FREE Full text] [Medline: 7020746]
20. Idro R, Opar B, Wamala J, Abbo C, Onzivua S, Mwaka DA, et al. Is nodding syndrome an Onchocerca volvulus-induced neuroinflammatory disorder? Uganda's story of research in understanding the disease. Int J Infect Dis 2016 Apr;45:112-117 [FREE Full text] [doi: 10.1016/j.ijid.2016.03.002] [Medline: 26987477]
21. Edwards G. Ivermectin: does P-glycoprotein play a role in neurotoxicity? Filaria J 2003 Oct 24;2 Suppl 1:S8 [FREE Full text] [doi: 10.1186/1475-2883-2-S1-S8] [Medline: 14975065]
22. Duke BO, Zea-Flores G, Castro J, Cupp EW, Munoz B. Comparison of the effects of a single dose and of four six-monthly doses of ivermectin on adult Onchocerca volvulus. Am J Trop Med Hyg 1991 Jul;45(1):132-137. [Medline: 1867345]
23. Basáñez M, Pion SDS, Boakes E, Filipe JAN, Churcher TS, Boussinesq M. Effect of single-dose ivermectin on onchocerca volvulus: a systematic review and meta-analysis. Lancet Infect Dis 2008 May;8(5):310-322. [doi: 10.1016/S1473-3099(08)70099-9] [Medline: 18471776]
24. Kipp W, Burnham G, Kamugisha J. Improvement in seizures after ivermectin. Lancet 1992 Sep 26;340(8822):789-790. [Medline: 1356196]
25. Colebunders R, Tepage F, Kashama J, Levick B, Mandro M, Komba M, et al. Onchocerciasis associated epilepsy in the Democratic Republic of the Congo. 2016 Presented at: International Congress for Tropical Medicine and Malaria; 2016; Brisbane.
26. Birbeck GL, Kalichi EMN. Epilepsy prevalence in rural Zambia: a door-to-door survey. Trop Med Int Health 2004 Jan;9(1):92-95. [Medline: 14728612]

27. Ngugi AK, Bottomley C, Kleinschmidt I, Wagner RG, Kakooza-Mwesige A, Ae-Ngibise K, SEEDS group. Prevalence of active convulsive epilepsy in sub-Saharan Africa and associated risk factors: cross-sectional and case-control studies. Lancet Neurol 2013 Mar;12(3):253-263 [FREE Full text] [doi: 10.1016/S1474-4422(13)70003-6] [Medline: 23375964]
28. Fisher RS, Acevedo C, Arzimanoglou A, Bogacz A, Cross JH, Elger CE, et al. ILAE official report: a practical clinical definition of epilepsy. Epilepsia 2014 Apr;55(4):475-482 [FREE Full text] [doi: 10.1111/epi.12550] [Medline: 24730690]
29. World Health Organisation. Epilepsy: a manuel for clinical officers in Africa. URL: http://www.who.int/mental_health/media/en/639.pdf [accessed 2017-07-06] [WebCite Cache ID 6rlL2Phk3]
30. Hopkins A. From "control to elimination:" a strategic change to win the end game. Int Health 2015;7(5):304-305.
31. Duke BO, Zea-Flores G, Gannon RT. On the reproductive activity of the female onchocerca volvulus. Trop Med Parasitol 1990 Dec;41(4):387-402. [Medline: 2075383]
32. Duke BO, Zea-Flores G, Muñoz B. The embryogenesis of onchocerca volvulus over the first year after a single dose of ivermectin. Trop Med Parasitol 1991 Sep;42(3):175-180. [Medline: 1801140]
33. Gardon J, Boussinesq M, Kamgno J, Gardon-Wendel N, Demanga N, Duke BOL. Effects of standard and high doses of ivermectin on adult worms of onchocerca volvulus: a randomised controlled trial. Lancet 2002 Jul 20;360(9328):203-210. [doi: 10.1016/S0140-6736(02)09456-4] [Medline: 12133654]
34. Kamgno J, Gardon J, Gardon-Wendel N, Demanga N, Duke BOL, Boussinesq M. Adverse systemic reactions to treatment of onchocerciasis with ivermectin at normal and high doses given annually or three-monthly. Trans R Soc Trop Med Hyg 2004 Aug;98(8):496-504. [doi: 10.1016/j.trstmh.2003.10.018] [Medline: 15186939]

Abbreviations

AED: antiepilepsy drug
CSF: cerebrospinal fluid
DRC: Democratic Republic of Congo
NS: nodding syndrome
OAE: onchocerciasis-associated epilepsy

Edited by G Eysenbach; submitted 19.12.16; peer-reviewed by A Hopkins, C Kaiser, C Chaccour; comments to author 19.01.17; revised version received 10.02.17; accepted 25.05.17; published 30.08.17.

Please cite as:
Colebunders R, Mandro M, Mukendi D, Dolo H, Suykerbuyk P, Van Oijen M
Ivermectin Treatment in Patients With Onchocerciasis-Associated Epilepsy: Protocol of a Randomized Clinical Trial
JMIR Res Protoc 2017;6(8):e137
URL: http://www.researchprotocols.org/2017/8/e137/
doi:10.2196/resprot.7186
PMID:28855148

Protocol

Project Stronger Together: Protocol to Test a Dyadic Intervention to Improve Engagement in HIV Care Among Sero-Discordant Male Couples in Three US Cities

Rob Stephenson[1,2*], MSc, PhD; Nicolas A Suarez[1,3*], BS; Robert Garofalo[4,5*], MD, MPH; Marco A Hidalgo[4,5*], PhD; Samuel Hoehnle[4,5*], BA; Jennie Thai[5*], BA; Matthew J Mimiaga[6,7,8,9*], ScD, MPH; Emily Brown[9*], BA; Anna Bratcher[10*], MSPH; Taylor Wimbly[10*], MPH; Patrick Sullivan[10*], PhD, DVM

[1]Center for Sexual and Health Disparities, University of Michigan, Ann Arbor, MI, United States

[2]Department of Health Behaviors and Biological Sciences, School of Nursing, University of Michigan, Ann Arbor, MI, United States

[3]Department of Health Behavior and Health Education, School of Public Health, University of Michigan, Ann Arbor, MI, United States

[4]Department of Pediatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[5]Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, United States

[6]Center for Health Equity Research, Brown University, Providence, RI, United States

[7]Departments of Behavioral and Social Health Sciences and Epidemiology, School of Public Health, Brown University, Providence, RI, United States

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

[9]The Fenway Institute, Fenway Health, Boston, MA, United States

[10]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[*]all authors contributed equally

Corresponding Author:
Rob Stephenson, MSc, PhD
Department of Health Behaviors and Biological Sciences
School of Nursing
University of Michigan
400 N Ingalls St
Ann Arbor, MI, 48109
United States
Phone: 1 734 615 0149
Email: rbsteph@med.umich.edu

Abstract

Background: An estimated one- to-two-thirds of new human immunodeficiency virus (HIV) infections among US men who have sex with men (MSM) occur within the context of primary partnerships. Despite this fact, there remains a lack of prevention interventions that focus on male sero-discordant dyads. Interventions that provide male couples with skills to manage HIV risk, and to support each other towards active engagement in HIV prevention and care, are urgently needed.

Objective: The objective of this paper is to describe the protocol for an innovative dyadic intervention (Stronger Together) that combines couples' HIV testing and dyadic adherence counseling to improve treatment adherence and engagement in care among HIV sero-discordant male couples in the United States.

Methods: The research activities involve a prospective randomized controlled trial (RCT) of approximately 165 venue- and clinic-recruited sero-discordant male couples (330 individuals: 165 HIV sero-negative and 165 HIV sero-positive). Couples randomized into the intervention arm receive couples' HIV counseling and testing plus dyadic adherence counseling, while those randomized to the control arm receive individual HIV counseling and testing. The study takes place in three cities: Atlanta, GA (study site Emory University); Boston, MA (study site The Fenway Institute); and Chicago, IL (study site Ann & Robert H. Lurie Children's Hospital of Chicago). Cohort recruitment began in 2015. Couples are followed prospectively for 24 months, with study assessments at baseline, 6, 12, 18, and 24 months.

Results: Stronger Together was launched in August 2014. To date, 160 couples (97% of the target enrollment) have been enrolled and randomized. The average retention rate across the three sites is 95%. Relationship dissolution has been relatively

low, with only 13 couples breaking up during the RCT. Of the 13 couples who have broken up, 10 of the 13 HIV-positive partners have been retained in the cohort; none of these HIV-positive partners have enrolled new partners into the RCT.

Conclusions: The intervention offers a unique opportunity for sero-discordant couples to support each other towards common HIV management goals by facilitating their development of tailored prevention plans via couples-based HIV testing and counseling, as well as problem-solving skills in Partner Strategies to Enhance Problem-solving Skills (STEPS).

Trial Registration: ClinicalTrials.gov NCT01772992; https://clinicaltrials.gov/ct2/show/NCT01772992 (Archived by WebCite at http://www.webcitation.org/6szFBVk1R)

(JMIR Res Protoc 2017;6(8):e170) doi:10.2196/resprot.7884

Introduction

Research has drawn attention to the role of male dyads in the US human immunodeficiency virus (HIV) epidemic, with primary partners identified as the source of approximately one-third [1] to two-thirds [2] of new HIV infections. Given these estimates, a significant paradigm shift in HIV prevention is needed, as efforts have traditionally focused on men who have sex with men (MSM), in particular gay-identifying men, as individuals rather than dyads. HIV prevention has predominantly emphasized HIV risks in the context of casual sex, largely ignoring the risk of HIV transmission that may occur within primary partnerships. Within male couples, various research findings have illustrated high rates of sexual risk behavior for HIV (with primary and casual partners), low rates of disclosure of potentially risky episodes with casual partners to primary partners, and reduced frequency of HIV testing [3-9]. Historically, HIV prevention efforts have focused on reducing the number of casual sex partners [10], indirectly messaging that primary partnerships pose reduced risk of HIV and conveying a misplaced sense of protection associated with primary partners [11,12]. However, there have been recent attempts to address this disproportionate focus on casual sex partners. The Office of the Global Acquired Immune Deficiency Syndrome (AIDS) Coordinator, through dissemination of prevention guidelines for MSM in the President's Emergency Plan for AIDS Relief-supported countries, now recommends couples' HIV testing and counseling (CHTC) for male couples [13].

CHTC has been used as an HIV prevention intervention for heterosexual couples in Africa for over 20 years [14]. Labeled as a, " *high leverage HIV prevention intervention* " by the US Centers for Disease Control and Prevention (CDC) [15], CHTC is considered to be an effective approach to HIV prevention among male couples. The difference between CHTC and individual HIV testing and counseling is that both partners receive counseling and testing at the same time, and do so together [16]. During the CHTC session, the counselor learns about the couple's relationship and provides tailored counseling and HIV prevention recommendations based on the characteristics of the couple's relationship and their joint HIV status [16]. Through the adaptation of CHTC and the high acceptability among MSM [14,17], preliminary data from MSM in three US cities (Atlanta, Chicago, Seattle) demonstrates the readiness of US MSM to receive and use CHTC [18,19]. Preliminary findings also suggest that male couples receiving CHTC, in which intimate partner violence (IPV) is not already present, do not report increased levels of IPV or relationship dissolution [20]. CHTC is now considered by the CDC to be an effective public health strategy, and is currently being implemented in over 40 US states [16,21].

A critical aspect of CHTC involves discussing a couple's sexual agreement. Sexual agreements are common among male couples and refer to mutually understood rules between two partners that describe the kinds of sexual behavior that is allowed within and outside of the relationship [4,22-29]. In CHTC, male couples discuss their sexual agreements, role-play with the counselor about how they would communicate about a breach in the sexual agreement, and develop an HIV prevention plan based on their agreement and couple sero-status. Research regarding male couples' agreements has demonstrated that men are less likely to practice condomless anal intercourse (CAI) with both primary and outside partners if they value and commit to their agreement and perceive their main partner to be dependable and invested in the relationship [28,30-32]. Promoting positive relationship dynamics has the potential to reduce couples' risk for HIV because higher trust, communication, commitment, and social support are associated with lower odds of breaking a sexual agreement, which can ultimately reduce unique HIV risks (eg, CAI) for the couple [3].

In the United States, CHTC has been utilized as a prevention strategy for male couples to learn about their HIV sero-status together and to develop a tailored prevention plan that reflects both their sero-status and their sexual agreement. However, the potential for CHTC to be used as an entry point into engagement in HIV care has been largely overlooked. Increasing evidence indicates that early initiation of antiretroviral therapy (ART) is beneficial for both the HIV-positive person and his partner. Observational cohort studies have demonstrated a 94% decreased mortality risk for the HIV-positive partner with initiation of ART [33,34], as well as significant protection for the HIV-negative partner from HIV infection due to early ART initiation and progression to viral suppression [35]. Adherence to ART is critical, as resultant viral suppression is correlated with increased health [35-38] and reduced likelihood of HIV transmission to an HIV-negative person. Approximately 95% of ART adherence is the threshold required to achieve viral suppression, which is a threshold that is considerably higher than levels of medication adherence observed in many of the observational studies that examined the impact of viral suppression on HIV progression and transmission [39-41]. Only 41% of HIV-positive persons in the United States are both aware of their HIV infection and are undertaking ongoing HIV care [42]. Estimates suggest that only 77% of HIV-positive persons are linked to care and approximately half (51%) remain in care, with only 28% of HIV-positive persons in the United States

achieving viral suppression [42]. Although a wide range of adherence rates (53-89%) have been documented in varied populations [40,43-52], the average rate of ART adherence is thought to be approximately 70% in the United States [42]. HIV-positive persons are also not receiving sufficient levels of HIV prevention counseling, and MSM are less likely than non-MSM to receive counseling in the preceding year [42].

There is evidence that dyadic interventions increase ART adherence when compared to individual adherence counseling [53]. In a randomized controlled trial (RCT) of 215 couples, including but not exclusively MSM couples, HIV-positive persons receiving ART adherence counseling with their partners had significantly higher levels of ART adherence [53]. The study examined the effects of joint counseling (in which counseling is delivered to the dyad) on adherence, but couples did not receive HIV testing together [53]. Social support among HIV-positive persons, including support from primary partners, is associated with fewer reported HIV risk behaviors with outside partners, greater self-efficacy to adhere, reported adherence, and lowered viral load after six months of follow-up [54,55]. Accordingly, couples who report higher relationship satisfaction are more likely to concur about the nature of their sexual agreement and to report not breaking the agreement [3,28].

In April 2012, World Health Organization released new guidelines for CHTC, including ART for treatment and prevention among sero-discordant couples [56]. The guidelines report a significant gap in evidence regarding the uptake and adherence to ART among sero-discordant couples, and highlight the role of CHTC in shaping uptake and adherence to ART [56]. This protocol outlines an intervention aimed at increasing engagement in HIV care among sero-discordant male couples in three US cities. The intervention draws upon two proven strategies to create a couples-focused package of care that incorporates dyadic HIV testing (CHTC) with dyadic adherence counseling. The intervention focuses on a couples-focused continuum of care, in which the couple is tested together and receives ART adherence counseling together, compared to a standard of care in which couples are tested and receive counseling individually. The RCT will examine and compare the intervention's effects on engagement in HIV care and achievement of viral suppression for the HIV-positive member of the dyad, as well as sexual risk-taking both within and outside the dyad. Understanding the efficacy of a couple's focused intervention for engagement in HIV care has the potential to inform the delivery of dyadic HIV prevention and care services for male couples, a group that is largely overlooked in current HIV prevention research and programming.

The intervention is grounded in Couple's Interdependence Theory [57], a framework that combines both interdependence theory and communal coping perspectives. The framework guides the selection of measures of behaviors and behavior change within the couple. These measures relate to our intervention in two ways. First, some aspects of communication and decision-making within the partnership may influence the efficacy of the intervention; couples with more constructive communication styles may benefit more from CHTC and achieve greater linkage to, and retention in, HIV care than couples with less constructive communication styles. Second, some aspects of partnerships, such as efficacy around implementing behavioral change, may actually be modified by the intervention. In these cases, changes in key characteristics of the partnerships may be in the causal pathway between the intervention and the adoption of ART, linkage to care, and safer behaviors within the partnership. We thus conceptualize the causal pathways as follows. Couples exposed to the intervention package will receive opportunities to talk about HIV, safer sex, and care-seeking within their relationship jointly with a qualified CHTC counselor. Relative to couples exposed to individually-focused testing and adherence counseling, exposure to CHTC may in turn impact communal coping, use of coping, and transformation of motivation, leading to initiation and maintenance of health-enhancing behavior (which we conceptualize as greater uptake and retention in care and ART adherence), lowering of sexual risk-taking inside and outside of the relationship, and a resultant lowering of sexually transmitted infection (STI) and HIV incidence. In our research design, predisposing factors, outcome efficacy, and couple efficacy will be collected separately from both partners before the HIV testing intervention is delivered, and will again be collected at each of the follow-up visits.

Methods

Design

The research activities involve a prospective RCT of approximately 165 venue- and clinic-recruited sero-discordant male couples (330 individuals: 165 HIV sero-negative and 165 HIV sero-positive). The study takes place in three cities: Atlanta, GA (study site Emory University); Boston, MA (study site The Fenway Institute); and Chicago, IL (study site Ann & Robert H. Lurie Children's Hospital of Chicago). Cohort recruitment began in 2015. Couples are followed prospectively for 24 months, with study assessments at baseline, 6, 12, 18, and 24 months.

Participants

Eligible participants are cis-gender male couples in which: (1) two men report having been in a relationship with each other for greater than one month, with a relationship defined as, "having a male partner who you are committed to above all others"; (2) both men are aged over 18 years; (3) individuals have lived in the Atlanta, Boston, or Chicago metro areas for greater than 3 months; (4) participants reported no recent history (in the past 12 months) of IPV or coercion; and (5) an HIV sero-discordant relationship exists, in which both partners have disclosed their sero-status to each other. Prevalent HIV sero-positive statuses are self-reported and are not confirmed by study staff.

Recruitment

Participants are recruited from the Atlanta, Boston, and Chicago metro areas via a multi-modal recruitment strategy. Recruitment takes place through physical and online/virtual spaces. Online sources include advertising on social media (eg, Facebook) and on geospatial dating apps (eg, Grindr). In-person recruitment is achieved by study staff attending lesbian, gay, bisexual, and

transgender (LGBT) events, visiting venues, meeting potential participants at clinic appointments, and posting flyers in gay-themed venues. All recruitment activities provide individuals with the study uniform resource locator (URL). The online screener can also be administered by study staff in-person or over the phone. When men visit the URL, a page is populated containing a short description of study activities; if they express an interest in participation and provide the metro area they live in, potential participants are taken to the study consent form, and if they consent, are directed to a short eligibility screener. Men who (1) do not consent or (2) do not meet the eligibility criteria are taken to a screen thanking them for their interest. Eligible men are directed to a registration process. During the registration process individuals provide their name, email address, and a cell phone number. Participants are also given the option to provide their main partner's email address and/or cell phone number so they can be contacted and screened to enroll the couple in the study together. Once both partners have (1) completed the consent forms, (2) finished the screening questionnaire, (3) been deemed eligible for the study, and (4) provided contact information, a staff member contacts the couple to schedule the couple for an in-person baseline visit.

Check-In and Informed Consent

When an eligible couple comes in for a baseline visit, they are assigned an identification number and administered a "Check In" survey. This survey generates a randomization number, confirms eligibility, verifies the couple is a real couple, and gathers further contact information and alternative contacts for the participants. A real couple is considered two individuals who are devoted to each other above all others. This language is used in all questions that gather self-report of relationship status. If a couple is no longer eligible or determined not to be partnered, they are dismissed without study staff specifying why, in order to avoid instigating IPV or revealing eligibility criteria. Eligible couples are taken into separate rooms and read the consent forms by study staff who answer any questions the participants may have. If both participants in the couple give consent, the couple is enrolled and they begin the baseline survey. If one or both members of the couple decline consent, the couple is dismissed without study staff specifying eligibility criteria.

Data Collection

After a couple gives informed consent, but before the couple is randomized to either the intervention or control group, each member of the couple is given a baseline survey. This survey is approximately 60-90 minutes long and collects data on demographics, relationship characteristics, sexual history, HIV care, and HIV prevention. In addition to survey data, biological samples are collected: STI testing, ART/preexposure prophylaxis (PrEP) drug adherence testing, and viral load testing. At the Boston site, however, STI testing is not conducted due to budgeting constraints.

Randomization

Upon individual completion of the check-in survey by both partners, couples are randomized to either the control arm (Individual Counseling) or the intervention arm (Couples Counseling). The treatment assignments are generated with the use of one pseudo-random number generator across all three study sites. The randomization process generates a random number between 0 and 100. If given an odd number, the couple is placed in the control treatment group; if given an even number, the couple is placed in the intervention treatment group.

Intervention

The intervention is a combination of CHTC and medication adherence counseling through the Partner Strategies to Enhance Problem-solving Skills (STEPS) method, which is an adaptation of an existing cognitive-behavioral intervention [16]. The intervention is comprised of three sessions. In the first session, lasting between 30-45 minutes, couples receive CHTC. The second and third sessions, lasting 60 minutes each, are held eight and ten weeks after the CHTC session, during which couples receive dyadic-focused ART adherence counseling. At the 6-, 12-, and 18-month follow-up visits, couples also receive Partner STEPS booster sessions.

Session One

CHTC sessions are conducted by a counselor who is trained in CHTC and last approximately 30-45 minutes. Only the HIV-negative partner is tested during the session. Posttest counseling focuses on dyadic prevention messages, and revisits the couple's HIV risk concerns and sexual agreements in light of their test results. While focusing on the needs of the HIV-positive partner is necessary, the discussion also emphasizes how the couple can work together to keep the HIV-positive partner healthy and reduce transmission risks within the relationship. The prevention counseling element of the CHTC session focuses on talking to the couple about prevention options (including PrEP) and asking them to consider which prevention options may work best based on their relationship needs, context, and unique risk profile.

Sessions Two and Three

Couples in the intervention arm will attend two additional visits consisting solely of adherence counseling at 8 and 10 weeks after their first CHTC visit. Based on the efficacious Partner STEPS intervention, CHTC counselors utilize motivational interviewing to improve ART adherence among HIV-positive individuals. The Partner STEPS intervention was developed by drawing from relationship-oriented theory, existing efficacious individual-level ART adherence interventions, couple-focused HIV prevention interventions, and expert consultation. New content was incorporated to address all aspects of the HIV care continuum (eg, linkage to, and retention in, care) and to draw on relationship strengths through interactive activities. The theory-based Partner STEPS intervention is delivered by a trained bachelors-level counselor (interventionist). Each session is designed to use relationship strengths to increase motivation for HIV care and treatment, and cover sequential intervention *steps* relating to specific challenges in HIV care engagement and barriers to ART adherence. For each step, couples work with a trained interventionist to identify their unique challenges, actively problem-solve with the interventionist, and articulate and commit to working together to implement a plan in which each partner agrees to complete specific tasks. Partner STEPS

counseling focuses on dyadic strategies to improve medication adherence and retention in care at each of ten *steps* for which Partner STEPS is named. Each step is a portion of HIV care that can present a challenge to those seeking care. The ten steps are: (1) transportation to appointments; (2) obtaining medications; (3) communicating with providers; (4) storing and transporting medications; (5) having a daily medication schedule; (6) coping with side effects; (7) adherence, self-care, and your relationship; (8) communicating within your relationship; (9) managing your social life and other relationships; and (10) dealing with privacy and disclosure. The counselors are trained to keep the focus on the couple by engaging both partners in problem solving and plan development. For sero-discordant couples, counselors are trained to focus the discussion on how *the couple* can *work together* to keep the HIV-positive partner healthy and to prevent transmission to the HIV-negative partner through medication adherence. Strategies to improve medication adherence and retention in care are tailored to the couple's unique relationship, as the counselor asks the couple to consider strategies that may work best based on their relationship needs, context, and unique health situation. Bazzi et al describe the protocol for developing and testing of the Partner STEPS intervention [16].

Control Group

Couples in the control group receive only one intervention visit, which is fewer than couples in the intervention. While it is possible that this aspect creates an attention effect, the control condition represents the current standard of care. At the baseline visit, the HIV-negative partner in the control group receives individual HIV counseling, testing, and referral (CTR). The HIV-positive partner receives information on the importance of ART uptake and adherence. Couples in the control arm do not receive Partner STEPS adherence counseling.

Follow-Up Visits

All couples in the prospective cohort return for follow-up visits at 6, 12, 18 and 24 months following the baseline survey visit. These visits consist of a survey similar to the baseline survey, biological sample collection, and either CTR for the HIV-negative partner in the control arm, information on adherence to ART for the HIV-positive partner in the control arm (but not counseling on adherence), or CHTC and Partner STEPS counseling for the intervention couples. CHTC is offered to control arm couples at the 24-month visit. At the 24-month visit, a sample of 30 couples (10 per city) is also invited to participate in a brief qualitative exit interview that inquires into evaluation of the study experience and perceived effectiveness of the counseling they received. Over the course of enrollment, couples also receive bi-monthly phone calls to check-in on their relationship and assess medication adherence. These phone calls also serve to help with retention.

Couple Dissolution

If a couple does not remain together throughout the follow-up period, HIV-positive partners will be retained in the study for the full 2-year follow-up period, while HIV-negative partners will return for one more follow-up visit and then will be censored from the cohort. If an HIV-positive partner in the intervention arm obtains a new partner who is eligible for the study, that partner will be invited to attend further follow-up visits, to continue providing CHTC to the original positive partner with the new partner. This new partner will participate in surveys, sample collections, counseling sessions and bi-monthly phone calls. At the Boston site, new negative partners are also invited to enroll if they are in the control arm.

Biomedical Measures

For HIV-negative partners, HIV sero-status is tested at each study assessment. For HIV-positive partners, viral load and ART levels are measured at each study assessment and collected via dried blood spots. All participants are tested for syphilis at each study assessment in the Atlanta and Chicago sites. Budgetary constraints prevented testing for other STIs, as well as syphilis testing, in the Boston site.

Outcomes

The primary outcomes are the HIV-positive partner's engagement in HIV care and his achievement of viral suppression. Engagement in care is conceptualized as also including linkage to care and retention in HIV care. For HIV-positive partners who report no engagement in HIV care at baseline, linkage to care is defined as attending at least one clinical care appointment, having at least one CD4 test performed, and having at least one viral load test performed within 3 months of the baseline visit. Retention in care is measured by determining participation in continuous care; that is, at least two or more routine HIV visits at least three months apart, receiving two or more CD4 tests, and receiving two or more viral load tests within a 12-month period [58]. At each study visit, a blood draw will be conducted to provide a measure of the HIV-positive partner's viral load and to test levels of ART. To supplement the viral load and ART-level biological markers, self-reported ART adherence is collected in each of the surveys. The AIDS Clinical Trial Group (ACTG) questionnaire includes items measuring adherence to medications for the past 4 days, adherence to scheduled instructions during the last weekend, and when any medication was last skipped. The questionnaire responses are weighted to calculate an adherence level from 0-100. The surveys also measure barriers to adherence via a 24-item scale based on the ACTG assessment for barriers to adherence to ART [59]. Participants are asked to note (using a scale ranging from *never* to *often*) if they missed their HIV medication over the past month for one of the provided reasons.

As secondary outcomes, the study measures sexual risk-taking and formation, and adherence of sexual agreements. For sexual agreements, participants are asked which of the following best describes their current sexual agreement with their main partner: (1) *both of us cannot have sex with outside partners*, (2) *we can have sex with outside partners, without any conditions or restrictions*, (3) *we can have sex with outside partners, but with some conditions or restrictions*, and (4) *we do not have an agreement*. Comparison of individual data will allow identification of discordant agreements. Participants will be asked if they have broken this agreement and whether this breakage was disclosed to their partner. To assess sexual behavior, measures adapted from the National HIV Behavioral

Surveillance (NHBS) behavioral inventory collect information both on sexual behaviors with the main sex partner in the 3 months before the interview, and on sexual behaviors with all sex partners outside the relationship. For sex with the main partner, men are asked to estimate the number of anal sex acts with the main partner and the number of those acts that were condom-protected. For each partner outside the relationship, men are asked a series of questions to include characteristics of the outside partner, HIV status (if known), whether the sex outside the relationship was disclosed to the main partner, the number and type of sex acts, and the proportion of those sex acts that were protected by condoms. Additional outcomes assessed are centered on Couple's Interdependence Theory.

Dyadic Characteristics

The four elements of Lewis' model [57] (*predisposing factors of couples*, *partner's transformation of motivation*, *process of communal coping*, and *use of communal coping*) are referred to as *dyadic characteristics*. In a recent RCT of CHTC, scales were developed to capture these constructs; all scales showed strong reliability, and evidence for construct validity was obtained for all scales [60]. In this intervention, each of the scales is collected in the baseline and follow-up surveys.

Predisposing Factors

Predisposing factors of couples uses several scales to measure this element. *Perceived Severity of HIV Scale:* this construct involves the perception of the personal, psychosocial, and physical consequences of a particular health threat. A total of 13 items were developed that crossed the three pertinent consequences of a particular health threat: personal, psychosocial, and physical. *Preferences for General Lifestyle Outcomes Scale*: this construct is defined as the degree to which interacting partners agree about the shared or joint outcomes in their relationship, which is composed of two subscales (*Preferences for General Lifestyle Scale* and *Preferences for Sexual Health Outcomes Scale*). The *Preferences for General Lifestyle Scale* includes six outcomes, including diet, nutrition, and social activities. *Preferences for Sexual Health Outcomes Scale* relates to sexual health (eg, reducing one's risk for HIV). In addition, scales to measure other predisposing factors of couples are proposed for inclusion. *Conflict Style* determines how respondents typically handle conflict in their relationships, so the Conflict Style Inventory will be included [61]. *Communication Style* is measured with the Communication Patterns Questionnaire Constructive Communication subscale [62]. Finally, *Problem-Solving Skills* are measured with the adherence problem solving/readiness scale [63].

Partner's Transformation of Motivation

In a recent RCT of CHTC, two measures were developed: ability of the participants to respond (1) cognitively and (2) emotionally to the health threat [20]. The scale for emotional response includes whether the respondent reports being fearful, nervous, or anxious about HIV. The scale for cognitive response includes whether the respondent reports understanding the risks of HIV transmission associated with being in a sero-discordant relationship, and the risks associated with outside sex partners.

Process and Use of Communal Coping

Several scales are used to measure this element. The *Outcome Efficacy to Reduce HIV Threat Scale* discusses how communal coping involves couples working together and making decisions to reduce the health threat. Three subscales were created to capture the full range of outcome efficacy related to these three processes of communal coping. For the first subscale, *Joint Effort*, the stem, "My partner and I believe that 'working together' versus on our own is an effective strategy" is used. For the second subscale, *Communication*, the stem, "Communicating with my partner is an effective strategy for" is used. For the third subscale, *Planning and Decision-making*, the stem, "My partner and I making decisions together rather than separately is an effective strategy" is used. The items for each of the three subscales were the same as the items used for the *Preferences for Sexual Health Outcomes Scale*. The *Couple Efficacy to Reduce HIV Threat Scale* defines couple efficacy as a couple's confidence that together they can engage in communal coping efforts. The study also assesses the occurrence of violence within the relationship using the Conflict Tactics Scale Revised [64] to assess both perpetration and experience of IPV.

Statistical Analyses

The analysis employs an intent-to-treat analysis design. The percentage of HIV-positive individuals who achieve viral suppression and report being fully engaged in care will be compared across arms, using Chi-square tests for significant difference. Retention in care is measured as the number of 6-month blocks during which at least one clinic visit was attended over the 2-year period following an initial attended visit; the percentage of HIV-positive individuals who report full retention in care (one clinic visit during each 6-month block) and the number of 6-month blocks during which care was received will be compared across study arms, using the appropriate tests for statistical significance. The capability of the intervention to yield longer-lasting effects in adherence endpoints over time will be examined. The visual analog scale data over the course of the study, at baseline and follow-up, will be analyzed using generalized linear models (GLMs) with properly-chosen (based on the distribution of dependent variables) link functions to analyze longitudinal adherence outcome data. The GLMs will be estimated using generalized estimating equations (GEE) with robust standard error estimates, which provide an extension of regression analyses to the cases of correlated or repeated observations and allows for inclusion of both categorical and count-dependent variables, and for appropriate modeling of covariance structures when observations are correlated across time. With appropriate link functions, GLMs can readily handle dependent variables with normal distributions, dichotomous outcomes, count data (Poisson distribution), and over-dispersed or zero-inflated count data (negative binomial models). The models will include the dyadic characteristics scales to examine the extent to which adherence is shaped by relationship functioning.

Dyadic Characteristics: Analysis of the Scale Data Over Time

Repeated-measures analyses using mixed linear models will be performed for scale data. These analyses will include participant-level characteristics and couple-level variation or clustering. Participants will be nested within couples with the participant as the experimental unit. Repeated-measures analyses for each scale will be analyzed with a means model with SAS Proc Mixed providing separate estimates of the means by time on study and treatment group. An unstructured variance-covariance form among the repeated measurements will be assumed for each outcome, and estimates of the standard errors of parameters will be used to perform statistical tests and construct 95% confidence intervals. T-tests will be used to compare the pairwise differences between the model-based treatment means (least-squares means) at each time point. Statistical tests will be 2-sided. The model-based means are unbiased with unbalanced and missing data if missing data are noninformative (missing at random). A dropout process is assumed to be missing at random if, conditional on the observed data, the dropout is independent of the unobserved measurements.

Incidence of Aggregate "Sex at Risk" Within Partnerships

At-risk sex will be defined as any CAI that occurs within the partnership during the follow-up period, even if the HIV-negative partner is taking PrEP or the HIV-positive partner has achieved viral suppression, in accordance with current CDC recommendations on PrEP and condom use [65]. The incidence of at-risk sex acts will be calculated as an incidence density, with the numerator being number of individual at-risk sex acts, and the denominator being person-years of follow time. Comparisons of the incidence of at-risk sex acts will be made by comparing incidence densities between the two arms. Incidence rates per couple-year of follow-up will be estimated and compared using exact methods based on the Poisson distribution when there are fewer than 15 events per subgroup or, when there are at least 15 events per group, by using the GEE approach. Baseline covariates include race, age, and duration of relationship. Period incidence rates (6-monthly incidence density rates) of at-risk sex will be estimated by performing a GEE Poisson regression analysis of the 6-monthly counts, implemented using the SAS PROC GENMOD procedure [66], and using an exchangeable correlation structure for the repeated observations of couples. The incidence density ratio (or incidence rate ratio; IRR) is the ratio of the incidence density in one treatment group (intervention arm) to that of a control group (standard of care). Results by each baseline covariate will be summarized as the IRR and the 95% confidence interval. In addition, we will tabulate data on disclosure of sex outside the relationship, the percentage of couples with agreements about sex outside the relationship, and the percentage of couples reporting agreement breakage or change in agreements. Prevalence of each outcome will be calculated, and prevalence of outcomes will be compared in the control and intervention groups using Chi-square tests or Fischer's exact tests, as appropriate.

The analysis will also examine conceptual mediators and epidemiologically identified moderators. If the Stronger Together intervention works to increase viral suppression and number of participants reporting full engagement in care among the intervention sample in significantly greater magnitude than the comparison condition, we will assess the extent to which this relationship works through several possible mediators, including dyadic factors and relationship functioning (ie, communication). For mediation analyses, we will employ MEDIATE procedures. MEDIATE estimates the total, direct, and indirect effects of causal variable(s) on the outcome variable through a proposed mediator variable or set of mediator variables. For effect modification (moderation) analyses, we will add interaction terms one-by-one for the intervention condition and the potential moderators (eg, age, race/ethnicity, and psychosocial factors such as depression and length of relationship). Significant or large interaction terms would suggest that the effects of the intervention differ for different subgroups, as defined by the moderators.

The analysis will also examine HIV sero-conversion among HIV-negative partners and syphilis as secondary outcomes: the prevalence and 24-month cumulative incidence of HIV and syphilis will be examined in aggregate and then by study arm. During the interim assessment visits (that occur before month 12 and then again before month 24) we will collect information on STI testing/diagnosis/treatment that participants received elsewhere since their last study assessment visit. These data will be used to adjust analyses. Using Cox proportional hazard regression models, we will assess if intervention status results in decreased odds (hazard ratio) of HIV and STI infection, separately, over the 24-month period.

The safety of the intervention at the individual level will be assessed by examining reported IPV within the relationship and relationship dissolution. Prevalence of each individual adverse outcome or any adverse outcome will be calculated, and prevalence of outcomes will be compared in the control and intervention groups using Chi-square tests or Fischer's exact tests, as appropriate.

Incentives

Individual participants receive US $50 for completing each study visit; this includes baseline, Partner STEPS sessions, and all four-to-six follow-up visits, depending on the study arm. If both members of the couple complete all visits, the total incentive amount is US $500 per couple ($250 per individual participant) in the control and US $700 per couple ($250-350 per individual participant) in the intervention arm.

Sample Size

Estimating 80% retention, we propose to enroll a sample of 165 male sero-discordant couples. The primary outcomes are engagement in HIV care and viral suppression. A sample of 165 sero-discordant couples provides statistical power (with 95% confidence and 80% power) to detect scientifically significant relative differences of 15%, 20%, and 25% in each of these outcomes between the two study arms. Additional health-enhancing behaviors include recent sexual risk-taking; using the data from our previous RCT of CHTC, this sample

of 165 sero-discordant couples provides statistical power to identify significant differences in sexual risk-taking between the two arms. Using the methods described by Rosner for sample size estimation for longitudinal studies [67], a sample size of 75 couples in each group will ensure statistical power (using a two-sided two-sample t-test) to identify differences in the dyadic scale constructs between arms. As secondary outcomes, the study will examine sero-conversion and syphilis incidence between the two arms, although the power is insufficient to detect significance differences between the study arms.

Trial Registration, Ethics, Consent, and Institutional Board Approval

The research and ethics presented in this study have been reviewed and approved by the Institutional Review Boards (IRBs) of Emory University (IRB #00065111), Lurie Children's (IRB #2014-15896) and The Fenway Institute (IRB #FWA00000145), in addition to a Data Safety Monitoring Board that meets annually. The study is also registered on ClinicalTrials.gov (NCT01772992).

Results

Stronger Together was launched in August 2014. To date, 160 couples (97% of the target enrollment) have been enrolled and randomized. The average retention rate across the surveys is 95%: current retention at 24 months is >90% across sites. Relationship dissolution has been relatively low, with only 13 couples breaking up during the RCT (13/160, 8%). Of the 13 couples who have broken up, 10 of the 13 (77%) HIV-positive partners have been retained in the cohort; none of these HIV-positive partners have enrolled new partners in the RCT.

Discussion

Dyadic interventions provide an opportunity for male sero-discordant couples to learn the skills necessary to work together to manage HIV in their relationships. By developing tailored prevention plans in CHTC and developing problem solving skills in Partner STEPS, the intervention offers a unique opportunity for sero-discordant couples to support each other towards common HIV management goals. The intervention tested in this protocol builds off the current success of CHTC. The current intervention extends the focus of dyadic interventions for male couples across the continuum of HIV care, allowing male couples to develop the skills necessary to support active and successful engagement of HIV care. The intervention has the potential to fill a critical gap in efficacious interventions for male sero-discordant couples who, despite evidence of high rates of transmission within partnerships, have been largely ignored by HIV research and programming.

Conflicts of Interest

None declared.

References

1. Goodreau SM, Carnegie NB, Vittinghoff E, Lama JR, Sanchez J, Grinsztejn B, et al. What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? PLoS One 2012;7(11):e50522 [FREE Full text] [doi: 10.1371/journal.pone.0050522] [Medline: 23209768]
2. Sullivan PS, Salazar L, Buchbinder S, Sanchez TH. Estimating the proportion of HIV transmissions from main sex partners among men who have sex with men in five US cities. AIDS 2009 Jun 1;23(9):1153-1162. [doi: 10.1097/QAD.0b013e32832baa34] [Medline: 19417579]
3. Gomez AM, Beougher SC, Chakravarty D, Neilands TB, Mandic CG, Darbes LA, et al. Relationship dynamics as predictors of broken agreements about outside sexual partners: implications for HIV prevention among gay couples. AIDS Behav 2012 Aug;16(6):1584-1588. [doi: 10.1007/s10461-011-0074-0] [Medline: 22020757]
4. Hoff CC, Beougher SC, Chakravarty D, Darbes LA, Neilands TB. Relationship characteristics and motivations behind agreements among gay male couples: differences by agreement type and couple serostatus. AIDS Care 2010 Jul;22(7):827-835 [FREE Full text] [doi: 10.1080/09540120903443384] [Medline: 20635246]
5. Stephenson R, White D, Darbes L, Hoff C, Sullivan P. HIV testing behaviors and perceptions of risk of HIV infection among MSM with main partners. AIDS Behav 2015 Mar;19(3):553-560 [FREE Full text] [doi: 10.1007/s10461-014-0862-4] [Medline: 25081599]
6. Stall R, Mills TC, Williamson J, Hart T, Greenwood G, Paul J, et al. Association of co-occurring psychosocial health problems and increased vulnerability to HIV/AIDS among urban men who have sex with men. Am J Public Health 2003 Jun;93(6):939-942. [Medline: 12773359]
7. Chakravarty D, Hoff CC, Neilands TB, Darbes LA. Rates of testing for HIV in the presence of serodiscordant UAI among HIV-negative gay men in committed relationships. AIDS Behav 2012 Oct;16(7):1944-1948 [FREE Full text] [doi: 10.1007/s10461-012-0181-6] [Medline: 22460227]
8. Mitchell JW, Petroll AE. Patterns of HIV and sexually transmitted infection testing among men who have sex with men couples in the United States. Sex Transm Dis 2012 Nov;39(11):871-876 [FREE Full text] [doi: 10.1097/OLQ.0b013e3182649135] [Medline: 23060078]
9. Phillips G, Hightow-Weidman LB, Arya M, Fields SD, Halpern-Felsher B, Outlaw AY, et al. HIV testing behaviors of a cohort of HIV-positive racial/ethnic minority YMSM. AIDS Behav 2012 Oct;16(7):1917-1925. [doi: 10.1007/s10461-012-0193-2] [Medline: 22555382]

10. Kelly JA, St Lawrence JS, Diaz YE, Stevenson LY, Hauth AC, Brasfield TL, et al. HIV risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. Am J Public Health 1991 Feb;81(2):168-171. [Medline: 1990853]
11. Pruitt KL, White D, Mitchell JW, Stephenson R. Sexual agreements and intimate-partner violence among male couples. Int J Sex Health 2015 May 07;27(4):429-441. [doi: 10.1080/19317611.2015.1037037]
12. Goldenberg T, Clarke D, Stephenson R. "Working together to reach a goal": MSM's perceptions of dyadic HIV care for same-sex male couples. J Acquir Immune Defic Syndr 2013 Nov 1;64 Suppl 1:S52-S61 [FREE Full text] [doi: 10.1097/QAI.0b013e3182a9014a] [Medline: 24126448]
13. The U.S. President's Emergency Plan for AIDS Relief (PEPFAR). 2011 May. Technical Guidance on Combination HIV Prevention URL: https://www.pepfar.gov/documents/organization/164010.pdf [accessed 2017-08-09] [WebCite Cache ID 6sbCm04x2]
14. Sullivan PS, Stephenson R, Grazter B, Wingood G, Diclemente R, Allen S, et al. Adaptation of the African couples HIV testing and counseling model for men who have sex with men in the United States: an application of the ADAPT-ITT framework. Springerplus 2014;3:249 [FREE Full text] [doi: 10.1186/2193-1801-3-249] [Medline: 24877036]
15. Painter TM. Voluntary counseling and testing for couples: a high-leverage intervention for HIV/AIDS prevention in sub-Saharan Africa. Soc Sci Med 2001 Dec;53(11):1397-1411. [Medline: 11710416]
16. Bazzi AR, Fergus KB, Stephenson R, Finneran CA, Coffey-Esquivel J, Hidalgo MA, et al. A dyadic behavioral intervention to optimize same sex male couples' engagement across the HIV care continuum: development of and protocol for an innovative couples-based approach (Partner Steps). JMIR Res Protoc 2016 Aug 25;5(3):e168 [FREE Full text] [doi: 10.2196/resprot.6271] [Medline: 27562905]
17. Wagenaar BH, Christiansen-Lindquist L, Khosropour C, Salazar LF, Benbow N, Prachand N, et al. Willingness of US men who have sex with men (MSM) to participate in Couples HIV Voluntary Counseling and Testing (CVCT). PLoS One 2012;7(8):e42953 [FREE Full text] [doi: 10.1371/journal.pone.0042953] [Medline: 22905191]
18. Stephenson R, Sullivan PS, Salazar LF, Gratzer B, Allen S, Seelbach E. Attitudes towards couples-based HIV testing among MSM in three US cities. AIDS Behav 2011 Apr;15 Suppl 1:S80-S87 [FREE Full text] [doi: 10.1007/s10461-011-9893-2] [Medline: 21336607]
19. Neme S, Goldenberg T, Stekler JD, Sullivan PS, Stephenson R. Attitudes towards couples HIV testing and counseling among Latino men who have sex with men in the Seattle area. AIDS Care 2015;27(10):1354-1359. [doi: 10.1080/09540121.2015.1058894] [Medline: 26272715]
20. Sullivan PS, White D, Rosenberg ES, Barnes J, Jones J, Dasgupta S, et al. Safety and acceptability of couples HIV testing and counseling for US men who have sex with men: a randomized prevention study. J Int Assoc Provid AIDS Care 2014;13(2):135-144 [FREE Full text] [doi: 10.1177/2325957413500534] [Medline: 23995295]
21. Stephenson R, Grabbe KL, Sidibe T, McWilliams A, Sullivan PS. Technical assistance needs for successful implementation of Couples HIV Testing and Counseling (CHTC) intervention for male couples at US HIV testing sites. AIDS Behav 2016 Apr;20(4):841-847. [doi: 10.1007/s10461-015-1150-7] [Medline: 26253221]
22. Hoff CC, Beougher SC. Sexual agreements among gay male couples. Arch Sex Behav 2008 Jun;39(3):774-787 [FREE Full text] [doi: 10.1007/s10508-008-9393-2] [Medline: 18686027]
23. Crawford JM, Rodden P, Kippax S, Van de Ven P. Negotiated safety and other agreements between men in relationships: risk practice redefined. Int J STD AIDS 2001 Mar;12(3):164-170. [Medline: 11231869]
24. Gass K, Hoff CC, Stephenson R, Sullivan PS. Sexual agreements in the partnerships of Internet-using men who have sex with men. AIDS Care 2012;24(10):1255-1263 [FREE Full text] [doi: 10.1080/09540121.2012.656571] [Medline: 22375729]
25. Hosking W. Australian gay men's satisfaction with sexual agreements: the roles of relationship quality, jealousy, and monogamy attitudes. Arch Sex Behav 2014 May;43(4):823-832. [doi: 10.1007/s10508-013-0197-7] [Medline: 24287963]
26. Kippax S, Noble J, Prestage G, Crawford JM, Campbell D, Baxter D, et al. Sexual negotiation in the AIDS era: negotiated safety revisited. AIDS 1997 Feb;11(2):191-197. [Medline: 9030366]
27. Mitchell JW. Characteristics and allowed behaviors of gay male couples' sexual agreements. J Sex Res 2013;51(3):316-328 [FREE Full text] [doi: 10.1080/00224499.2012.727915] [Medline: 23514544]
28. Mitchell JW, Harvey SM, Champeau D, Moskowitz DA, Seal DW. Relationship factors associated with gay male couples' concordance on aspects of their sexual agreements: establishment, type, and adherence. AIDS Behav 2012 Aug;16(6):1560-1569 [FREE Full text] [doi: 10.1007/s10461-011-0064-2] [Medline: 22012148]
29. Prestage G, Jin F, Zablotska I, Grulich A, Imrie J, Kaldor J, et al. Trends in agreements between regular partners among gay men in Sydney, Melbourne and Brisbane, Australia. AIDS Behav 2008 May;12(3):513-520. [doi: 10.1007/s10461-007-9351-3] [Medline: 18188690]
30. Mitchell JW, Champeau D, Harvey SM. Actor-partner effects of demographic and relationship factors associated with HIV risk within gay male couples. Arch Sex Behav 2012 Oct;42(7):1337-1345 [FREE Full text] [doi: 10.1007/s10508-012-9985-8] [Medline: 22875716]
31. Mitchell JW, Garcia L, Champeau D, Harvey SM, Petroll AE. HIV-negative seroconcordant gay male couples' attitudes, intentions, and perceived behavioral control for planned condom use within and outside of their relationships. Int J Sex Health 2012 Oct;24(4):239-253. [doi: 10.1080/19317611.2012.707168]

32. Mitchell JW, Harvey SM, Champeau D, Seal DW. Relationship factors associated with HIV risk among a sample of gay male couples. AIDS Behav 2012 Feb;16(2):404-411 [FREE Full text] [doi: 10.1007/s10461-011-9976-0] [Medline: 21614560]
33. Kitahata MM, Gange SJ, Abraham AG, Merriman B, Saag MS, Justice AC, NA-ACCORD Investigators. Effect of early versus deferred antiretroviral therapy for HIV on survival. N Engl J Med 2009 Apr 30;360(18):1815-1826 [FREE Full text] [doi: 10.1056/NEJMoa0807252] [Medline: 19339714]
34. When To Start Consortium, Sterne JAC, May M, Costagliola D, de Wolf F, Phillips AN, et al. Timing of initiation of antiretroviral therapy in AIDS-free HIV-1-infected patients: a collaborative analysis of 18 HIV cohort studies. Lancet 2009 Apr 18;373(9672):1352-1363 [FREE Full text] [doi: 10.1016/S0140-6736(09)60612-7] [Medline: 19361855]
35. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. 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]
36. Hogg RS, Heath KV, Yip B, Craib KJ, O'Shaughnessy MV, Schechter MT, et al. Improved survival among HIV-infected individuals following initiation of antiretroviral therapy. JAMA 1998 Feb 11;279(6):450-454. [Medline: 9466638]
37. Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med 1998 Mar 26;338(13):853-860. [doi: 10.1056/NEJM199803263381301] [Medline: 9516219]
38. Mocroft A, Ledergerber B, Katlama C, Kirk O, Reiss P, d'Arminio Monforte A, EuroSIDA study group. Decline in the AIDS and death rates in the EuroSIDA study: an observational study. Lancet 2003 Jul 05;362(9377):22-29. [Medline: 12853195]
39. Singh N, Berman SM, Swindells S, Justis JC, Mohr JA, Squier C, et al. Adherence of human immunodeficiency virus-infected patients to antiretroviral therapy. Clin Infect Dis 1999 Oct;29(4):824-830. [doi: 10.1086/520443] [Medline: 10589897]
40. Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squier C, et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann Intern Med 2000 Jul 4;133(1):21-30. [Medline: 10877736]
41. Bartlett JA. Addressing the challenges of adherence. J Acquir Immune Defic Syndr 2002 Feb 01;29 Suppl 1:S2-10. [Medline: 11832696]
42. Centers for Disease Control and Prevention (CDC). Vital signs: HIV prevention through care and treatment--United States. MMWR Morb Mortal Wkly Rep 2011 Dec 2;60(47):1618-1623 [FREE Full text] [Medline: 22129997]
43. Mannheimer SB, Matts J, Telzak E, Chesney M, Child C, Wu AW, The Terry Beirn Community Programs for Clinical Research on AIDS. Quality of life in HIV-infected individuals receiving antiretroviral therapy is related to adherence. AIDS Care 2005 Jan;17(1):10-22. [Medline: 15832830]
44. Golin CE, Liu H, Hays RD, Miller LG, Beck CK, Ickovics J, et al. A prospective study of predictors of adherence to combination antiretroviral medication. J Gen Intern Med 2002 Oct;17(10):756-765 [FREE Full text] [Medline: 12390551]
45. Bangsberg DR, Hecht FM, Charlebois ED, Zolopa AR, Holodniy M, Sheiner L, et al. Adherence to protease inhibitors, HIV-1 viral load, and development of drug resistance in an indigent population. AIDS 2000 Mar 10;14(4):357-366. [Medline: 10770537]
46. Wagner GJ, Kanouse DE, Koegel P, Sullivan G. Adherence to HIV antiretrovirals among persons with serious mental illness. AIDS Patient Care STDs 2003 Apr;17(4):179-186. [doi: 10.1089/108729103321619782] [Medline: 12737641]
47. Howard AA, Arnsten JH, Lo Y, Vlahov D, Rich JD, Schuman P, Her Study Group. A prospective study of adherence and viral load in a large multi-center cohort of HIV-infected women. AIDS 2002 Nov 08;16(16):2175-2182. [Medline: 12409739]
48. McNabb J, Ross JW, Abriola K, Turley C, Nightingale CH, Nicolau DP. Adherence to highly active antiretroviral therapy predicts virologic outcome at an inner-city human immunodeficiency virus clinic. Clin Infect Dis 2001 Sep 01;33(5):700-705. [doi: 10.1086/322590] [Medline: 11486292]
49. Barroso PF, Schechter M, Gupta P, Bressan C, Bomfim A, Harrison LH. Adherence to antiretroviral therapy and persistence of HIV RNA in semen. J Acquir Immune Defic Syndr 2003 Apr 1;32(4):435-440. [Medline: 12640203]
50. Ammassari A, Murri R, Pezzotti P, Trotta MP, Ravasio L, De Longis P, AdICONA Study Group. Self-reported symptoms and medication side effects influence adherence to highly active antiretroviral therapy in persons with HIV infection. J Acquir Immune Defic Syndr 2001 Dec 15;28(5):445-449. [Medline: 11744832]
51. Ickovics JR, Meade CS. Adherence to HAART among patients with HIV: breakthroughs and barriers. AIDS Care 2002 Jun;14(3):309-318. [doi: 10.1080/09540120220123685] [Medline: 12042076]
52. Arnsten JH, Demas PA, Farzadegan H, Grant RW, Gourevitch MN, Chang CJ, et al. Antiretroviral therapy adherence and viral suppression in HIV-infected drug users: comparison of self-report and electronic monitoring. Clin Infect Dis 2001 Oct 15;33(8):1417-1423 [FREE Full text] [doi: 10.1086/323201] [Medline: 11550118]
53. Remien RH, Stirratt MJ, Dolezal C, Dognin JS, Wagner GJ, Carballo-Dieguez A, et al. Couple-focused support to improve HIV medication adherence: a randomized controlled trial. AIDS 2005 May 20;19(8):807-814. [Medline: 15867495]
54. Simoni JM, Frick PA, Huang B. A longitudinal evaluation of a social support model of medication adherence among HIV-positive men and women on antiretroviral therapy. Health Psychol 2006 Jan;25(1):74-81 [FREE Full text] [doi: 10.1037/0278-6133.25.1.74] [Medline: 16448300]

55. Darbes LA, Chakravarty D, Beougher SC, Neilands TB, Hoff CC. Partner-provided social support influences choice of risk reduction strategies in gay male couples. AIDS Behav 2012 Jan;16(1):159-167 [FREE Full text] [doi: 10.1007/s10461-010-9868-8] [Medline: 21221756]
56. WHO. Guidance on couples HIV counseling and testing including antriretroviral therapy for treatment and prevention in serodiscordant couples. WHO HIV/AIDS Programme 2012 Apr. [Medline: 23700649]
57. Lewis MA, McBride CM, Pollak KI, Puleo E, Butterfield RM, Emmons KM. Understanding health behavior change among couples: an interdependence and communal coping approach. Soc Sci Med 2006 Mar;62(6):1369-1380. [doi: 10.1016/j.socscimed.2005.08.006] [Medline: 16146666]
58. Committee to Review Data Systems for Monitoring HIV Care. Monitoring HIV care in the United States: IndicatorsData Systems. National Academies Press 2012 Mar. [Medline: 24830072]
59. Chesney MA, Ickovics JR, Chambers DB, Gifford AL, Neidig J, Zwickl B, Patient care committee & adherence working group of the outcomes committee of the Adult AIDS Clinical Trials Group (AACTG). Self-reported adherence to antiretroviral medications among participants in HIV clinical trials: the AACTG adherence instruments. AIDS Care 2000 Jun;12(3):255-266. [doi: 10.1080/09540120050042891] [Medline: 10928201]
60. Salazar LF, Stephenson RB, Sullivan PS, Tarver R. Development and validation of HIV-related dyadic measures for men who have sex with men. J Sex Res 2013;50(2):164-177. [doi: 10.1080/00224499.2011.636845] [Medline: 22206480]
61. Levinger G, Pietromonaco P. Conflict Style Inventory. University of Massachusetts, Amherst. University of Massachusetts, Amherst; 1989. URL: http://people.umass.edu/monaco/ConflictStyleQue.pdf?_ga=2.112951236.1133740566.1503436941-696446127.1503436941 [accessed 2017-08-22] [WebCite Cache ID 6suqmNWAQ]
62. Heavey C, Larson B, Zumtobel D, Christensen A. The communication patterns questionnaire: the reliability and validity of a constructive communication subscale. J Marriage Fam 1996 Aug;58(3):796. [doi: 10.2307/353737]
63. Balfour L, Tasca GA, Kowal J, Corace K, Cooper CL, Angel JB, et al. Development and validation of the HIV Medication Readiness Scale. Assessment 2007 Dec;14(4):408-416. [doi: 10.1177/1073191107304295] [Medline: 17986658]
64. Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The Revised Conflict Tactics Scales (CTS2). J Fam Issues 1996 May 01;17(3):283-316 [FREE Full text] [doi: 10.1177/019251396017003001]
65. Centers for Disease Control and Prevention. 2017. HIV Basics - PrEP URL: https://www.cdc.gov/hiv/basics/prep.html [accessed 2017-06-26] [WebCite Cache ID 6rVnSWw6x]
66. Stokes M, Davis C, Koch G. Categorical data analysis using the SAS system. Cary, NC: SAS Institute; 2001.
67. Rosner B. Fundamentals of Biostatistics. Pacific Grove, CA: Duxbury Press; 2010.

Abbreviations

ACTG: AIDS Clinical Trial Group
AIDS: acquired immune deficiency syndrome
ART: antiretroviral therapy
CAI: condomless anal intercourse
CHTC: couples' HIV testing and counseling
CTR: counseling, testing, and referral
GEE: generalized estimating equations
GLM: generalized linear model
HIV: human immunodeficiency virus
IPV: intimate partner violence
IRB: Institutional Review Board
IRR: incidence rate ratio
LGBT: lesbian, gay, bisexual, and transgender
MSM: men who have sex with men
NHBS: National HIV Behavioral Surveillance
PrEP: preexposure prophylaxis
RCT: randomized controlled trial
STEPS: Strategies to Enhance Problem-solving Skills
STI: sexually transmitted infection
URL: Uniform Resource Locator

Edited by G Eysenbach; submitted 18.04.17; peer-reviewed by J Jones, B Koblin; comments to author 18.05.17; revised version received 10.07.17; accepted 21.07.17; published 31.08.17.

Please cite as:
Stephenson R, Suarez NA, Garofalo R, Hidalgo MA, Hoehnle S, Thai J, Mimiaga MJ, Brown E, Bratcher A, Wimbly T, Sullivan P
Project Stronger Together: Protocol to Test a Dyadic Intervention to Improve Engagement in HIV Care Among Sero-Discordant Male Couples in Three US Cities
JMIR Res Protoc 2017;6(8):e170
URL: http://www.researchprotocols.org/2017/8/e170/
doi:10.2196/resprot.7884
PMID:28860107

Protocol

A Mobile Device App to Reduce Medication Errors and Time to Drug Delivery During Pediatric Cardiopulmonary Resuscitation: Study Protocol of a Multicenter Randomized Controlled Crossover Trial

Johan N Siebert[1], MD; Frederic Ehrler[2], PhD; Christian Lovis[2], MPH, FACMI, MD; Christophe Combescure[3], MSPH, MSc (Biostatistics), PhD; Kevin Haddad[1], RN; Alain Gervaix[1], MD; Sergio Manzano[1], MD

[1]Department of Pediatric Emergency Medicine, Geneva Children's Hospital, University Hospitals of Geneva, Geneva, Switzerland

[2]Division of Medical Information Sciences, Department of Radiology and Medical Informatics, University Hospitals of Geneva, Geneva, Switzerland

[3]Division of Clinical Epidemiology, Department of Health and Community Medicine, University of Geneva and University Hospitals of Geneva, Geneva, Switzerland

Corresponding Author:
Johan N Siebert, MD
Department of Pediatric Emergency Medicine
Geneva Children's Hospital
University Hospitals of Geneva
Avenue de la Roseraie 47
Geneva,
Switzerland
Phone: 41 22 372 45 55
Fax: 41 22 372 54 23
Email: Johan.Siebert@hcuge.ch

Abstract

Background: During pediatric cardiopulmonary resuscitation (CPR), vasoactive drug preparation for continuous infusions is complex and time-consuming. The need for individual specific weight-based drug dose calculation and preparation places children at higher risk than adults for medication errors. Following an evidence-based and ergonomic driven approach, we developed a mobile device app called Pediatric Accurate Medication in Emergency Situations (PedAMINES), intended to guide caregivers step-by-step from preparation to delivery of drugs requiring continuous infusion. In a prior single center randomized controlled trial, medication errors were reduced from 70% to 0% by using PedAMINES when compared with conventional preparation methods.

Objective: The purpose of this study is to determine whether the use of PedAMINES in both university and smaller hospitals reduces medication dosage errors (primary outcome), time to drug preparation (TDP), and time to drug delivery (TDD) (secondary outcomes) during pediatric CPR when compared with conventional preparation methods.

Methods: This is a multicenter, prospective, randomized controlled crossover trial with 2 parallel groups comparing PedAMINES with a conventional and internationally used drug infusion rate table in the preparation of continuous drug infusion. The evaluation setting uses a simulation-based pediatric CPR cardiac arrest scenario with a high-fidelity manikin. The study involving 120 certified nurses (sample size) will take place in the resuscitation rooms of 3 tertiary pediatric emergency departments and 3 smaller hospitals. After epinephrine-induced return of spontaneous circulation, nurses will be asked to prepare a continuous infusion of dopamine using either PedAMINES (intervention group) or the infusion table (control group) and then prepare a continuous infusion of norepinephrine by crossing the procedure. The primary outcome is the medication dosage error rate. The secondary outcome is the time in seconds elapsed since the oral prescription by the physician to drug delivery by the nurse in each allocation group. TDD includes TDP. Stress level during the resuscitation scenario will be assessed for each participant by questionnaire and recorded by the heart rate monitor of a fitness watch. The study is formatted according to the Consolidated Standards of Reporting Trials Statement for Randomized Controlled Trials of Electronic and Mobile Health Applications and Online TeleHealth (CONSORT-EHEALTH) and the Reporting Guidelines for Health Care Simulation Research.

Results: Enrollment and data analysis started in March 2017. We anticipate the intervention will be completed in late 2017, and study results will be submitted in early 2018 for publication expected in mid-2018. Results will be reported in line with recommendations from CONSORT-EHEALTH and the Reporting Guidelines for Health Care Simulation Research .

Conclusions: This paper describes the protocol used for a clinical trial assessing the impact of a mobile device app to reduce the rate of medication errors, time to drug preparation, and time to drug delivery during pediatric resuscitation. As research in this area is scarce, results generated from this study will be of great importance and might be sufficient to change and improve the pediatric emergency care practice.

Trial Registration: ClinicalTrials.gov NCT03021122; https://clinicaltrials.gov/ct2/show/NCT03021122 (Archived by WebCite at http://www.webcitation.org/6nfVJ5b4R)

(JMIR Res Protoc 2017;6(8):e167) doi:10.2196/resprot.7901

KEYWORDS

resuscitation; medication errors; pharmaceutical preparations; pediatrics; biomedical technology; children; emergency; simulation

Introduction

While many drugs can be directly injected without prior preparation, others require a fastidious and complicated preparation and dilution by nurses before administration as continuous infusions. During pediatric cardiopulmonary resuscitation (CPR), quick, accurate, and safe preparation and administration of intravenous (IV) vasoactive drugs for continuous infusion is both complex and time-consuming [1-3]. In some critical situations such as postcardiac arrest return of spontaneous circulation (ROSC) or septic shock, preparing those drugs is particularly challenging. Contrarily to adults, children require individual specific weight-based drug dose calculation and preparation. The lower dosing-error tolerance [4] places children at higher risk than adults for life-threatening errors [5-7]. Medication errors have been reported in up to 41% of pediatric resuscitations, the most common being incorrect medication dosage, found in up to 65% of cases [8]. Proper preparation and delivery of these drugs could favorably affect pediatric resuscitation outcomes.

In resuscitations, time is also a decisive success criterion. It is well established that during the first 15 minutes of pediatric resuscitation, survival and favorable neurological outcome decrease linearly by 2.1% and 1.2% per minute, respectively [9]. They are negatively affected by time to drug preparation (TDP) and time to drug delivery (TDD) [10]. In a study with adults in cardiac arrest, the chance of ROSC was decreased by 4% for every 1-minute delay in delivery of a vasopressor [11].

To address these problems, we followed an evidence-based and ergonomic approach [12] to develop an innovative and customizable mobile device app called Pediatric Accurate Medication in Emergency Situations (PedAMINES). This app was designed to support nurses and physicians step-by-step from order to delivery of a wide range of drugs in real time, including those requiring continuous infusion [13]. In a previous single center simulation-based randomized controlled trial, we have shown that medication errors reduced significantly from 70% to 0% by using PedAMINES when compared with conventional methods [14]. PedAMINES also dramatically reduced TDP and TDD. The purpose of this protocol is to investigate whether the use of PedAMINES might similarly reduce medication errors, TDP, and TDD in other university hospitals and in smaller hospitals where nurses and physicians are exposed to a much lower extent to pediatric resuscitations. Simulation was used as an investigational method to assess these outcomes. We hypothesize that PedAMINES might reduce medication errors and delays to TDP and delivery independently of the existing conventional preparation methods or nurse skills. We expect even bigger improvements in smaller hospitals due to strong disparities between nurses less exposed to pediatric resuscitations.

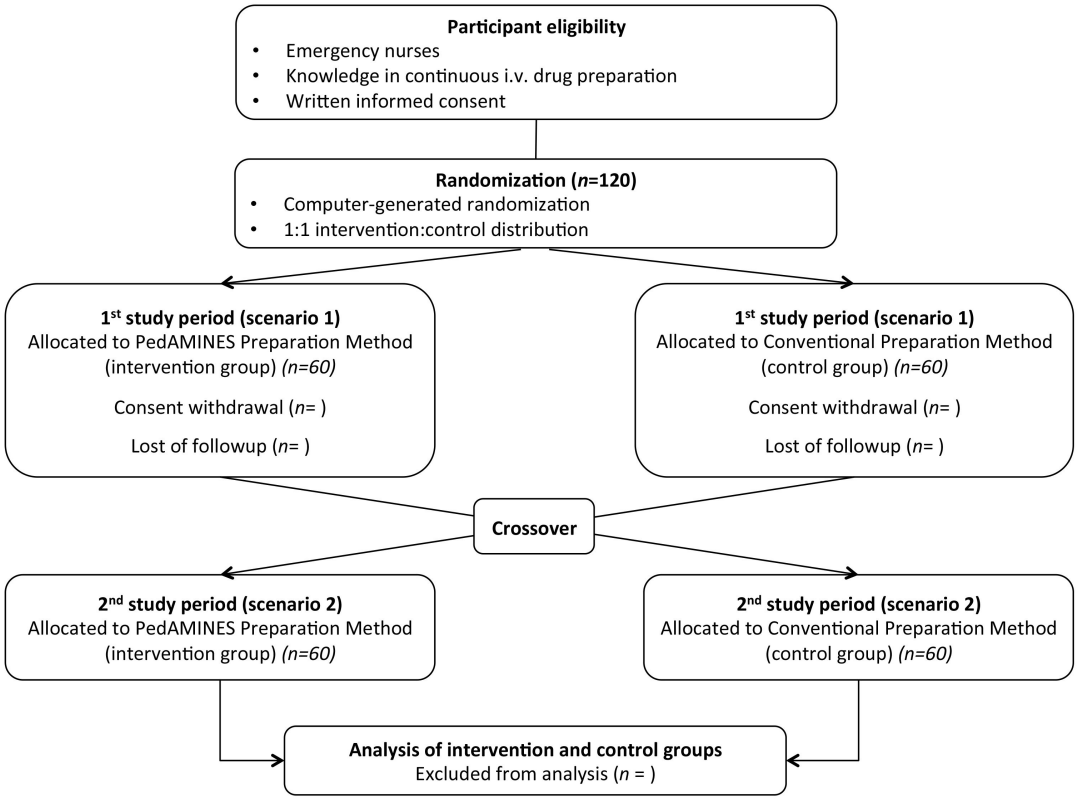
Methods

Study Design

The study is a prospective, multicenter, randomized controlled crossover trial with 2 parallel groups (Figure 1) comparing PedAMINES [13,14] with a conventional and internationally used drug infusion rate table method [15] (derived from the rule of six [16]) in the preparation of continuous drug infusion during a standardized simulation-based pediatric postcardiac arrest scenario. The infusion rate table is presented as a spreadsheet, enabling the preparation of the commonly ordered concentrations of vasoactive drugs at varying dose ranges based on the patient's weight (Figure 2). To calculate the composition of the drug infusion, one first selects the desired drug dosage to be delivered in µg/kg/min (first column). The next step is to select the initial infusion rate in mL/h (center of the table). Finally, one calculates the milligrams of drug based on the weight of the patient to be diluted with compatible fluids (sodium chloride 0.9%, etc) in a total volume of 50 mL (first row).

The study is formatted according to the Consolidated Standards of Reporting Trials Statement for Randomized Controlled Trials of Electronic and Mobile Health Applications and Online TeleHealth (CONSORT-EHEALTH) [17] and the Reporting Guidelines for Health Care Simulation Research [18].

Figure 1. Study design and flowchart per study center.



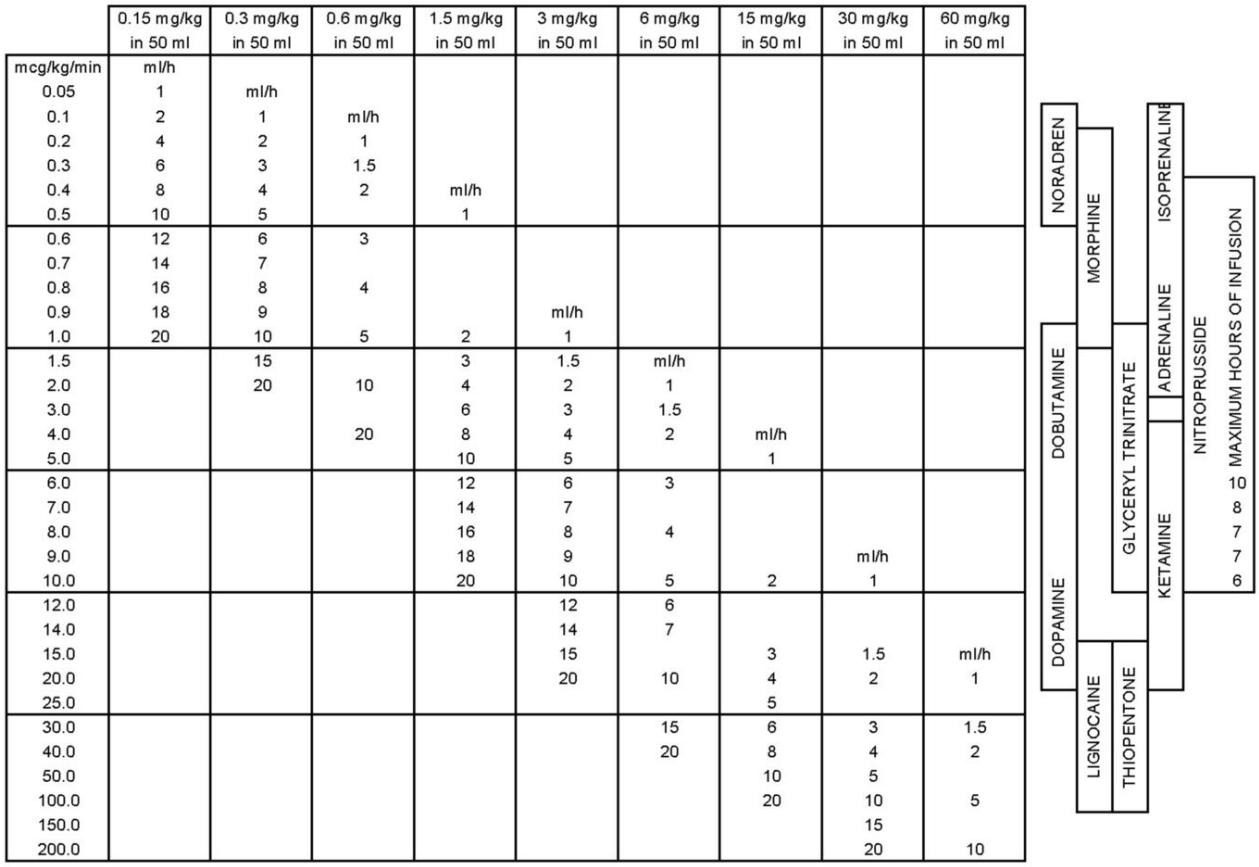
Selection of Participants

Certified nurses are eligible for inclusion in this study. Participants must know how to prepare continuous intravenous drug injections, have previously completed a standardized 5-minute introductory course to the use of PedAMINES dispensed by an investigator of the study (SM), and be willing and able to grant written informed consent. Written informed consent will be obtained from all the participants before their voluntary involvement. Participants will be excluded if they have previously used a numeric device aimed at helping vasoactive drug preparation. Shift-working nurses will be randomly recruited on the day of the study by a blinded noninvestigator.

Setting

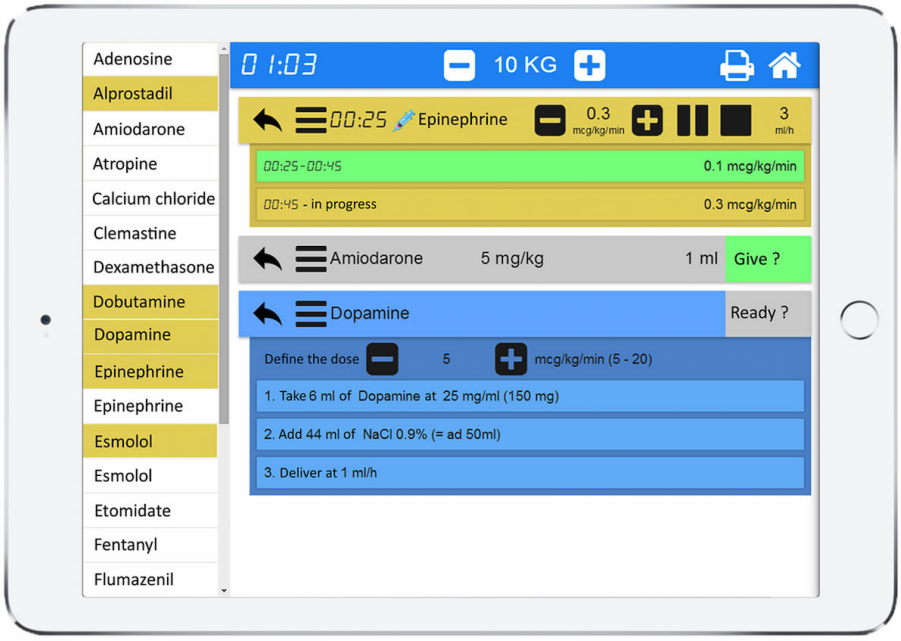
The study will be conducted in 3 university hospitals and 3 smaller hospitals with a total of approximately 150,000 visits per year. The development of PedAMINES followed a user-centered and evidence-based [12] approach with emergency department (ED) caregivers as well as software developers and ergonomists. Based on pediatric resuscitation observations and focus groups, the team worked closely together to identify the key functionalities and processes to be implemented [13]. The PedAMINES app lists all the available resuscitation drugs with doses automatically adapted to the weight or age of the patient based on information entered when starting the app. At the time of the study, 15 drugs for continuous infusion and 19 drugs for direct IV injection will be listed in the PedAMINES app and at the nurse's disposal. By a simple touch, any of the listed drugs can be selected and preparation detailed according to a standardized and simplified path. In the case of a continuous infusion, this path is composed of 3 steps: (1) drug selection, (2) dilution of the initial drug concentration, and (3) conversion of the prescribed dose rate in µg/kg/min into infusion pump rate in mL/h. For each drug, the exact amount to prepare is clearly displayed and thus avoids the necessity for calculations (Figure 3). This is based on the app's ability to automatically calculate the optimal weight-based final infusion pump rate and describe the preparation sequence required to achieve it independently of the nurse competency in this domain. The nurse may at any time interact with the app. The user can start, pause, stop, increase, or diminish the perfusion rate. Multiple drugs can be prepared and run in parallel. All actions performed by the nurses will be sequentially saved locally on the device in historic files to preserve information that can be retrieved at any time for debriefing or medicolegal purposes. Historic files can also be erased or safely exported and saved in institutional electronic health records.

Figure 2. Calculation of the Composition of Drug Infusions (Syringe Pump). Frank Shann Drug Doses [15].



mcg/kg/min	0.15 mg/kg in 50 ml	0.3 mg/kg in 50 ml	0.6 mg/kg in 50 ml	1.5 mg/kg in 50 ml	3 mg/kg in 50 ml	6 mg/kg in 50 ml	15 mg/kg in 50 ml	30 mg/kg in 50 ml	60 mg/kg in 50 ml
0.05	1	ml/h							
0.1	2	1	ml/h						
0.2	4	2	1						
0.3	6	3	1.5						
0.4	8	4	2	ml/h					
0.5	10	5		1					
0.6	12	6	3						
0.7	14	7							
0.8	16	8	4						
0.9	18	9			ml/h				
1.0	20	10	5	2	1				
1.5		15		3	1.5	ml/h			
2.0		20	10	4	2	1			
3.0				6	3	1.5			
4.0			20	8	4	2	ml/h		
5.0				10	5		1		
6.0				12	6	3			
7.0				14	7				
8.0				16	8	4			
9.0				18	9			ml/h	
10.0				20	10	5	2	1	
12.0					12	6			
14.0					14	7			
15.0					15		3	1.5	ml/h
20.0					20	10	4	2	1
25.0							5		
30.0						15	6	3	1.5
40.0						20	8	4	2
50.0							10	5	
100.0							20	10	5
150.0								15	
200.0								20	10

Figure 3. PedAMINES screenshot.



Intervention

The scenario in this study is a short, approximately 15-minute, standardized highly realistic pediatric resuscitation simulation on a high-fidelity manikin (Laerdal SimJunior). SimJunior represents a realistic 6-year-old boy and simulates a wide range of conditions from a healthy, talking child to an unresponsive, critical patient with no vital signs. This simulator provides all functionalities that are relevant to assess the research questions, including chest compression and ventilation and vascular access, as well as realistic and interactive vital signs. To date, high-fidelity simulation has become essential to study resuscitations skills and technologies that cannot be practiced during CPR because interindividual diversity among patients and their diseases make CPR studies hard to standardize in critical situations [19]. Moreover, by standardizing the scenario and environment, we will avoid effect modifiers by preventing the influence of undesired variables on the outcomes.

Consistent with standard emergency medicine practice, we will create resuscitation teams. Two members from the study team, remaining the same along the whole study period, will assist the study participant. A Pediatric Advanced Life Support (PALS) instructor-certified emergency pediatrician will lead the resuscitation, and a nurse will assist with resuscitation by performing chest compressions and bag-valve mask ventilation according to the pediatrician instructions. Participants will be informed before the scenario starts that these 2 people are study team members. Both study team members will guide each participant through a series of predefined key steps, blinded to the participant, following a standardized resuscitation scenario (Multimedia Appendix 1). The physician will order the medications and allow progression through the scenario only once predefined milestones have been reached, regardless of the occurrence of errors or time to achieve them. The study-specific training and standardization of both study team members is ensured through their previous involvement in the pilot study [14] and by following the predefined scenario. A certified technician will operate the simulator. The study will take place during a 1-week period in every hospital and will occupy each nurse for a single 30-minute period. Thus, we anticipate a very low rate of drop-outs or loss of follow-up. To ensure the presence of participants on the day of participation, shift-working nurses will be randomly recruited 1 month before the start of the study by a blinded noninvestigator. They will be informed of the upcoming simulation study but not of its purpose and outcomes.

On the day of participation, nurses will complete a survey collecting data regarding their demographics, nursing training, and simulation and computer experience. After random allocation, each participating nurse will receive a standardized 5-minute training session on how to use the PedAMINES app. Nurses will remain unaware of specific endpoints of the study during this learning session. The simulation manikin characteristics are then presented. The nurses will then be asked to perform the pediatric resuscitation scenario, including post-ROSC. This scenario is standardized to follow the same chronological progression and range of difficulty to ensure each participant is exposed to exactly the same case, with similar challenges in decision making and treatment preparation provided on the same manikin. The uniform delivery of the scenario all along the study will minimize confounders. Study team members will only adapt to progression speed of participants through the scenario by maintaining a stressful resuscitation atmosphere. The scenario will be videorecorded and conducted in situ in the pediatric resuscitation rooms of each collaborating center to increase realism. High levels of realism are known to immerse participants in the simulated experience and prevent confounding variables that might potentially affect the way individuals perform [19]. On the day of participation, the resuscitation room will be exclusively devoted to the simulation to prevent unexpected interruption or external stimuli. Monitoring alarms will be activated to increase realism. All resuscitation equipment in the resuscitation room will be at the disposal of the nurse. In both allocation groups, the decision to use or not use any equipment will remain personal as in real life. Neither pilot testing nor repetitions will be permitted. There will be neither interventions nor educational adjuncts prior to or after the study period. Once the experiment is completed, the nurse will be required to recall and describe precisely how she prepared dopamine and norepinephrine in order to verify that the drug names and original doses prescribed by the physician were correctly understood to assess the presence of comprehension bias. Then, a quick oral debriefing, shorter than 5 minutes, on how the drugs were actually prepared and how they should have been prepared without errors will be provided to each participant by one of the study team members. The debriefing will focus on both preparation methods.

Outcome Measures

The primary outcome of the study is the medication dosage error rate in each allocation group during the sequence from drug preparation to drug injection. A secondary outcome is the elapsed time in seconds, in each allocation group, between the oral prescription by the physician and drug delivery by the nurse. TDP completion by the nurse is included in TDD. Comparisons of medication errors and TDD between allocation groups in each hospital and between each of the 6 hospitals will be assessed as secondary outcomes. At the end of the scenario, a 4-item questionnaire using a 10-point Likert scale will be provided to the participants. The questionnaire measures (1) the overall stress perceived (On a scale of 1 to 10, how much stress did you feel during the whole resuscitation scenario?), (2) the stress perceived with PedAMINES (On a scale of 1 to 10, how much stress did you feel using PedAMINES?), (3) the stress perceived with the infusion rate table (On a scale of 1 to 10, how much stress did you feel using the infusion rate table?), and (4) the satisfaction about the preparation method used during the resuscitation scenario (On a scale of 1 to 10, how much satisfaction did you get during the resuscitation scenario with the help of PedAMINES and with the help of the infusion rate table?). A baseline heart rate (HR) before the beginning of the scenario as well as continuous HR monitoring of each participant reflecting his or her stress level will be recorded during the resuscitation scenario. Mean delta HR values (difference between HR peak values and baseline HR) will be obtained during some small segments of scenario and correlated to the scenario phases and the preparation methods used. The segments of interest are (1) when prompted to start the resuscitation

because of asystole just before massage and ventilation, (2) when prompted to prepare dopamine either with PedAMINES or the infusion rate table, (3) the first 30 seconds when dopamine is being prepared, (4) the last 30 seconds when dopamine is being prepared, (5) the setting up of the pump, (6) the washout period, (7) when prompted to prepare norepinephrine either with PedAMINES or the infusion rate table, (8) the first 30 seconds when norepinephrine is being prepared, (9) the last 30 seconds when norepinephrine is being prepared, (10) at announcement of successful resuscitation achievement, and (11) 1 minute later. Acceptability and usability testing of the app will be assessed using a 52-item questionnaire based on the unified theory of acceptance and use of technology (UTAUT) model [20]. UTAUT provides a useful tool to assess the likelihood of success for new technology introductions and helps to understand the drivers of its acceptance.

Methods of Measurement and Data Collection

Final delivered drug concentration in µg/mL and infusion rate in µg/kg/min, TDP, and TDD will be collected during the scenario. All the actions (ie, primary and secondary outcomes) performed by the nurses during the scenario will be automatically recorded and stored by the responsive simulator detectors (Leardal SimJunior) and by 3 GoPro Hero 5 Black edition (GoPro Inc) video cameras. To avoid assessment bias, 2 evaluators will then independently review these video recordings. In case of disagreement, a third independent evaluator will help reach a consensus. The setup of the 3 cameras will be standardized to record at a resolution of 1080p, at 25 frames per second, wide field of view, and with a 16:9 aspect ratio. Cameras position will be standardized. The first camera will be placed in a harness on the nurse's chest with an inclination of 10° downwards to film the front scene. The second camera will be placed on a tripod in front of the nurse, slightly above the head height, with an inclination of 90° downwards to film the bench where the drugs will be prepared. The third camera will be placed on a tripod 1 meter away from the nurse on his or her left (if right-handed) or right (if left-handed) and at the level of the navel to film the scene on the side. All actions performed with PedAMINES will be automatically saved locally in log files for further analysis. The validity and reliability of the app has been assessed in a prior study [14]. Stress level of each participant will be recorded during the entire resuscitation scenario with the HR monitor on a Polar A360 watch (Polar Electro Oy). The data will be stored on the wristwatch itself with further analysis being accomplished offline. Data collection will be carried out using Excel spreadsheet version 2011 (Microsoft Corp).

This study offers the great advantage of being very short in duration, around 30 minutes per nurse. Therefore, neither follow-up nor retention plans will be necessary. The intervention protocol is highly standardized, and nurse deviation from protocol in terms of drug preparation is a parameter that is of interest in our study (ie, in terms of medication errors or delays in drug preparation).

Sample Size

The primary objective of this study is to detect a difference in rates of medication errors between groups (with PedAMINES vs without PedAMINES). In our crossover pilot study conducted in a single university hospital, the rate of medication errors was 70% (14/20, 95% CI 46%-88%) without PedAMINES and 0% (0/20, 95% CI 0%-17%) with PedAMINES [14]. We aim to reproduce these results in 3 university hospitals and 3 smaller hospitals. The sample size was calculated to have a power of 90% to detect a difference of at least 30% in rates of medication error between intervention groups in both sets of hospitals (university and smaller hospitals). We consider that a difference of 30% in the rate of medication errors is sufficient to modify the practice. We assumed a rate of error of 15% with PedAMINES and 45% without PedAMINES. Since this study is a crossover, observations will be paired and a McNemar test will be used to compare the rate of errors between intervention groups. For the sample size calculation, we assumed that 5% of nurses will commit a medication error with PedAMINES but not without PedAMINES and that 35% of nurses will commit a medication error without PedAMINES but not with PedAMINES (Multimedia Appendix 2). With a 2-sided risk alpha of 0.05, the needed sample size is 43 nurses for each set of hospitals. Finally, 8 nurses per randomized group must be recruited in each participating center (16 nurses per center). To prevent a potential loss of power due to misspecification of assumptions, 10 nurses will be recruited per randomized group and per participating center (total sample size: 120 nurses).

Randomization and Blinding

Nurses will be randomized using a stratified, single, constant 1:1 allocation ratio determined with Web-based software [21]. Blinding to the purpose of the study during recruitment will be maintained to minimize preparation bias. Nurses will be unblinded after randomization. Allocation concealment will be ensured with sealed envelopes and will not be released until the nurses start the scenario. The study team members will be revealed just before the scenario starts, and videoreviewing will be done without blinding by both video reviewers but independently and blindly from one another.

Statistical Analysis

Primary Outcome

The rate of medication errors is the proportion of nurses committing a medication error. The rate of medication errors will be reported with 95% CI (Clopper-Pearson method) with each method and by study period to investigate a potential carryover effect. The error rates for each method will be compared using McNemar test for paired data. Differences in error rates will be reported with 95% CIs. Potentially, the efficacy of PedAMINES can be different according to the first method used in the crossover design (PedAMINES first: group A or the conventional preparation method first: group B). To investigate this potential effect, the difference in error rates (PedAMINES vs conventional method) will also be reported by randomized group (paired observations) and by period (independent observations). McNemar, chi-square, or Fisher exact test will be used to compare interventions. Errors will also be measured as the deviation in percent from the amount of delivered drug compared with the original dose prescribed by the physician. Absolute deviation will be analyzed. The mean (SD) difference in deviation obtained with each method will be

reported with 95% CI. A *t* test for paired data will be used to compare interventions. Mean differences will be also reported by randomized group and by crossover period. We will assess the model for the presence of carryover and period effects following the Hills-Armitage approach to crossover study analysis [22]. The analysis of the primary outcome will be conducted for the university hospitals and for smaller hospitals.

Secondary Outcomes

For TDD and TDP, the mean times will be reported with 95% CI for each arm and each study period to investigate a potential carryover effect. In case of a carryover effect, intervention arms will be compared within each study period using *t* tests for independent groups. If no carryover effect is found, data will be paired and analyzed using *t* tests for paired data.

For the primary and secondary outcomes, logistic regression analyses will be conducted to test a difference in rates of errors between university hospitals and smaller hospitals with PedAMINES and with the conventional preparation method. In a generalized estimating equation logistic regression model, an interaction between interventions and university/smaller hospitals will be tested to investigate a potential modification of the efficacy of PedAMINES in smaller hospitals compared with university hospitals. Analyses of primary and secondary outcomes will be conducted in nurses with more than 10 years of experience and in nurses with less than 10 years of experience. Means and standard deviations will be determined for stress and satisfaction scores of individuals for each questionnaire item as well as for the UTAUT questionnaire and reported with descriptive statistics. Pearson correlations will be computed between the HR measures obtained with the watch and the scenario phases for each of the preparation methods used. Data analysis will be carried out using GraphPad Prism version 6.0 (GraphPad Software, Inc) for graph figures, Stata/IC version 14 (StataCorp LLC) for descriptive analyses, and R version 2.15.2 (R Foundation) for statistical tests and 95% CI. Due to the nature of the interventions, we expect to have no missing data. In case of missing data, a complete case analysis will be conducted and no multiple imputation is planned. In order to assess interrater reliability on video reviewing, a kappa score—which provides a measure of interrater agreement independent of chance—will be calculated.

Ethics and Informed Consent

Ethical approval has been obtained from the institutional ethics committee, and the trial was registered at ClinicalTrials.gov [NCT03021122]. The study will be conducted in accordance with the principles of the Declaration of Helsinki, the standards of Good Clinical Practice, and Swiss regulatory requirements.

Results

Enrollment and data analysis started in March 2017. We anticipate that the intervention will be complete in late 2017, and study results will be submitted in early 2018 for publication expected in mid-2018. Results will be reported in line with recommendations in the CONSORT-EHEALTH [17] and the Reporting Guidelines for Health Care Simulation Research [18].

Discussion

Principal Findings

Despite many advances in the medical field in recent years and in particular in emergency medicine, suboptimal quality of resuscitation is still common for both adult and pediatric patients [23]. Currently, the median hospital survival rate from pediatric in-hospital cardiopulmonary arrest is 36% [23], whereas it is below 10% for out-of-hospital cardiopulmonary arrest [24,25]. Survival from resuscitation is time-sensitive and relies in part on administration of certain drugs without delay [10]. While many drugs can be directly injected without prior preparation, others require correct, precise, and fast preparation and dilution by nurses before administration as a continuous infusion. The latter warrants titration of the drugs and the maintenance of steady blood levels within therapeutic ranges. However, despite the availability of conversion methods intended to simplify the infusions such as infusion rate tables or nomograms [26], these methods remain difficult to use and infusions subject to medication errors.

In a recent randomized trial in a pediatric ED, the use of a reference book providing weight-based precalculated doses was associated with a lower proportion of prescribing errors for drugs administered by infusion [27]. The study, however, didn't look at the preparation errors. Errors with infusions frequently result from mistakes during preparation due to wrong drug-volume calculations, imprecision with volume measurements, or incorrect mixing during dilution [1,28,29]. Children are particularly at risk for such medication errors because they require individual specific weight-based drug dose calculation and preparation. Moreover, disruptive anxiety and exogenous conditions encountered during resuscitation increase the nurse's cognitive workload and may add to the risk of errors. At this stage, even small errors either in drug calculation or infusion pump flow rate may have a large detrimental impact on the amount of drug delivered [30-32]. This can be deleterious to critically ill and unstable patients [7].

Being able to reduce medication errors and TDD in resuscitation is imperative. If ROSC is quickly achieved and maintained after the onset of cardiac arrest, survival might be improved [33] since early hemodynamic optimization improves patient outcome [34,35].

Some authors have advocated replacing as much as possible tasks inducing cognitive load during pediatric resuscitation by automated actions in order to optimize patient care and diminish medication errors [36,37]. Numerous interventions involving information technologies have been developed to improve the security of the 3 major steps of medication process: prescription, compounding, and administration [38]. However, apart from computerized physician order entry systems, few robust data are available to measure their real impact on patient safety [39]. In addition, there have been few studies assessing information technologies during resuscitation in both adults and pediatric patients. In particular, there has been no multicenter randomized controlled trial evaluating the impact of a mobile app to reduce medication errors and decrease TDP and TDD during pediatric resuscitation. As research in this area is scarce, it is anticipated

that the results generated from this study will therefore be of great importance and might be sufficient to change and improve the pediatric emergency care practice. Given that most of the results obtained from simulation-based resuscitation studies agree with those obtained from studies in real life, we are confident that PedAMINES could be of great interest for real situations.

Limitations

A limitation of this study is that it will be conducted during a resuscitation simulation-based scenario. This choice was related to the ethical and organizational difficulties of conducting studies with patients in critical situations. However, several studies have demonstrated the benefit of simulation as an investigative research methodology to answer research questions that otherwise could not be answered during resuscitation [19]. Simulation-based CPR scenarios may overcome these limitations by providing a standardized and controlled environment, detailed feedback analysis of the resuscitation stages using audiovisual recordings, and reproducibility. To date, high-fidelity simulation has become essential to study resuscitations skills or technologies.

Our study is not intended to compare PedAMINES with smart IV pumps. As recently reported, there is no conclusive evidence showing that smart pumps prevent medication errors and adverse drug events [40,41]. In addition, little is known about the kind of errors that still occur with their use. Moreover, the lack of specialized pharmacy facilities in many smaller hospitals or in other countries around the world limit their use. PedAMINES does not have these limitations for use and can be used in smaller hospitals and worldwide.

It should be noted that the Likert-type questionnaire used in this study to measure stress has not been assessed for validity, internal consistency, reliability, or generalizability. Although it cannot objectively measure the stress perceived, it can be used to measure the difference of perceived stress.

Acknowledgments

The proposed trial was awarded external peer-reviewed funding by the Swiss National Science Foundation (Switzerland) on September 21, 2016 (number 32003B_169348/1).

Authors' Contributions

All authors contributed to the design and coordination of the study and have read, commented on, and approved the study protocol manuscript.

Conflicts of Interest

JNS, FE, CL, AG, and SM are the owners of the PedAMINES app that will be available in the near future in Google Play and the App Store. The authors therefore declare a direct financial interest to market this app.

Multimedia Appendix 1

Detailed description of the resuscitation scenario used during the simulation.

[PDF File (Adobe PDF File), 60KB - resprot_v6i8e167_app1.pdf]

Multimedia Appendix 2

Contingency table for the expected McNemar test comparing the rate of errors between intervention groups. Percentages in bold denote the values hypothesized for the sample size calculation.

[JPG File, 497KB - resprot_v6i8e167_app2.jpg]

Multimedia Appendix 3

Peer-review document: Acceptance letter by the Swiss National Science Foundation.

[PDF File (Adobe PDF File), 101KB - resprot_v6i8e167_app3.pdf]

References

1. Larsen GY, Parker HB, Cash J, O'Connell M, Grant MC. Standard drug concentrations and smart-pump technology reduce continuous-medication-infusion errors in pediatric patients. Pediatrics 2005 Jul;116(1):e21-e25. [doi: 10.1542/peds.2004-2452] [Medline: 15995017]
2. Moyen E, Camiré E, Stelfox HT. Clinical review: medication errors in critical care. Crit Care 2008;12(2):208 [FREE Full text] [doi: 10.1186/cc6813] [Medline: 18373883]

3. Polischuk E, Vetterly CG, Crowley KL, Thompson A, Goff J, Nguyen-Ha P, et al. Implementation of a standardized process for ordering and dispensing of high-alert emergency medication infusions. J Pediatr Pharmacol Ther 2012 Apr;17(2):166-172 [FREE Full text] [doi: 10.5863/1551-6776-17.2.166] [Medline: 23118669]
4. Kaushal R, Bates DW, Landrigan C, McKenna KJ, Clapp MD, Federico F, et al. Medication errors and adverse drug events in pediatric inpatients. JAMA 2001 Apr 25;285(16):2114-2120. [Medline: 11311101]
5. Gonzales K. Medication administration errors and the pediatric population: a systematic search of the literature. J Pediatr Nurs 2010 Dec;25(6):555-565. [doi: 10.1016/j.pedn.2010.04.002] [Medline: 21035020]
6. Hoyle JD, Davis AT, Putman KK, Trytko JA, Fales WD. Medication dosing errors in pediatric patients treated by emergency medical services. Prehosp Emerg Care 2012;16(1):59-66. [doi: 10.3109/10903127.2011.614043] [Medline: 21999707]
7. Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Dtsch Arztebl Int 2012 Sep;109(38):609-616 [FREE Full text] [doi: 10.3238/arztebl.2012.0609] [Medline: 23093991]
8. Porter E, Barcega B, Kim TY. Analysis of medication errors in simulated pediatric resuscitation by residents. West J Emerg Med 2014 Jul;15(4):486-490 [FREE Full text] [doi: 10.5811/westjem.2014.2.17922] [Medline: 25035756]
9. Matos RI, Watson RS, Nadkarni VM, Huang H, Berg RA, Meaney PA, et al. Duration of cardiopulmonary resuscitation and illness category impact survival and neurologic outcomes for in-hospital pediatric cardiac arrests. Circulation 2013 Jan 29;127(4):442-451. [doi: 10.1161/CIRCULATIONAHA.112.125625] [Medline: 23339874]
10. Andersen LW, Berg KM, Saindon BZ, Massaro JM, Raymond TT, Berg RA, et al. Time to epinephrine and survival after pediatric in-hospital cardiac arrest. JAMA 2015 Aug 25;314(8):802-810. [doi: 10.1001/jama.2015.9678] [Medline: 26305650]
11. Hubble MW, Johnson C, Blackwelder J, Collopy K, Houston S, Martin M, et al. Probability of return of spontaneous circulation as a function of timing of vasopressor administration in out-of-hospital cardiac arrest. Prehosp Emerg Care 2015;19(4):457-463. [doi: 10.3109/10903127.2015.1005262] [Medline: 25909945]
12. Ehrler F, Haller G, Sarrey E, Walesa M, Wipfli R, Lovis C. Assessing the usability of six data entry mobile interfaces for caregivers: a randomized trial. JMIR Hum Factors 2015 Dec 15;2(2):e15 [FREE Full text] [doi: 10.2196/humanfactors.4093] [Medline: 27025648]
13. Hagberg H, Siebert J, Gervaix A, Daehne P, Lovis C, Manzano S, et al. Improving drugs administration safety in pediatric resuscitation using mobile technology. Stud Health Technol Inform 2016;225:656-657. [Medline: 27332296]
14. Siebert JN, Ehrler F, Combescure C, Lacroix L, Haddad K, Sanchez O, et al. A mobile device app to reduce time to drug delivery and medication errors during simulated pediatric cardiopulmonary resuscitation: a randomized controlled trial. J Med Internet Res 2017 Feb 01;19(2):e31 [FREE Full text] [doi: 10.2196/jmir.7005] [Medline: 28148473]
15. Shann F. Drug Doses RCH Intensive Care Unit, 16th Edition. Melbourne: The Royal Children's Hospital; 2014.
16. McLeroy PA. The rule of six: calculating intravenous infusions in a pediatric crisis situation. Hosp Pharm 1994 Oct;29(10):939-943. [Medline: 10137852]
17. Eysenbach G. 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]
18. Cheng A, Kessler D, Mackinnon R, Chang TP, Nadkarni VM, Hunt EA, International Network for Simulation-based Pediatric Innovation, Research,Education (INSPIRE) Reporting Guidelines Investigators. Reporting guidelines for health care simulation research: extensions to the CONSORT and STROBE statements. Simul Healthc 2016 Aug;11(4):238-248. [doi: 10.1097/SIH.0000000000000150] [Medline: 27465839]
19. Cheng A, Auerbach M, Hunt EA, Chang TP, Pusic M, Nadkarni V, et al. Designing and conducting simulation-based research. Pediatrics 2014 Jun;133(6):1091-1101 [FREE Full text] [doi: 10.1542/peds.2013-3267] [Medline: 24819576]
20. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478 [FREE Full text]
21. sealedenvelope.com. URL: http://www.sealedenvelope.com [accessed 2017-04-14] [WebCite Cache ID 6pj1kvMba]
22. Hills M, Armitage P. The two-period cross-over clinical trial. Br J Clin Pharmacol 1979 Jul;8(1):7-20 [FREE Full text] [Medline: 552299]
23. Meaney PA, Bobrow BJ, Mancini ME, Christenson J, de Caen AR, Bhanji F, et al. Cardiopulmonary resuscitation quality: [corrected] improving cardiac resuscitation outcomes both inside and outside the hospital: a consensus statement from the American Heart Association. Circulation 2013 Jul 23;128(4):417-435 [FREE Full text] [doi: 10.1161/CIR.0b013e31829d8654] [Medline: 23801105]
24. Jayaram N, McNally B, Tang F, Chan PS. Survival after out-of-hospital cardiac arrest in children. J Am Heart Assoc 2015 Oct 08;4(10):e002122 [FREE Full text] [doi: 10.1161/JAHA.115.002122] [Medline: 26450118]
25. Kämäräinen A. Out-of-hospital cardiac arrests in children. J Emerg Trauma Shock 2010 Jul;3(3):273-276 [FREE Full text] [doi: 10.4103/0974-2700.66531] [Medline: 20930972]
26. Sherry E, Burton GW, Wilkins DG. Infusion nomograms. Anaesthesia 1993 May;48(5):396-401 [FREE Full text] [Medline: 8317648]
27. Larose G, Levy A, Bailey B, Cummins-McManus B, Lebel D, Gravel J. Decreasing prescribing errors during pediatric emergencies: a randomized simulation trial. Pediatrics 2017 Mar;139(3). [doi: 10.1542/peds.2016-3200] [Medline: 28246338]

28. Adapa RM, Mani V, Murray LJ, Degnan BA, Ercole A, Cadman B, et al. Errors during the preparation of drug infusions: a randomized controlled trial. Br J Anaesth 2012 Nov;109(5):729-734 [FREE Full text] [doi: 10.1093/bja/aes257] [Medline: 22850220]
29. Ameer A, Dhillon S, Peters M, Ghaleb M. Systematic literature review of hospital medication administration errors in children. IPRP 2015 Nov:153. [doi: 10.2147/IPRP.S54998]
30. Allen EM, Van Boerum DH, Olsen AF, Dean JM. Difference between the measured and ordered dose of catecholamine infusions. Ann Pharmacother 1995 Nov;29(11):1095-1100. [Medline: 8573951]
31. Lehmann CU, Kim GR, Gujral R, Veltri MA, Clark JS, Miller MR. Decreasing errors in pediatric continuous intravenous infusions. Pediatr Crit Care Med 2006 May;7(3):225-230. [doi: 10.1097/01.PCC.0000216415.12120.FF] [Medline: 16575355]
32. Parshuram CS, To T, Seto W, Trope A, Koren G, Laupacis A. Systematic evaluation of errors occurring during the preparation of intravenous medication. CMAJ 2008 Jan 1;178(1):42-48 [FREE Full text] [doi: 10.1503/cmaj.061743] [Medline: 18166730]
33. Neumar RW, Nolan JP, Adrie C, Aibiki M, Berg RA, Böttiger BW, et al. Post-cardiac arrest syndrome: epidemiology, pathophysiology, treatment, and prognostication. A consensus statement from the International Liaison Committee on Resuscitation (American Heart Association, Australian and New Zealand Council on Resuscitation, European Resuscitation Council, Heart and Stroke Foundation of Canada, InterAmerican Heart Foundation, Resuscitation Council of Asia, and the Resuscitation Council of Southern Africa); the American Heart Association Emergency Cardiovascular Care Committee; the Council on Cardiovascular Surgery and Anesthesia; the Council on Cardiopulmonary, Perioperative, and Critical Care; the Council on Clinical Cardiology; and the Stroke Council. Circulation 2008 Dec 2;118(23):2452-2483 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.108.190652] [Medline: 18948368]
34. López-Herce J, del Castillo J, Matamoros M, Canadas S, Rodriguez-Calvo A, Cecchetti C, et al. Post return of spontaneous circulation factors associated with mortality in pediatric in-hospital cardiac arrest: a prospective multicenter multinational observational study. Crit Care 2014 Nov 03;18(6):607 [FREE Full text] [doi: 10.1186/s13054-014-0607-9] [Medline: 25672247]
35. Topjian AA, French B, Sutton RM, Conlon T, Nadkarni VM, Moler FW, et al. Early postresuscitation hypotension is associated with increased mortality following pediatric cardiac arrest. Crit Care Med 2014 Jun;42(6):1518-1523 [FREE Full text] [doi: 10.1097/CCM.0000000000000216] [Medline: 24561563]
36. Luten R, Wears RL, Broselow J, Croskerry P, Joseph MM, Frush K. Managing the unique size-related issues of pediatric resuscitation: reducing cognitive load with resuscitation aids. Acad Emerg Med 2002 Aug;9(8):840-847 [FREE Full text] [Medline: 12153892]
37. Moreira ME, Hernandez C, Stevens AD, Jones S, Sande M, Blumen JR, et al. Color-coded prefilled medication syringes decrease time to delivery and dosing error in simulated emergency department pediatric resuscitations. Ann Emerg Med 2015 Aug;66(2):97-106 [FREE Full text] [doi: 10.1016/j.annemergmed.2014.12.035] [Medline: 25701295]
38. Bonnabry P. Information technologies for the prevention of medication errors. CHIMIA Int J Chem 2005;59(6):359-361.
39. Acheampong F, Anto BP, Koffuor GA. Medication safety strategies in hospitals—a systematic review. Int J Risk Saf Med 2014;26(3):117-131. [doi: 10.3233/JRS-140623] [Medline: 25214157]
40. Franklin BD. 'Smart' intravenous pumps: how smart are they? BMJ Qual Saf 2016 Apr 7. [doi: 10.1136/bmjqs-2016-005302] [Medline: 27055758]
41. Schnock KO, Dykes PC, Albert J, Ariosto D, Call R, Cameron C, et al. The frequency of intravenous medication administration errors related to smart infusion pumps: a multihospital observational study. BMJ Qual Saf 2016 Feb 23. [doi: 10.1136/bmjqs-2015-004465] [Medline: 26908900]

Abbreviations

CONSORT-EHEALTH: Randomized Controlled Trials of Electronic and Mobile Health Applications and Online TeleHealth
CPR: cardiopulmonary resuscitation
ED: emergency department
HR: heart rate
IV: intravenous
PALS: Pediatric Advanced Life Support
PedAMINES: Pediatric Accurate Medication in Emergency Situations
ROSC: return of spontaneous circulation
TDD: time to drug delivery
TDP: time to drug preparation
UTAUT: unified theory of acceptance and use of technology

Edited by G Eysenbach; submitted 21.04.17; peer-reviewed by T Chang, J Adjedj; comments to author 22.06.17; revised version received 05.08.17; accepted 09.08.17; published 22.08.17.

Please cite as:
Siebert JN, Ehrler F, Lovis C, Combescure C, Haddad K, Gervaix A, Manzano S
A Mobile Device App to Reduce Medication Errors and Time to Drug Delivery During Pediatric Cardiopulmonary Resuscitation: Study Protocol of a Multicenter Randomized Controlled Crossover Trial
JMIR Res Protoc 2017;6(8):e167
URL: http://www.researchprotocols.org/2017/8/e167/
doi:10.2196/resprot.7901
PMID:28830854

Protocol

Medicine Goes Female: Protocol for Improving Career Options of Females and Working Conditions for Researching Physicians in Clinical Medical Research by Organizational Transformation and Participatory Design

Joachim Hasebrook[1], PhD; Klaus Hahnenkamp[2], MD; Wolfgang F.F.A Buhre[3], MD; Dianne de Korte-de Boer[3], PhD; Ankie E.W Hamaekers[3], PhD; Bibiana Metelmann[2], MD; Camila Metelmann[2], MD; Marina Bortul[4], MD; Silvia Palmisano[4], MD; Jannicke Mellin-Olsen[5], MD, DPH; Andrius Macas[6], MD; Janusz Andres[7], MD; Anna Prokop-Dorner[8], MA; Tomáš Vymazal[9], MD; Juergen Hinkelmann[10], MD; Sibyll Rodde[11], PhD; Bettina Pfleiderer[12], MD, PhD

[1]zeb.business school, Steinbeis University Berlin, Muenster, Germany

[2]Clinic for Anesthesiology, University Medicine Greifswald, Greifswald, Germany

[3]Department of Anaesthesia & Pain therapy, Maastricht UMC+, Maastricht, Netherlands

[4]General Surgery Clinic, Department of Medical, Surgical and Health Sciences, University Hospital of Trieste, Trieste, Italy

[5]Bærum Hospital, Department of Anaesthesia and Intensive Care Medicine, Vestre Viken Health Trust, Vestre Viken, Norway

[6]Department of Anaesthesiology, Lithuanian University of Health Sciences, Kaunas, Lithuania

[7]University Hospital, Clinic for Anasethesiology, Uniwersytet Jagiellonski, Cracow, Poland

[8]Chair of Epidemiology and Preventive Medicine, Department of Medical Sociology, Jagiellonian University Medical College, Cracow, Poland

[9]Clinic for Anaesthesiology, University Hospital Praha, Praha, Czech Republic

[10]Board of the University Hospital Frankfurt, University Hospital Frankfurt, Frankfurt, Germany

[11]zeb.health care, zeb, Muenster, Germany

[12]Department of Clinical Radiology, Chair of the Research Group Cognition & Gender, University Hospital Muenster, Muenster, Germany

Corresponding Author:
Joachim Hasebrook, PhD
zeb.business school
Steinbeis University Berlin
Hammer Str. 165
Muenster, 48153
Germany
Phone: 49 251 97128 ext 940
Fax: 49 25197128102
Email: jhasebrook@zeb-bs.de

Abstract

Background: All European countries need to increase the number of health professionals in the near future. Most efforts have not brought the expected results so far. The current notion is that this is mainly related to the fact that female physicians will clearly outnumber their male colleagues within a few years in nearly all European countries. Still, women are underrepresented in leadership and research positions throughout Europe.

Objectives: The MedGoFem project addresses multiple perspectives with the participation of multiple stakeholders. The goal is to facilitate the implementation of Gender Equality Plans (GEP) in university hospitals; thereby, transforming the working conditions for women working as researchers and highly qualified physicians simultaneously. Our proposed innovation, a crosscutting topic in all research and clinical activities, must become an essential part of university hospital strategic concepts.

Methods: We capture the current status with gender-sensitive demographic data concerning medical staff and conduct Web-based surveys to identify cultural, country-specific, and interdisciplinary factors conducive to women's academic success. Individual expectations of employees regarding job satisfaction and working conditions will be visualized based on "personal construct theory" through repertory grids. An expert board working out scenarios and a gender topic agenda will identify culture-, nation-, and discipline-specific aspects of gender equality. University hospitals in 7 countries will establish consensus groups, which work

on related topics. Hospital management supports the consensus groups, valuates group results, and shares discussion results and suggested measures across groups. Central findings of the consensus groups will be prepared as exemplary case studies for academic teaching on research and work organization, leadership, and management.

Results: A discussion group on gender equality in academic medicine will be established on an internationally renowned open-research platform. Project results will be published in peer-reviewed journals with high-impact factors. In addition, workshops on gender dimension in research using the principles of Gendered Innovation will be held. Support and consulting services for hospitals will be introduced in order to develop a European consulting service.

Conclusions: The main impact of the project will be the implementation of innovative GEP tailored to the needs of university hospitals, which will lead to measurable institutional change in gender equality. This will impact the research at university hospitals in general, and will improve career prospects of female researchers in particular. Simultaneously, the gender dimension in medical research as an innovation factor and mandatory topic will be strengthened and integrated in each individual university hospital research activity. Research funding organizations can use the built knowledge to include mandatory topics for funding applications to enforce the use and implementation of GEP in university hospitals.

(JMIR Res Protoc 2017;6(8):e152) doi:10.2196/resprot.7632

KEYWORDS

gender equality; gender equality plan; research in academic medicine; working conditions; skills shortage

Introduction

Situation of Women in Medicine

Women are underrepresented in leadership and research positions in all European countries—albeit the extent of inequalities differs widely with the overall inequality (as expressed in The European Institute for Gender Equality's Gender Equality Index) [1]. Moreover, women remain underrepresented in EU research and science despite numerous attempts to address the imbalance, according to the European Union's analyses "She Figures: Gender in Research and Innovation" and well as "Report on Equality between women and men – 2015" [2,3]. The EU report on "Enhancing excellence, gender equality and efficiency in research and innovation" states that we need to change our approach from "... no longer fixing women but fixing institutions" [4].

The situation for female physicians in the medical environment is even worse in comparison with other fields, and most programs introduced to tackle gender inequalities [5,6] did not work as expected in medicine; in particular, in university hospitals. University hospitals are involved in the care of critically ill patients, pre- and postgraduate education, and last, but not least, medical research. Besides providing health care at the highest possible level, it is a necessity and privilege of university hospitals to carry out medical research in addition to teaching medical students to ensure excellent health care at present and in the future. In that vein, heads of medical departments of a university hospital are usually full professors at medical schools and chief physicians at the same time. In contrast to most other research areas where full-time scientists focus exclusively on their research, physicians at university hospitals hardly ever work full-time in research. Because university hospitals are subject to the same economic constraints as other hospitals without teaching and research responsibilities, most research activities are usually done after regular clinical work.

High-research productivity as measured by publication of scientific articles with high-impact levels is crucial for career advancement. However, research productivity of male researchers is often overrated and of females underrated [7]. For instance, it was shown that female researchers are less likely to be listed as first or last authors, and get published less [8]. In biomedical sciences, women get smaller grants than men in the United States, in addition to applying less for competitive grants in the life sciences [9]. In cases of female first authorship in combination with last authorship taken by her male thesis supervisor, it is automatically assumed that most of the work was done by the male supervisor [7]. This leads to frustration, and as a consequence diminished interest in research of young female physicians. Therefore, female doctors decide against a promising academic career to stronger focus on family goals. This heavily jeopardizes their career prospects to achieve leadership positions in a medical setting.

In addition to unequal research conditions between men and women, highly qualified (young) women [10] often do not find appropriate working conditions in mostly hierarchically structured university hospitals with their male-dominated management. Institutional structures involuntarily erect barriers against the recruitment, retention, and career progression of (young) women. Gendered working conditions remain firmly fixed, and this is even more challenging–overt discrimination has been replaced by less visible, mostly implicit stereotypes and prejudices against women (eg, women have less career motivations, assumed lack of confidence, and ambition to take on leadership positions) [11-13]. Moreover, having children is an additional "career stopper" for female physicians: those with children are less likely to be promoted and have a lower income [14]. To summarize, female physicians receive fewer promotions than their male colleagues, and on average earn less than their male colleagues for the same work [14,15]. As a consequence, young female physicians are less willing to work at university hospitals under the given conditions and are less interested in doing research. They decide to walk different paths because they cannot see how to combine a career and having a family at the same time.

Approaches to Reach Gender Equality

Three strategic approaches to gender equality over the past decades have been taken by governments and research organizations like universities [16]:

1. The “Fix the Number of Women” approach focuses on increasing women's participation.
2. The “Fix the Institutions” approach promotes gender equality in careers through structural change in research organizations [4].
3. The “Fix the Knowledge” (or “gendered innovations”) approach stimulates excellence in science and technology by removing “gender bias” by male-dominated researchers, and by a better integration of the gender dimension in basic life sciences research [5,16].

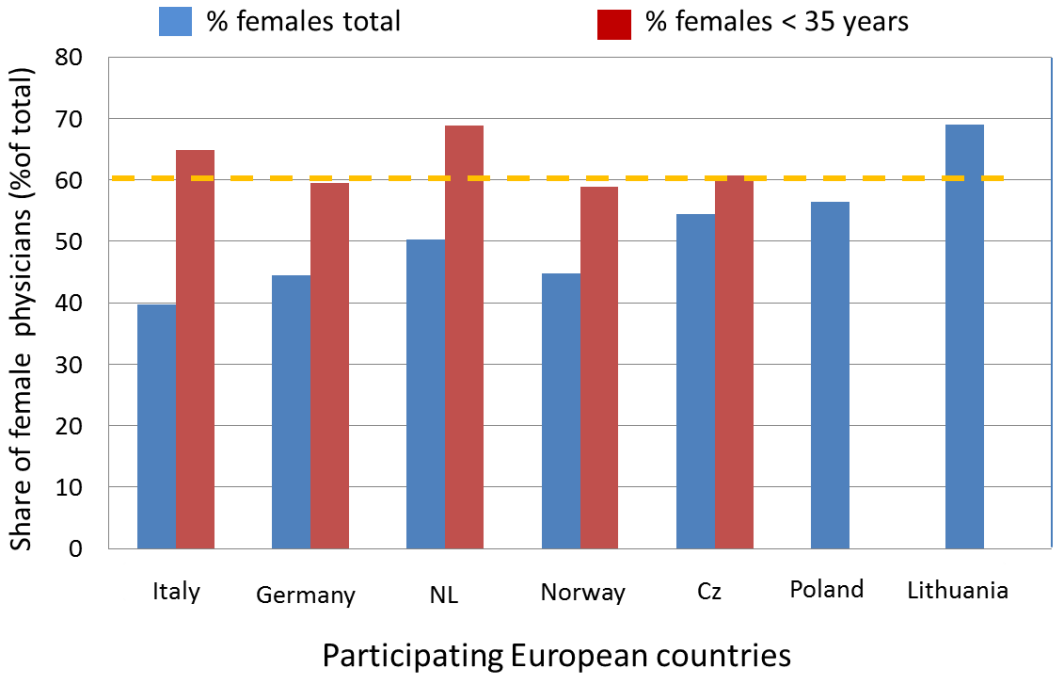
Most efforts targeted to women’s career needs are either top-down (eg, new flexibility policies [6]) or bottom-up (eg, research skills training for women faculty). Up to now, all these strategies were not successful to retain and keep women in the workforce at university hospitals and research.

Shortage of Health Professionals and Impact on European Health Care

Can society really afford to lose those highly skilled and well-trained female medical doctors? The aging of society and the rising prevalence of diseases related to sedentary lifestyles will increase the burden of disease by 10% to 15% until 2030. The growth of noncommunicable diseases is especially alarming [17]. Consequently, all European countries need to increase the number of health professionals in the near future.

Female physicians will clearly outnumber their male colleagues within a few years in nearly all European countries, because the vast majority of physicians younger than 35 years are female (Figure 1) [18]. The increase in the number of female physicians, of which many prefer to work part-time [15], and/or decide to leave the workforce at university hospitals and research, will result in a serious lack of experienced medical leaders and researchers over the next 10 years [19].

Figure 1. Percentage of female physicians in age group <35 years and overall (status: 2013 [18]) for those European countries taking party in the proposed study. Cz = Czech Republic; NL = Netherlands. (No data for Lithuania regarding number of females < 35 have been published, but it can extrapolated from the high number of female physicians already that young female physicians will outnumber males as well.).



Advantages of Increasing the Number of Women

Due to the demographic change, the shortage of experienced female medical leaders at hospitals and in research is of great concern, and keeping women at university hospitals is a necessity. A higher number of women in university departments will also broaden the horizon regarding optimal treatment strategies, and thus will maximize treatment success: as shown previously, sex of the physician as well as the sex of the patient is of major importance for the physician-patient interaction, and subsequently for the success of therapies [20-25]. In an American study, it was reported that female physicians were more successful as compared with male colleagues regarding prevention measures (P<.001) [24]. The communication style was also of major importance: females usually present a more patient-centered, empathic style, while men show less emotion and appear to be more goal directed [25]. Remarkably, this played only a minor role in the interaction between male physicians and female patients, in contrast to the interaction between female physicians and female patients. Female patients were less satisfied with their treatment when they felt that their

female physician was not empathic enough; thereby, not fulfilling expectations with regard to gender roles [25]. This had a negative impact on the treatment success. Last, but not least, keeping women in the hospitals will prevent "burn-out" of male and female physicians alike with a more equally distributed work load.

Having more women in the work place, not only for patient care, but also research has high advantages, too. The increased presence of women will advocate different views at the senior level, and will not only remove the male bias in decision-making, but also the "fix the women" perspective (conventional gender expectations). Additionally, it will encourage women to explore different career strategies by having more role models present. As a consequence, the quality and output of research will be strengthened, because there will be a general benefit from mixed teams consisting of men and women [26]. Moreover, it will broaden research's diversity and innovation, because women choose different specializations in medicine, and have a high competence and knowledge level in these fields. Also, potentially biased views in research will be reduced, because female researchers will bring in new perspectives and ideas when initiating projects, analyzing data, and discussing results. Having a mainly male-biased view does not imply that all results are wrong, but bringing a female onto research teams will increase innovation and excellence to medicine.

To summarize, an increase in female participation and leadership will create new chances for an innovative change in hospital working environments and in research settings.

Possible Solutions

Rather than focus on the "women lack confidence" perspective, attention will be directed toward an altered gendered organizational system [10]. Transforming the existing organizational structures in mostly male-dominated and strictly hierarchically structured university hospitals, including the underlying systems of values and appreciation, is not an easy task. This transformation is the single most relevant challenge at present and in the near future socioeconomically, and it is of importance to develop novel concepts now to tackle the challenges ahead.

To attract women to medical research and leadership positions, the working environment needs to be adapted and new models need to be developed, taking into account the transformation of work values (Textbox 1).

Textbox 1. Measures that have shown to have a positive impact on attracting females in medicine [9, 26]

- Mentoring female juniors by experienced female leaders.
- Building and strengthening professional networks of female medical leaders.
- Working toward gender parity in leadership; it was shown, that (1) this provided a more welcoming social environment for women, (2) women participated more in prestigious international collaborative initiatives increasing their scientific output and their reputation, and (3) papers coming out of these mixed groups received more citations from their peers and were considered to be of higher quality [26]. Even though one has to keep in mind that this reflects the US perspective and ecology, it is feasible to assume that these findings can be translated to the life sciences and Europe.
- Foster special competitive grant programs for junior and senior female physician and researchers [9].
- Altering the career pipeline by increasing the number of junior female scientists not only as principal investigators by also as participants and co-authors in prestigious collaborative research projects [26].
- Developing promotion criteria that focus on quality rather than quantity, and have men as ambassadors of change on board [9].

Methods

Overall Concept

All too often, Gender Equality Plans (GEPs) and gender equality offices are seen as aspects being an additional, sometimes even annoying part of everyday (work) life. Innovative strategies, such as implementing a gender dimension in workforce planning, research content, and academic education, are being treated as a 'costly add-on' to the normal work practice. In a model approach to be adopted later by other disciplines of medicine, we will bring together anesthesiology, a clinical specialty with a high share of female physicians and high research activity, and surgery, with a low percentage of female medical doctors, but high research activity as well. We want to overcome simplistic approaches, in which GEPs are developed on a management level and applied on an administration level, only, and forge a Pan-European research platform and consultancy service in order to sustainably promote gender equality in academic medicine.

From our research [27], we know that high-quality management needs to be developed by and within the organization itself and cannot be copied as 'best practice' from outside. Therefore, we want to use a top-down and bottom-up implementation approach in parallel. We will use 'consensus groups' and those concepts, which will lead to a positive outcome, will be shared.

Med *Go* Fem is a complex project addressing multiple perspectives with the participation of multiple stakeholders. On the one hand, this is a challenge, but on the other hand, the changes will clearly outweigh the risks. Our project brings together novel research and innovative implementation methods, which have not yet been applied to GEP or gender action plans in health care. Nonetheless, all research and implementation methods are scientifically sound and have been proven to work in projects of consortium members, such as the German FacharztPlus project [28].

Main Project Goals

The objectives of the study is to develop innovative concepts for transforming GEPs, which will be implemented in 8 university hospitals in 7 countries (Germany, Norway, Lithuania, Poland, the Netherlands, Czech Republic, and Italy) to tailor to the needs of female medical doctors in university hospitals. This should spur a deep transformation of working conditions and research cultures in an innovative and creative way to make institutional structures more attractive and flexible for career plans in academic medicine and biomedical research. We believe that one of the key issues here is that culture is deeply embedded in all layers of the organizational environment (eg, how work, life, and work-life balance is perceived on the national organizational hierarchy).

This proposal has the goal to assess the impact of the following measures in 4 model approaches to instill sustainable change in a best practice scenario.

First, an increase in the number of university hospitals implementing GEPs, the main impact of the Med *Go* Fem project will be the implementation of innovative GEPs tailored to the needs of 7 countries and 8 university hospitals, which will lead to measurable institutional change in gender equality. The partners with a tradition of gender management like Norway and the Netherlands will serve as role models for partners in the start-up phase. This kind of transfer offers the opportunity to identify cultural, country specific, and interdisciplinary factors, and filter out relevant gender-specific factors of interest for the building and implementation of GEPs.

Second, an increase in the number of female researchers: this will improve career perspectives and mobility of female physician/researchers in university hospitals. This will also increase female leadership in the medium run.

Third will be an integration of the gender dimension in research programs. Research performing organizations, like university hospitals, and research funding organizations (RFO) will understand—through the best practices and results of consensus groups—how they can take advantage of GEP, and which type of measures are mandatory to improve the situation of female researchers in academic medicine. RFO can thus use the built knowledge to include mandatory topics for funding applications to enforce the use and implementation of GEP in university hospitals.

Fourth is an optimization in research content to increase the social value of innovations. Gendered innovation [16], as a necessary approach, will be promoted by Med *Go* Fem, and thus will become a part of university hospital strategic concepts as a crosscutting topic to be dealt with in all research and clinical activities. The project will lead to innovative methodologies on implementing GEP in university hospitals. Due to the specific challenges of female physician/researchers in health care, education, and research, this process will need specific adaptations when compared with other research environments. Accordingly, monitoring and assessment will need different indicator parameters.

The project will investigate in more detail:

1. Equal access: the extent to which women have equal access to the resources that contribute to career success, compared with men (eg, career development, protected time for research, role in decision making).
2. Support for work-life balance: the extent to which women are supported in their efforts to balance work and family for the achievement of both personal and professional success (eg, support for temporary reduction of work load, events/meeting schedules consider family demands).
3. Freedom from gender bias: the extent to which women are able to work in an environment in which they are able to voice concerns regarding subtle and overt gender biases (eg, raising issues about the supportiveness of the work environment for women, concerns about biases against women).
4. Chair/chief support: the extent to which the unit leader supports important aspects of women's careers (eg, access to resources and office space, participation in formal and informal meetings, coverage on maternity leave): "Equality will not be achieved without proactive support from key organisations" [9].

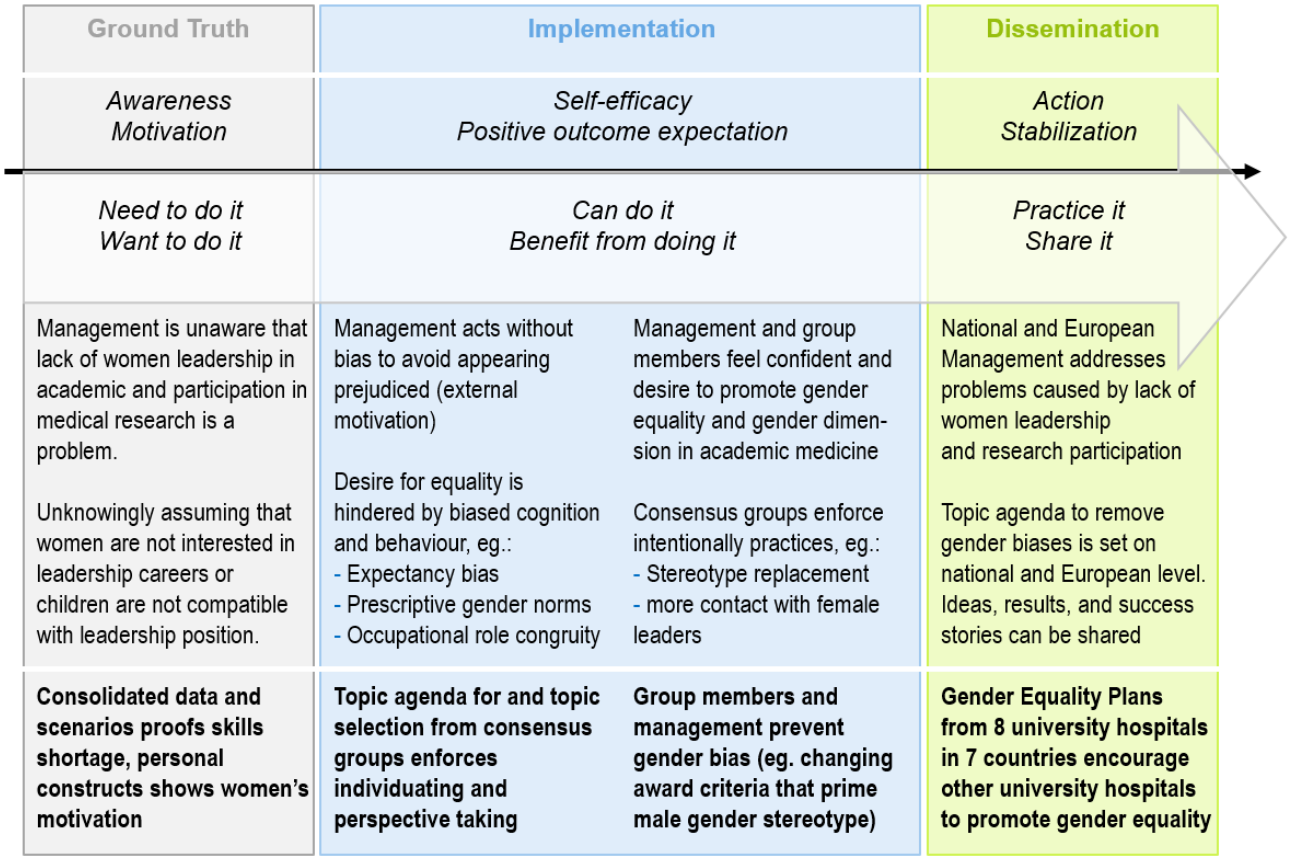
Main Project Stages

We will go through 3 major project stages:

1. We will define 'ground truth' regarding personal as well as organizational constructs by integrating personal construct theory to participatory gender audits (PGA).
2. We will implement and validate effective transformative gender equality by participatory design in "consensus groups" and 'living labs,' which fuses creation and implementation as well as employees' and employers' perspectives.
3. We will disseminate positive effects of gender equality among relevant stakeholders and shareholders by applying female and male 'gender lenses' on all management levels rather than simplistic gender concepts on an administrative level.

All 3 phases of the project have in common that they have a focus on participatory methods (Figure 2). The step-wise implementation will facilitate the move of employees from the motivational state ("I want to do it") to action, supporting both, self-efficacy ("I can do it") and positive outcome expectation ("I benefit from doing it"). The application of consensus groups helps to change behavior, because it generates deliberate practice of the desired behavior [29].

Figure 2. MedGoFem applies a step-wise approach to promote and implement gender equality.

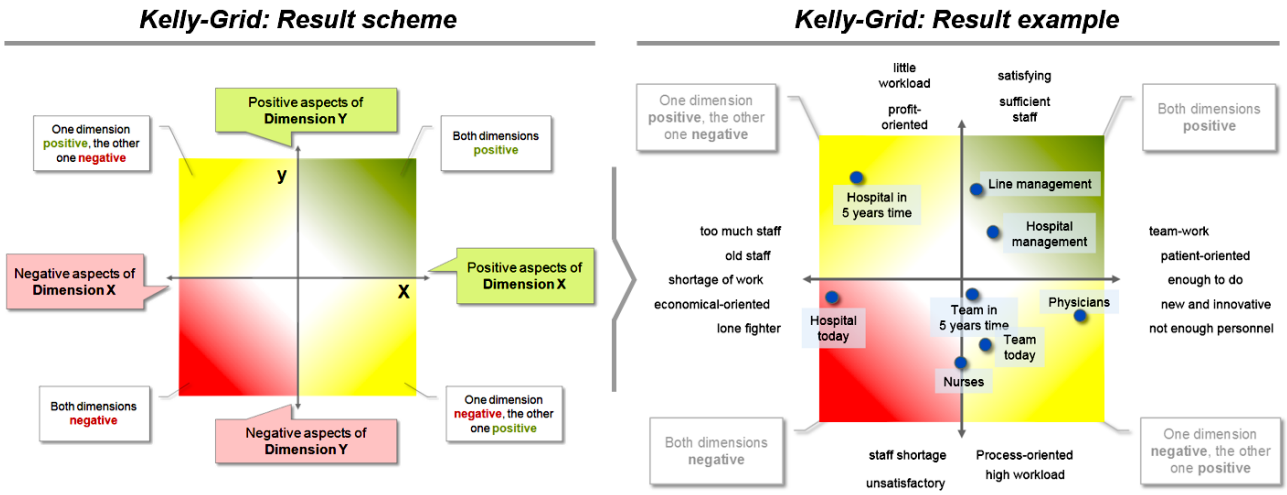


Results

Ground Truth and Personal Constructs

Practice needs to be embedded in the cultural and practical context. For example, values concerning family life and childcare vary across European countries, as does public support and social appreciation of working mothers [30,31]. In the first project step, already existing culture-sensitive measurements developed in academic medicine will be applied, namely Culture Conducive to Women's Academic Success (CCWAS) Survey [32]. In addition, methods that combine qualitative and quantitative aspects as well as personal views and institutional variables by referring to the PGA framework of the International Labor Office(s) [33] of the United Nations. As a start, we will search and review national statistics of gender aspects in academic medicine to identify and fill-in relevant knowledge gaps, taking into account cultural factors. Then, we will use the CCWAS Survey and add data from structured interviews, workshops, and personnel databases recommended in International Labor Office's framework. In a last step, we aim to understand "personal constructs": "gender" is viewed as a personal construct that shows individual differences that are also strongly influenced by people's own internalized cultural values [34,35]. We will use Kelly Grids (or repertory grids) to visualize personal construct spaces (Figure 3). This will lead to national data sets on female physicians working conditions and projected workforce developments on a national basis in 7 European countries. It will also result in PGA of at least university hospitals.

Figure 3. Visualization scheme for Kelly with typical result (taken from FacharztPlus project [35]).



Transformation and Participation

Taking into account the results of the first project step as mapped out before, we will apply, in a second step, the 'nominal group technique' (NGT) to reach and document consensus about necessary organizational transformational measures. Research has shown that NGT is superior to moderated discussions in focus groups, and other more structured techniques, because it regards individual differences, accounts for the strength of conviction, and documents the progress toward achievement of consensus [36]. Moreover, NGT is well documented [37], easy to learn and apply [38], as well as largely applied in health care and nursery [39]. The different steps of the NGT document individual inputs, as a list of ideas, and group decisions in the form of ranking tables. Moreover, comparative studies reveal that nominal group members produced a significantly larger amount of enhancements than respondents in focus groups, and show greater levels of group member satisfaction [36].

Each university hospital will establish 2 to 4 of these consensus groups consisting of 6 to 8 participants, and taking 3 to 5 meetings to come up with full consensus. In consequence, 120 to 240 persons from 8 university hospitals will take part in the consensus groups. The university hospitals' management use approved gender topic lists to elicit interest and initiate consensus groups. Each group will pick 2 to 4 related topics it wants to work on. Once a group reaches consensus, the team members might decide to stay together in order to test the suggested measures (eg, team agreements replacing central holiday planning). The management supports consensus groups, shares ideas, and results among groups, and approves the measures suggested or tested by the groups. A tool package for content- and research-related work in the consensus group will be developed.

Consensus groups may be extended to serve as "living labs" [40], not just defining but also testing their ideas in practice. Using NGT in this process, ideas and decisions can be easily validated, shared, and used as input for further refinement. They also create a sense of self-efficacy ("We can do it") and positive expectation ("We shall benefit from it") in the group [29]. Figure 4 summarizes all elements of our implementation plan.

To assess to what extent gender equality is promoted within the hospital after implementation of measures to improve gender equality, management of the hospital measures the percentage of persons promoting gender equality as compared with the percentage of skeptical people. This proportion is derived from customer loyalty metrics called "net promoter score" [41] (Figure 5).

The "Implementation and Validation" phase will create a wealth of data from participatory and accompanying research, such as interviews, individual mental models (Kelly Grids), change progress (net promoter score), and consolidated data about best practices as case studies and teaching stories.

Figure 4. Information flow in project and central position of consensus groups using "nominal group technique" to create measures to improve gender equality.

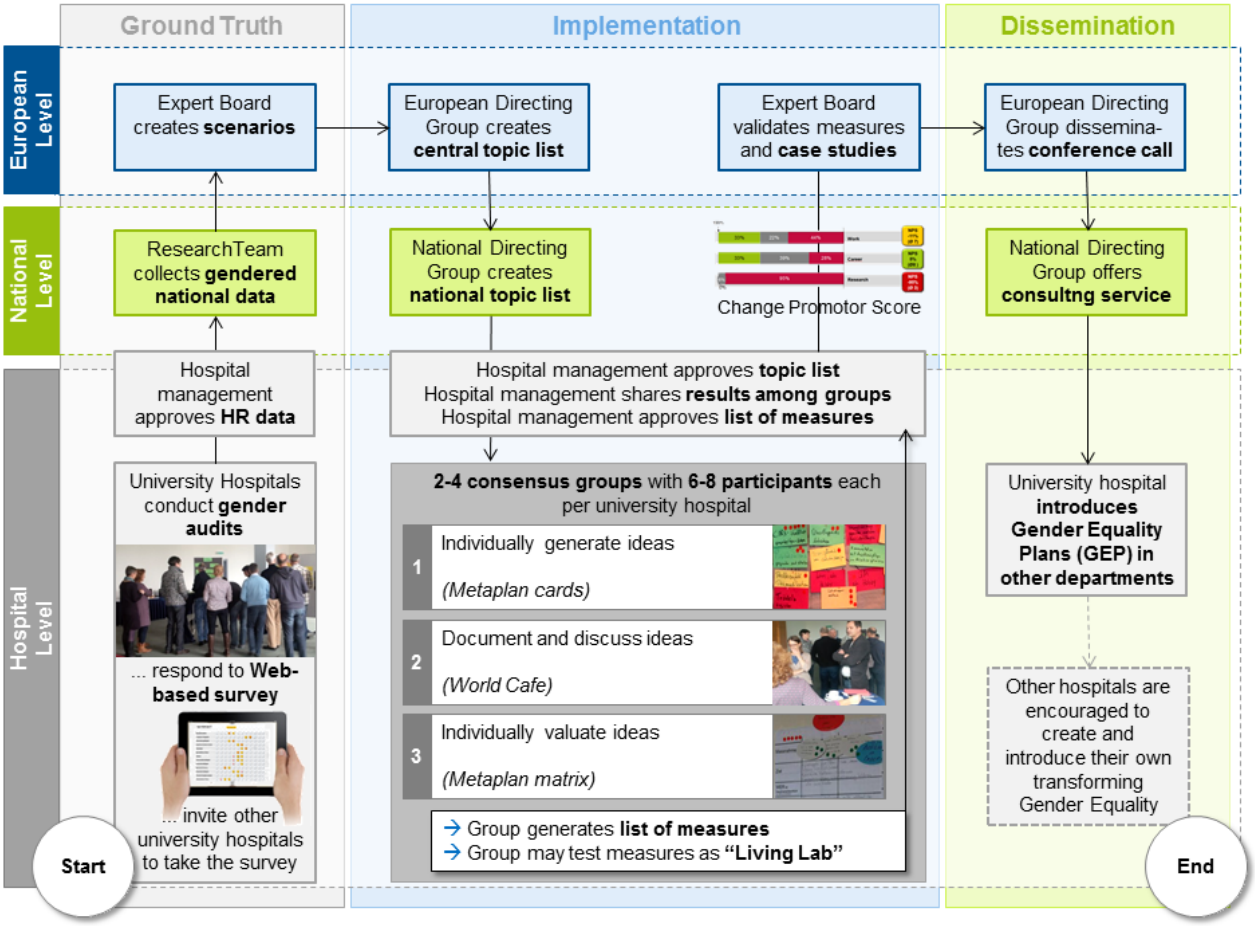
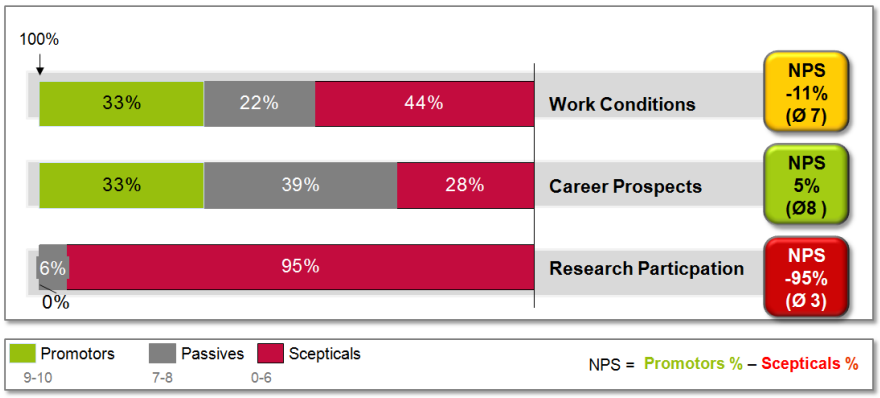


Figure 5. Net Promoter Score (NPS) showing the percentage of physicians in favour or against working, career, and research conditions in a university hospital (taken from FacharztPlus project [35]).



zeb/

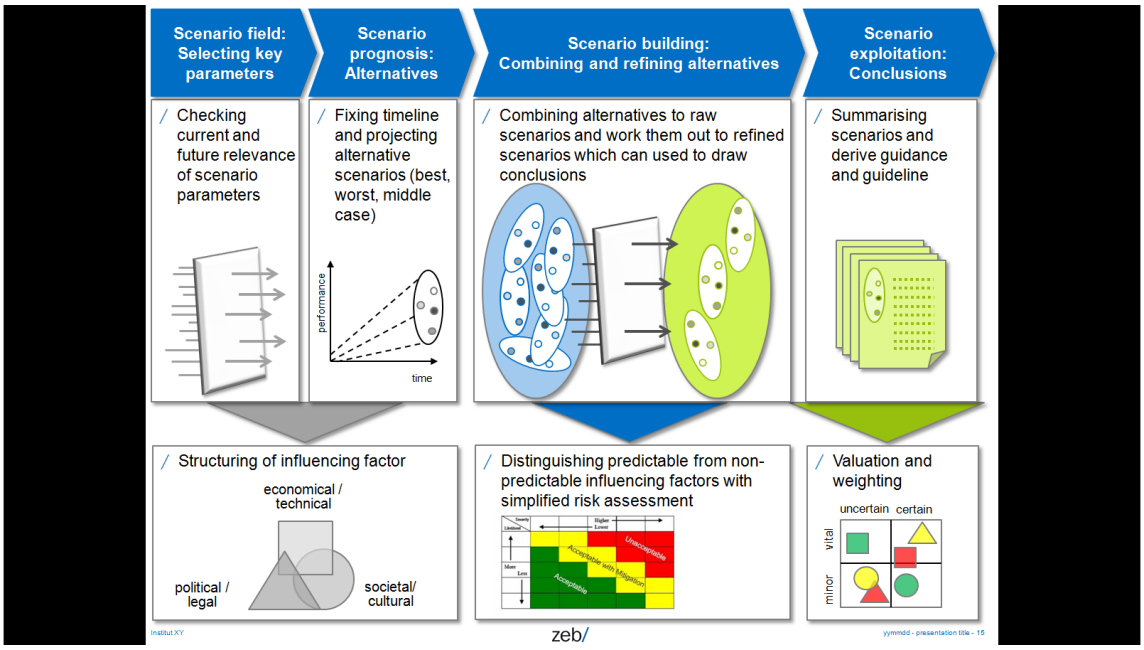
Expert Panels and Dissemination

In a last step, discourse about project results will start on a national level. Cross-national results will be discussed as soon as the national consensus groups and "Directing groups" documented the national outputs and the expert panel. An abbreviated form of scenario analysis will be applied, which has been developed and used by the German Ministry of Education and Research for technical options assessment [42] (Figure 6).

A successful adoption of those measures suggested by the consensus groups in university hospitals is based on the approach developed by the European Foundation for Quality Management [43] and its assessment method 'RADAR' (short for Results, Approaches, Deploy, Assess, and Refine) [44]. The RADAR quality assessments are weighted by their importance so that a weighted score card is available for decision-making within the hospital, and as an input to the expert panel validating the results from all 15 to 20 consensus groups in all 8 university hospitals.

In the last step of the dissemination, a catalogue of services will be developed, presented at the European Conference "Medicine Goes Female," and used to launch a European consulting service for hospitals that wish to implement GEP and gender actions plans after the end of the Med *Go* Fem project. This catalogue will be derived from service gap and service quality models, which—like net promoter score—have also been adopted from marketing research [45,46]. Based on consensus group reports and management assessment relevant service gaps will be identified and service quality levels will be defined. The "dissemination" of the project results will also create output in itself: besides invited lectures, project-related national and international publications, and guidelines and handbooks will be written, and workshops or symposia will be held.

Figure 6. Abbreviated form of scenario analysis for workshops and expert panels.



Discussion

Implementation Strategy and Work Packages

Consistent with the latest research and first-evidence based GEP in academic medicine [32,47] we shall apply a stepwise implementation strategy (Figure 2).

The entire project will be organized in 5 consecutive work packages.

Work Package 1: Understanding Gender Factors

The objectives are to analyze gendered national data and perform participatory gender audits in university hospitals including inquiry of personal constructs.

The tasks performed relate to gender-specific factors for design and success of introduction and implementation of GEP in university hospitals with a focus on their role as a research performing organization will be investigated. In particular, individual mental constructs of physicians will be inquired and cultural, as well as societal aspects taken into account, which may have an influence on the design and success of GEP implementation. Initially, all partner nations will collect comprehensive gender-sensitive demographic data concerning medical staff in research, education and patient care, age and gender distribution, emigration and immigration of physicians, as well as sociodemographic data. These data will enable us to build gender-tailored future scenarios (see objective 2) as well as provide us with up-to-date data to be integrated in medical research concepts. PGA will be analyzed as well.

Work Package 2: Building Future Scenarios

The objectives will be meta-analysis of national data and gender audits as preparation for an expert panel to create 2-fold

scenarios (predicted and requested) for all participating university hospitals

The tasks will include finding an expert panel to transform the results from Work Package 1 into hospital-specific, national, and transnational 2-fold scenarios: one part of the scenario predicts how performance requirements, as well as structure and demands of medical and research staff will change over the next 5 to 10 years (predicted scenario). Another part will describe what the work situation, leadership, and management in university hospitals should look like in order to attract and retain—mostly female—physicians (required scenario). Gaps between predicted and required scenarios form the basis of suggestions for the "living labs," which will be implemented in Work Package 3 in order to develop and test the necessary structural change in university hospitals.

Work Package 3: Consensus Groups and Living Labs

The objectives will be to establish topic-related consensus groups in each university hospital trying to reach consensus about the implementation of relevant aspects of gender equality applying NGT in order to equally regard individual opinions, document ideas and consensus, and report progress to National Directing Groups and share it among consensus groups. Consensus groups working together longer than 6 months may also serve as "living labs," which do not only define but also test selected measures.

The tasks will be to gather national data and hospital-specific PGA (from Work Package 1) as well as predicted and required scenarios (from Work Package 2), and all participating university hospitals set up their own "topic agenda" of relevant gender issues and present these topics to their employees in kick-off workshops. During and after the workshop, employees from medical, as well as administrative, staff are encouraged to join topic-related consensus groups (eg, flexible work schedules, open research platforms). Each group goes through predefined steps: (1) generating, (2) recording, (3) discussing ideas, and (4) voting on them. The National Directing Group collects all generated ideas, validates them, and shares them among consensus groups. Once votes have been made, the results are summarized and validated by the National Directing Group, which also reports to the European Directing Group. If it useful to work out ideas and votes in more detail or test certain measures in practices, the Directing Group may ask the existing consensus groups to stay together as a "living lab" or form a new consensus group.

Work Package 4: Validation and Generalization

The objectives will be to identify culture-, nation-, and discipline-specific success factors for the implementation of GEPs in university hospitals, deriving flexible adaptable models for structural transforming GEP in academic medicine, and the development of 'teaching stories' for academic teaching

The tasks will be to gather national and hospital-specific research data and results. Major culture-, nation-, discipline-specific aspects will be identified by the same expert panel, which has been invited in Work Package 2. The expert panel will work out recommendations and a transnational model for GEPs, which can be adapted to national and disciplinary aspects in all researching hospitals in Europe. Central findings of the "living labs" will be prepared as "teaching stories" for academic teaching in a way that they can become part of the academic curricula in medical faculties and medical management schools on excellence in research, work organization, leadership, and management.

Work Package 5: Dissemination Through Academic Publication and Management Consulting

The objectives will be to develop a foundation of an international European discussion forum in an open research platform, publication of relevant results in at least 4 international academic journals (impact factor > 2), 7 national journals, organization of national workshops and symposia, networking with existing international and national projects, associations, and institutions, and founding a European consulting service for advancement of GEP in hospitals and medical research.

The tasks will focus on the results from research (Work Packages 1, 2, and 4) and implementation (Work Package 3) will be published in relevant international, European, and national journals, which are listed in PubMed, and with an impact factor of 2 or higher [48]. Simultaneously, an academic discussion forum will be initiated on an international renowned research platform. A European conference will be organized in order to invite leading medical experts, research managers, politicians, and partner projects. During this conference, a support services for hospitals, which wish to implement their own GEP, will be introduced and developed into a European consulting service. Participating university hospitals and National Directing Groups will seek to convince other university hospitals in their country to develop and implement GEP.

Expected Impact of MedGoFem

Each work package of the Med *Go* Fem project will create relevant output, generating its own specific impact and close a highly relevant research gap, because most research related to gendered equality that measures working conditions, organizational structures, and research settings has not been conducted in medicine, let alone in academic medicine, but for so-called "MINT" jobs (mathematics, informatics, natural sciences, and technology). In that line, almost all relevant research about the working conditions of women in academic medicine has been conducted in the United States and cannot readily be transferred due to cultural, societal, and economical differences of other countries. Consequently, in the European Union, gender issues in academic medicine are a heavily under-researched issue. Med *Go* Fem will close this gap in all of its project phases.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Evaluation report of the European Commission's reviewing board, which shows that the proposal has been evaluated as "above threshold". Due to budgetary reasons the proposal has been postponed to a "waiting list" and has not yet been funded.

[PDF File (Adobe PDF File), 95KB - resprot_v6i8e152_app1.pdf]

References

1. European Institute for Gender Equality. Gender Equality Index 2012. Brussels: European Union; 2012. URL: http://eige.europa.eu/gender-statistics/gender-equality-index [accessed 2017-07-22] [WebCite Cache ID 6s8y5Szou]
2. European Union. SHE FIGURES 2015, Gender in Research and Innovation: Statistics and Indicators. In: SHE FIGURES. Brussels: Europan Union; 2015.
3. European Union. In: European Union , editor. Publications of the European Union: Report on equality between women and men 2015. Brussels: Publications Office of the European Union; 2016.
4. Expert Group on Structural Change. Structural change in research institutions: Enhancing excellence, gender equality and efficiency in research and innovation. In: European Commission , editor. Report of the Expert Group on Structural Change. Brussels: European Commission; 2016.
5. Centrum voor Gender en Diversiteit. GenderBasic. GenderBasic URL: http://www.genderbasic.nl/ [accessed 2017-04-20] [WebCite Cache ID 6prPlsNbF]
6. De RD. How to add more “family” to the work-life-balance? - family friendliness in medical under- and postgraduate studies and the workplace. GMS Z Med Ausbild 2012;29(2):Doc22 [FREE Full text] [doi: 10.3205/zma000792] [Medline: 22558028]
7. Fridner A, Norell A, Åkesson G, Gustafsson SM, Tevik LL, Schenck-Gustafsson K. Possible reasons why female physicians publish fewer scientific articles than male physicians - a cross-sectional study. BMC Med Educ 2015;15:67 [FREE Full text] [doi: 10.1186/s12909-015-0347-9] [Medline: 25889674]
8. Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. Bibliometrics: global gender disparities in science. Nature 2013;504:211-213. [Medline: 24350369]
9. Boyle PJ, Smith LK, Cooper NJ, Williams KS, O'Connor H. Gender balance: Women are funded more fairly in social science. Nature 2015;525:181-183. [doi: 10.1038/525181a] [Medline: 26354468]
10. O'Neil DA, Hopkins MM. The impact of gendered organizational systems on women's career advancement. Front Psychol 2015;6:905 [FREE Full text] [doi: 10.3389/fpsyg.2015.00905] [Medline: 26175708]
11. Ely RJ, Ibarra H, Kolb DM. Taking gender into account: theory and design for women's leadership development programs. Academy of Management Learning & Education 2011;10:474-493. [doi: 10.5465/amle.2010.0046]
12. Ibarra H, Ely R, Kolb D. Harvard Business Review. Harvard (MA): Harvard Business Manager; 2013. Women Rising: The Unseen Barriers URL: https://hbr.org/2013/09/women-rising-the-unseen-barriers [accessed 2017-07-22] [WebCite Cache ID 6s8yRs5mM]
13. Hoobler J, Lemmon G, Wayne S. Women’s managerial aspirations. J Manag 2011;40:703-730 [FREE Full text] [doi: 10.1177/0149206311426911]
14. Tijdens K, de Vries DH, Steinmetz S. Health workforce remuneration: comparing wage levels, ranking, and dispersion of 16 occupational groups in 20 countries. Hum Resour Health 2013;11:11 [FREE Full text] [doi: 10.1186/1478-4491-11-11] [Medline: 23448429]
15. Reed DA, Enders F, Lindor R, McClees M, Lindor KD. Gender differences in academic productivity and leadership appointments of physicians throughout academic careers. Acad Med 2011;86:43-47. [doi: 10.1097/ACM.0b013e3181ff9ff2] [Medline: 21099390]
16. European Commission. Science with and for society. Brussels: European Commission; 2015. Gender Innovations in Science, Health & Medicine, Engineering and Environment URL: http://ec.europa.eu/research/science-society/gendered-innovations/index_en.cfm [accessed 2017-06-01] [WebCite Cache ID 6prPSo8Rk]
17. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006;3:e442 [FREE Full text] [doi: 10.1371/journal.pmed.0030442] [Medline: 17132052]
18. OECD. Organisation for economic co-operation and development. Stats URL: http://stats.oecd.org/Index.aspx?DataSetCode=HEALTH_REAC [accessed 2017-04-20] [WebCite Cache ID 6prKJPjNv]
19. European Commission. An Agenda for New Skills and Jobs: A European contribution towards full employment? In: European Commission , editor. European Commission Communication, Document 0682. Brussels: European Commission; 2010.
20. Berthold HK, Gouni-Berthold I, Bestehorn KP, Böhm M, Krone W. Physician gender is associated with the quality of type 2 diabetes care. J Intern Med 2008;264:340-350 [FREE Full text] [doi: 10.1111/j.1365-2796.2008.01967.x] [Medline: 18397244]
21. Gouni-Berthold I, Berthold HK, Mantzoros CS, Böhm M, Krone W. Sex disparities in the treatment and control of cardiovascular risk factors in type 2 diabetes. Diabetes Care 2008;31:1389-1391 [FREE Full text] [doi: 10.2337/dc08-0194] [Medline: 18375411]

22. Christen RN, Alder J, Bitzer J. Gender differences in physicians' communicative skills and their influence on patient satisfaction in gynaecological outpatient consultations. Soc Sci Med 2008;66:1474-1483. [doi: 10.1016/j.socscimed.2007.12.011] [Medline: 18222586]
23. Nicolai J, Demmel R. The impact of gender stereotypes on the evaluation of general practitioners' communication skills: an experimental study using transcripts of physician-patient encounters. Patient Educ Couns 2007;69:200-205. [doi: 10.1016/j.pec.2007.08.013] [Medline: 17942269]
24. Reid RO, Friedberg MW, Adams JL, McGlynn EA, Mehrotra A. Associations between physician characteristics and quality of care. Arch Intern Med 2010;170:1442-1449 [FREE Full text] [doi: 10.1001/archinternmed.2010.307] [Medline: 20837830]
25. Schmid MM, Hall JA, Roter DL. Disentangling physician sex and physician communication style: their effects on patient satisfaction in a virtual medical visit. Patient Educ Couns 2007;68:16-22. [doi: 10.1016/j.pec.2007.03.020] [Medline: 17482418]
26. Campbell LG, Mehtani S, Dozier ME, Rinehart J. Gender-heterogeneous working groups produce higher quality science. PLoS One 2013;8:e79147 [FREE Full text] [doi: 10.1371/journal.pone.0079147] [Medline: 24205372]
27. Hasebrook J, Singer M. Exzellenz zahlt sich aus (excellence pays off). Harvard Business Manager 2015;2015:2-3 [FREE Full text]
28. Hasebrook JP, Hinkelmann J, Volkert T, Rodde S, Hahnenkamp K. Securing the continuity of medical competence in times of demographic change: a proposal. JMIR Res Protoc 2016;5:e240 [FREE Full text] [doi: 10.2196/resprot.5897] [Medline: 28003176]
29. Carnes M, Devine P, Baier ML, Byars-Winston A, Fine E, Ford C, et al. The effect of an intervention to break the gender bias habit for faculty at one institution: a cluster randomized, controlled trial. Acad Med 2015;90:221-230 [FREE Full text] [doi: 10.1097/ACM.0000000000000552] [Medline: 25374039]
30. Inglehart R, Norris P. Rising Tide: Gender Equality and Cultural Change Around the World. New York: Cambridge University Press; 2003.
31. World Value Survey. URL: http://www.worldvaluessurvey.org/wvs.jsp [accessed 2017-04-20] [WebCite Cache ID 6prSXo8Al]
32. Westring A, Speck R, Sammel M, Scott P, Tuton L, Grisso J, et al. A culture conducive to women's academic success: development of a measure. Acad Med 2012;87:1622-1631 [FREE Full text] [doi: 10.1097/ACM.0b013e31826dbfd1] [Medline: 23018337]
33. International Labor Office (ILO). The ILO participatory gender audit methodology, 2nd Edition; 2012. New York: International Labor Office (ILO); 2012. A manual for gender audit facilitators: The ILO participatory gender audit methodology URL: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---gender/documents/publication/wcms_187411.pdf [accessed 2017-07-22] [WebCite Cache ID 6s8zVByD3]
34. Bourne D, Özbilgin MF. Strategies for combating gendered perceptions of careers. Career Dev Int 2008;13:320-332. [doi: 10.1108/13620430810880817]
35. Barker L. Effective leadership within hospice and specialist palliative care units. J Manag Med 2000;14:291-309. [Medline: 11200297]
36. de Ruyter K. Focus versus nominal group interviews: a comparative analysis. Mrkting Intelligence & Plan 1996;14:44-50. [doi: 10.1108/02634509610131153]
37. Centers for Disease Control and Prevention (CDC). CDC Evaluation Briefs No 7. Atlanta (GA): Centers for Disease Control and Prevention (CDC); 2006. Gaining Consensus Among Stakeholders Through the Nominal Group Technique URL: https://www.cdc.gov/HealthyYouth/evaluation/pdf/brief7.pdf [accessed 2017-07-22] [WebCite Cache ID 6s8zp9Ype]
38. Sim J. Collecting and analysing qualitative data: issues raised by the focus group. J Adv Nurs 1998;28:345-352. [Medline: 9725732]
39. O'Donovan A, Mohile SG, Leech M. Expert consensus panel guidelines on geriatric assessment in oncology. Eur J Cancer Care (Engl) 2015;24:574-589 [FREE Full text] [doi: 10.1111/ecc.12302] [Medline: 25757457]
40. Schumacher J, Niitamo V. European Living Labs - A New Approach for Human Centric Regional Innovation. Berlin: Wissenschaftlicher Verlag; 2008.
41. Reichheld FF. The one number you need to grow. Harv Bus Rev 2003;81:46-54, 124. [Medline: 14712543]
42. German Ministry of Education and Research (BMBF). Zukunftsbild Industrie 4. Grolman results. Berlin: German Ministry of Education and Research (BMBF); 2014. URL: https://www.bmbf.de/pub/Zukunftsbild_Industrie_4.0.pdf [accessed 2017-07-22] [WebCite Cache ID 6s903OmB4]
43. European Foundation for Quality Management (EFQM). EFQM - Leading Excellence. Brussels: European Foundation for Quality Management (EFQM); 2017. URL: http://www.efqm.org/ [accessed 2017-04-20] [WebCite Cache ID 6prVAn8Re]
44. European Foundation for Quality Management (EFQM). EFQM Leading Excellence. Brussels: European Foundation for Quality Management (EFQM); 2017. Radar Logic URL: http://www.efqm.org/efqm-model/radar-logic [accessed 2017-04-20] [WebCite Cache ID 6prVJRMnc]
45. Parasuraman A, Zeithaml VA, Berry LL. A conceptual model of service quality and its implications for future research. Journal of Marketing 1985;49:23-41. [doi: 10.2307/1251430]

46. Parasuraman A, Zeithaml VA, Berry LL. Reassessment of expectations as a comparison standard in measuring service quality: implications for further research. Journal of Marketing 1994;58:111. [doi: 10.2307/1252255]
47. Westring AF, Speck RM, Dupuis SM, Scott P, Conant EF, Tuton LW, et al. Culture matters: the pivotal role of culture for women's careers in academic medicine. Acad Med 2014;89:658-663 [FREE Full text] [doi: 10.1097/ACM.0000000000000173] [Medline: 24556773]
48. Kianifar H, Sadeghi R, Zarifmahmoudi L. Comparison between impact factor, eigenfactor metrics, and scimago journal rank indicator of pediatric neurology journals. Acta Inform Med 2014;22:103-106 [FREE Full text] [doi: 10.5455/aim.2014.22.103-106] [Medline: 24825934]

Abbreviations

CCWAS: Culture Conducive to Women's Academic Success
GEP: gender equality plans
NGT: nominal group technique
PGA: participatory gender audits
RADAR: results, approaches, deploy, assess, and refine
RFO: research funding organizations

Edited by G Eysenbach; submitted 02.03.17; peer-reviewed by P Ovseiko, A Punnett; comments to author 30.03.17; revised version received 04.06.17; accepted 27.06.17; published 02.08.17.

Please cite as:
Hasebrook J, Hahnenkamp K, Buhre WF, de Korte-de Boer D, Hamaekers AE, Metelmann B, Metelmann C, Bortul M, Palmisano S, Mellin-Olsen J, Macas A, Andres J, Prokop-Dorner A, Vymazal T, Hinkelmann J, Rodde S, Pfleiderer B
Medicine Goes Female: Protocol for Improving Career Options of Females and Working Conditions for Researching Physicians in Clinical Medical Research by Organizational Transformation and Participatory Design
JMIR Res Protoc 2017;6(8):e152
URL: http://www.researchprotocols.org/2017/8/e152/
doi:10.2196/resprot.7632
PMID:28768613

Protocol

Combination Analgesic Development for Enhanced Clinical Efficacy (CADENCE Trial): Study Protocol for a Double-Blind, Randomized, Placebo-Controlled Crossover Trial of an Alpha-Lipoic Acid – Pregabalin Combination for the Treatment of Fibromyalgia Pain

Ian Gilron[1], MSc, MD; Dongsheng Tu[2], PhD; Ronald Holden[2], PhD; Tanveer Towheed[2], MD; Elizabeth Vandenkerkhof[2], DrPH; Roumen Milev[2], MD, PhD

[1]Queen's University, Department of Anesthesiology and Perioperative Medicine, Queen's University, Kingston, ON, Canada

[2]Queen's University, Kingston, ON, Canada

Corresponding Author:
Ian Gilron, MSc, MD
Queen's University
Department of Anesthesiology and Perioperative Medicine
Queen's University
Kingston General Hospital
76 Stuart Street, Victory 2 Pavillion
Kingston, ON,
Canada
Phone: 1 613 549 6666 ext 3963
Fax: 1 613 548 1375
Email: gilroni@queensu.ca

Abstract

Background: Fibromyalgia is a clinical disorder commonly presenting with chronic widespread pain as well as sleep disturbance, fatigue, depression, and cognitive dysfunction. There is an urgent need for treatment strategies that provide better pain relief and fewer adverse effects (AEs). Efforts to develop rational combinations of specific fibromyalgia treatments have demonstrated potential for measurable improvements in pain relief, quality of life, and health care utilization. More than half of fibromyalgia patients receive 2 or more analgesics but current combination use is based on limited evidence. As an early proof-of-concept project from the Canadian Institutes of Health Research–Strategy on Patient-Oriented Research Chronic Pain Network, this trial protocol is expected to advance the field by rigorously evaluating a new treatment combination for fibromyalgia.

Objective: We will test the hypothesis that analgesic combinations containing at least one nonsedating agent would be as safe but more effective than either monotherapy because of additive pain relief without increasing overall AEs. Pregabalin (PGB), a sedating anticonvulsant, is proven effective for fibromyalgia, and the antioxidant, alpha-lipoic acid (ALA), one of the only nonsedating systemic agents proven effective for neuropathic pain, is currently being evaluated in fibromyalgia. Thus, we will conduct a clinical trial to compare a PGB+ALA combination to each monotherapy for fibromyalgia.

Methods: Using a double-blind, double-dummy, crossover design, 54 adults with fibromyalgia will be randomly allocated to 1 of 6 sequences of treatment with PGB, ALA, and PGB+ALA combination. During each of 3 different treatment periods, participants will take 2 sets of capsules containing (1) ALA (or placebo) and (2) PGB (or placebo) for 31 days, followed by an 11-day taper/washout period. The primary outcome will be mean daily pain intensity (0 to 10 scale) at maximal tolerated doses (MTDs) during each period. Secondary outcomes, assessed at MTD, will include global improvement, adverse events, mood, and quality of life.

Results: This trial attained ethics approval March 6, 2017 (Queen's University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board protocol number ANAE-313-17), and recruitment is set to start in August 2017.

Conclusions: This trial will provide rigorous evidence comparing the efficacy of a PGB-ALA combination to PGB alone and ALA alone in the treatment of fibromyalgia.

Trial Registration: International Standard Randomized Controlled Trial Number ISRCTN14939460; https://www.isrctn.com/ISRCTN1493946 (Archived by WebCite at http://www.webcitation.org/6sFqAjxkt)

(JMIR Res Protoc 2017;6(8):e154) doi:10.2196/resprot.8001

KEYWORDS

fibromyalgia; alpha-lipoic acid; antioxidant; pregabalin; anticonvulsant

Introduction

Chronic pain affects 20% to 25% of the population [1] and is one of the most common reasons to see a health care provider and to miss work [2]. In North America alone, chronic pain costs over $650 billion in health care and lost productivity [3]. Fibromyalgia is a complex clinical disorder characterized by chronic widespread pain that is also associated with sleep disturbance, fatigue, irritable bowel syndrome, depressed mood, and, possibly cognitive dysfunction [4,5]. Patients suffering with fibromyalgia very frequently report functional disability and impaired quality of life [6]; furthermore, fibromyalgia is a common disorder estimated to affect 1.6% of men and 4.9% of women [7].

Hundreds of randomized controlled trials (RCTs) have evaluated various drug (eg, nonsteroidal anti-inflammatory drugs, antidepressants, opioids, and anticonvulsants) and nondrug (eg, exercise, acupuncture, cognitive behavioral therapy) therapies for fibromyalgia [6]. In addition to exercise and cognitive behavioral therapy, pharmacotherapy remains an important treatment for fibromyalgia. Evidence-based treatment recommendations from various groups including the European League Against Rheumatism and Canadian Pain Society have included amitriptyline, cyclobenzaprine, tramadol, gabapentin/pregabalin, fluoxetine, and duloxetine [8-10].

Available drugs used for fibromyalgia reduce pain on average by only 25% to 40%, and meaningful relief occurs in only 40% to 60% of patients, in part due to incomplete efficacy as well as commonly encountered dose-limiting adverse effects (AEs) (eg, sedation, cognitive dysfunction, and dizziness). Combining two drugs with different pharmacological mechanisms has the potential to provide superior relief over monotherapy without increasing side effects [11]. A recent trial has demonstrated greater analgesic efficacy with a pregabalin-duloxetine combination versus either monotherapy without an increase in side effect profile [12]. Although this was a positive finding, the additive benefit was submaximal because these two agents cause some similar AEs, and doses must be reduced during combination therapy to maintain safety and tolerability.

Thus, we hypothesize that analgesic combinations containing at least one nonsedating agent would provide even greater additive benefits because of additive pain relief but nonadditive AEs. Both pregabalin (PGB) and alpha-lipoic acid (ALA) are approved by Health Canada and proven for the treatment of neuropathic pain [13,14]. Based on a rationale for the use of ALA to treat fibromyalgia pain [15], a placebo-controlled RCT of ALA monotherapy is currently under way in this population [ISRCTN58259979].

An important pharmacological mechanism of PGB is the blockade of α-2-δ subunits of N-type voltage gated calcium channels, resulting in decreased calcium influx and neurotransmitter release [16,17]. ALA has been studied in both preclinical and clinical neuropathic pain conditions. In a rat model of streptozocin-induced diabetes, ALA delayed the onset of polyneuropathy [18]. Mechanistic studies suggest decreased nociceptive sensitivity by inhibition of T-type calcium (Cav3.2) channels [19], distinct from that of PGB which inhibits N-type calcium channels [17], suggesting potential for synergy at these different sites of action.

At least 16 trials of over 1320 patients have reported reductions in pain and other symptoms [13,20] and a recent meta-analysis reported a number needed to treat of 6.3 [13]. Also, 1 trial reported improvement in neuropathic pain symptoms after 4 years of treatment [21]. AEs of nausea, vomiting, headache, and vertigo have been reported in studies involving >1200 mg per day of ALA. There have also been rare reports of hypoglycemia (low blood sugar) in diabetic patients taking ALA and reporting symptoms of sweating, paleness, chills, headache, dizziness, and/or confusion. We identified only 1 study of a combination similar to ALA+PGB—ALA plus gabapentin (related to PGB) in the treatment of burning mouth syndrome [22]. Despite the study having major methodological flaws, greater benefit with this combination was suggested versus monotherapy, and AEs were reported overall as very mild [22].

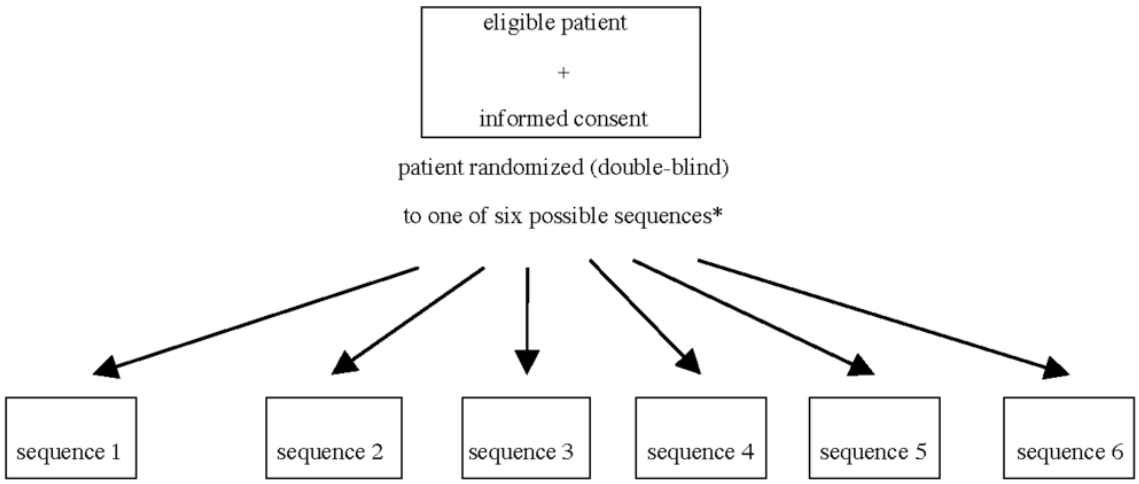
Thus, our goal is to conduct a novel double-blind RCT to compare the combination of the anticonvulsant PGB with the nonsedating antioxidant ALA to each monotherapy for the treatment of pain in fibromyalgia.

Methods

Ethics

This study underwent ethics review and received a compliance notice by the Queen's University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board on March 6, 2017. This trial will be conducted at one site, Providence Care Hospital, Kingston, Ontario, Canada. This trial is registered with the International Standard Randomized Controlled Trial Number Registry [ISRCTN14939460].

Figure 1. Trial design.



This will each be a 3-period, active treatment-controlled randomized double-blind trial, using a double-dummy, balanced Latin Square crossover design in which patients will be allocated to one of 6 treatment sequences of the three treatments: ALA, PGB and ALA-PGB combination

*see below for specific treatment sequences

TREATMENT SEQUENCES

(each patient is randomized to ONE of these sequences)

[All patients complete all three treatment periods (i.e. A, B and C) as per the treatment sequence they were randomized to (i.e. 1, 2, 3, 4, 5 or 6)]

Baseline 7 day washout of prohibited medications (e.g. gabapentin, pregabalin, alpha-lipoic acid)	Sequence	A 8 weeks (24 day titration; 7 day fixed; 7 day taper; 4 day washout)	B 8 weeks (24 day titration; 7 day fixed; 7 day taper; 4 day washout)	C 8 weeks (24 day titration; 7 day fixed; 7 day taper; 4 day washout)
	1	COMBINATION (A + P)	Alpha-lipoic acid (A+ Pp)	PREGABALIN (Pa + P)
	2	PREGABALIN (Pa + P)	COMBINATION (A + P)	Alpha-lipoic acid (A + Pp)
	3	Alpha-lipoic acid (A + Pp)	PREGABALIN (Pa + P)	COMBINATION (A + P)
	4	COMBINATION (A + P)	PREGABALIN (Pa + P)	Alpha-lipoic acid (A+ Pp)
	5	PREGABALIN (Pa + P)	Alpha-lipoic acid (A + Pp)	COMBINATION (A + P)
	6	Alpha-lipoic acid (A + Pp)	COMBINATION (A + P)	PREGABALIN (Pa + P)

Legend: Symbols in parentheses indicate the content of the corresponding blinded study drug capsules as per the double-dummy design. A=alpha-lipoic acid; Pa= "alpha-lipoic acid" placebo; P=pregabalin; Pp= "pregabalin" placebo

Aims and Hypothesis

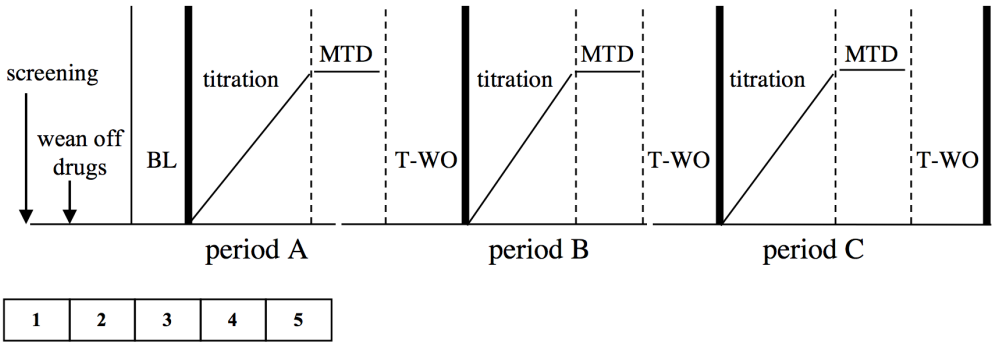
The objective of this trial is to compare the safety and efficacy of a PGB+ALA combination to each monotherapy in treating participants with fibromyalgia. Our primary hypothesis is that PGB+ALA has greater analgesic efficacy versus either monotherapy.

Trial Design

We have designed a single-center, double-blind, double-dummy, randomized, controlled 3-period crossover trial comparing a PGB+ALA combination to monotherapy in treating fibromyalgia (Figures 1 and 2). This trial is compliant with Health Canada/International Conference on Harmonization guidelines and incorporates outcome measures recommended by the

Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials. Using a flexible dose titration, Latin Square crossover design, treatments will be titrated during each of 3 treatment periods to maximal tolerated dose (MTD) with primary and secondary trial analyses comparing the 3 treatments using end-of-period outcomes. Internal validity of our crossover design is supported by stability of fibromyalgia over time [23-26] and the risk of carryover from 1 period to the next is very low because each period is followed by an 11-day dose taper and drug washout, and the final MTD week for each period (from which the primary outcome is obtained) is separated from the next period's final week by 7 weeks (ie, ≥20 half-lives of the drugs studied). Nevertheless, exploratory analyses will be conducted to identify if any low-order carryover effect does exist.

Figure 2. Trial design, continued.



1. Screening

2. Wean off current alpha-lipoic acid, gabapentin and pregabalin for ≥1 week prior to commencing

   - patients taking, and perceiving benefit from opioids (<90/day morphine equivalents), antidepressants, nonsteroidal anti-inflammatory agents or acetaminophen may continue these at a steady dose during the study.

   - patients required to avoid any procedural pain therapies (e.g. neurosurgical interventions, nerve blocks or acupuncture) during the entire study as these treatments may be unevenly distributed across treatment periods and could skew the study's results.

3. BL – 7 day baseline period

4. Double-blind randomization to one of six possible treatment sequences (e.g. sequence 1: combination>alpha-lipoic acid>pregabalin) such that each patient progresses through each of three 6 week treatment periods (i.e. periods "A", "B" and "C").

5. Each treatment period starts with a 24 day study drug dose titration towards maximal tolerated dose (MTD). If MTD is reached before the end of this 24 day period, that dose will be continued up to and including days 25-31 of the treatment period and then concluded with a 7 day taper/ 4 day washout (T-WO).

   A study nurse will contact patients by telephone at least twice a week to evaluate adverse effects, guide study drug titration and encourage compliance. Study patients will be encouraged to contact a study physician, as needed, who will be available 24 hours a day by pager in order to deal with study-related problems that may occur between scheduled study nurse telephone calls.

   With each dosage increase of study medication in the titration schedule, if mild to moderate treatment-emergent adverse effects are encountered, patients will be asked if they can tolerate continuing at that dose for another 2-3 days. If so, this dosage will be continued with the expectation that tolerance to side effects will occur. If side effects are intolerable or do not improve, both study medications will be decreased to the next lowest possible dose and an increase will be attempted one more time at the next scheduled dose increase. If this again results in intolerable side effects, both study drugs will be decrease back to the previous dose, which will be defined as the maximally tolerated dose (MTD) for that individual.

Participant Allocation

As per the 3-period Latin Square crossover, patients are randomly allocated to 1 of 6 sequences of ALA, PGB, and combination (Figures 1 and 2). Before the trial, an independent pharmacist and biostatistician will prepare a concealed allocation schedule using a computer-generated block randomization process to randomly assign treatment sequences to a consecutive

series of numbers within a block. Each patient will be assigned to the next consecutive number, and the corresponding sequence of medications will be dispensed. All study personnel will be blinded to the block sizes to preserve allocation concealment.

Protecting Against Bias

Medications will be encapsulated (ALA: blue, PGB: gray) in an identical fashion across all periods. As per a double-blind, double-dummy design, patients will take both sets of medications so treatment conditions will be identical across all 3 treatment periods. Treatment codes will be generated by the investigational pharmacist and concealed until trial completion. In case of emergency, individual codes will be disclosed by an investigational pharmacist to a nonstudy clinician. A questionnaire completed by every participant at the end of each period will ask patients to guess the treatment they received to assess blinding.

Participants

Men and women aged 18 years and older meeting the 2016 updated American College of Rheumatology diagnostic criteria for fibromyalgia [27] will be considered for recruitment following informed consent. See Textbox 1 for selection criteria.

Textbox 1. Selection criteria.

Inclusion criteria:

- Adults aged 18 years and older
- Diagnosed with fibromyalgia using the 2016 updated American College of Rheumatology diagnostic criteria [27]
- Experience daily moderate pain (≥4/10 on a numerical rating scale) for at least 3 months
- Women of childbearing potential must have a negative serum beta–human chorionic gonadotropin test and are required to use a highly effective form of contraception while on trial
- Have the necessary abilities, visual acuity, and English language skills to complete questionnaires and pain diaries and to participate in telephone communication with study nurses to permit titration of the study drugs

Exclusion criteria:

- Presence of a painful condition, including inflammatory rheumatic disease, more than 50% as severe as but distinct from fibromyalgia
- Women who are pregnant or lactating
- Women of childbearing potential not using adequate contraceptives
- End-stage kidney or liver disease
- Unstable cardiovascular disease (myocardial infarction within the preceding year, unstable angina, or congestive heart failure) or clinically relevant abnormal 12-lead electrocardiogram
- Any poorly controlled medical condition that, in the opinion of the investigator, would interfere with proper conduct of the trial
- Severe depression, as determined by a Beck Depression Inventory–II score of 29 or more; suicidal ideation, as determined by a Beck Depression Inventory–II item 9 score of 2 or more; or any current major psychiatric disorder (eg, schizophrenia, bipolar disorder) that is not well controlled
- Hypersensitivity to any of the study medications
- Any current alcohol or drug abuse or dependence (except nicotine and caffeine). Participants with a history of abuse or dependence with more than 1 year of abstinence may be considered for inclusion
- Those taking more than 90 mg morphine equivalents per day

Study Interventions

During each of 3 trial periods, using a double-blind randomized crossover design, patients will receive 2 sets of capsules (Figure 3): (1) blue capsules (ALA 300 mg or placebo) and (2) gray capsules (PGB 75 mg or placebo). During the combination period, blue will contain ALA and gray will contain PGB. During the ALA alone period, blue will contain ALA and gray will contain placebo. During the PGB alone period, blue will contain placebo and gray will contain PGB.

Consenting patients on ALA or PGB (or gabapentin) pretrial will agree to be weaned gradually for a washout of at least 7 days. Co-interventions: patients taking and perceiving benefit from opioids (<90 mg morphine equivalents), antidepressants (tricyclic, selective serotonin reuptake inhibitors or serotonin-norepinephrine reuptake inhibitors), nonsteroidal anti-inflammatory drugs, or acetaminophen may continue these at a steady dose for the entire study. Any cognitive behavioral therapy or exercise programs may continue only if they can be scheduled evenly across all treatment periods. Research staff will monitor and advise patients weekly about prohibited co-interventions throughout the study. A thorough understanding of the threats to validity of using forbidden co-interventions (gabapentin, PGB, ALA, and any other newly initiated analgesic intervention) is heavily emphasized to participants. Patients will not be allowed to start new cognitive behavioral therapy or exercise programs after study initiation and must avoid any procedural therapies (eg, nerve blocks or acupuncture) during the entire study. Any pain exacerbations that in the opinion of the patient warrant initiation of a new therapy would necessitate trial discontinuation and immediate weaning from study

medications, but these patients would still be included in the trial analyses.

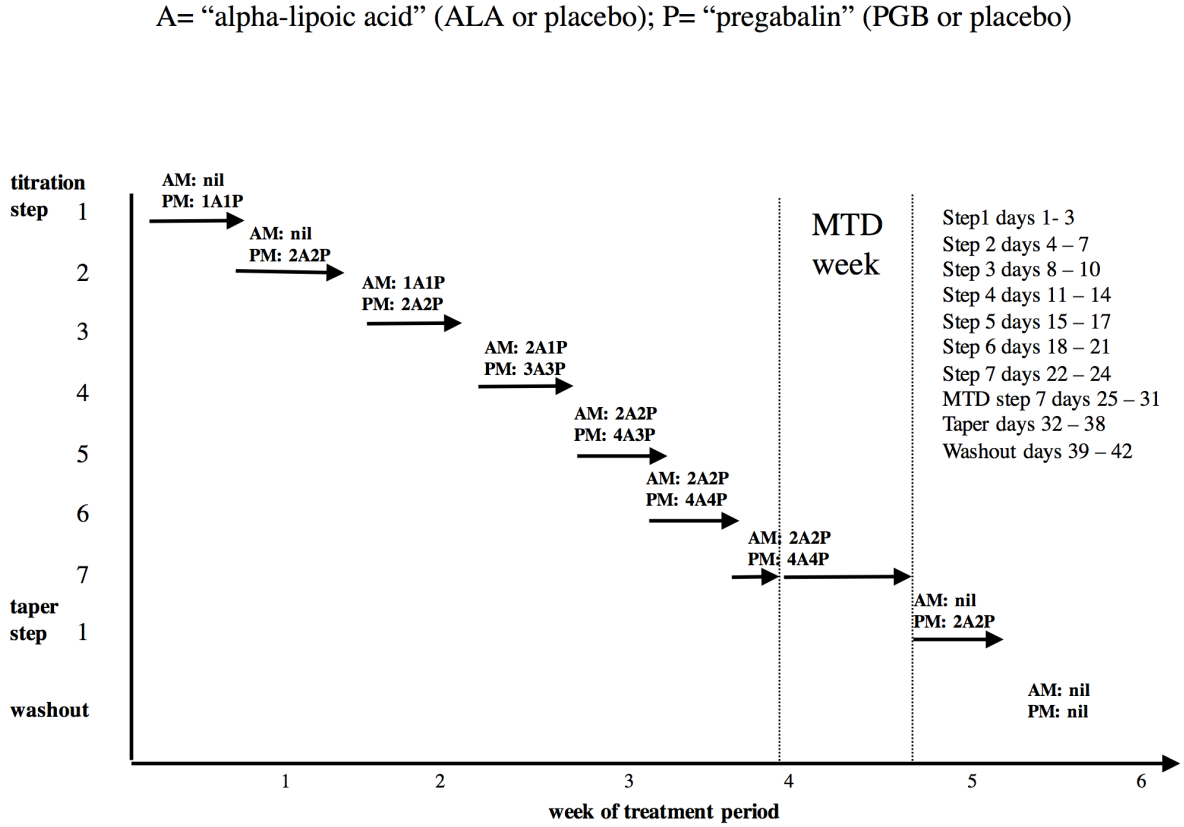
Dose Titration

Study medication will follow a flexible dose titration to MTD to balance tolerability and relief, with regularly weekly calls by research personnel. This means that doses of study medication will not be further increased if intolerable AEs are encountered at higher doses or if "a lot" or "complete" pain relief is achieved. The MTD fixed dose week will be from days 25 to 31. However, if MTD is reached before day 25, that MTD dose will continue up to and including the day 25 to 31 period. The MTD fixed dose week will be followed by a 7-day dose taper and 4-day complete washout. Daily pain ratings will be completed throughout the trial. During dose taper and washout periods only, patients may take acetaminophen ≤8 tablets per day (325 mg per tablet) as needed. This rescue medication is very unlikely to affect the primary outcome measure of pain intensity during the MTD phase of each treatment period.

Trial Duration and Follow-Up Frequency

Each of the 3 treatment periods will be 6 weeks, for a total trial duration of 18 weeks. The nurse will phone patients weekly to evaluate AEs, guide drug titration, and encourage compliance. Patients will be seen in clinic at the end of each treatment period for assessment of vital signs and measurement of secondary outcomes (Figure 4). Patients will be followed up by phone 2 weeks and 3 months after trial completion (including patients who were withdrawn from the trial prematurely) to document any subsequent AEs.

Figure 3. Study drug schedules.



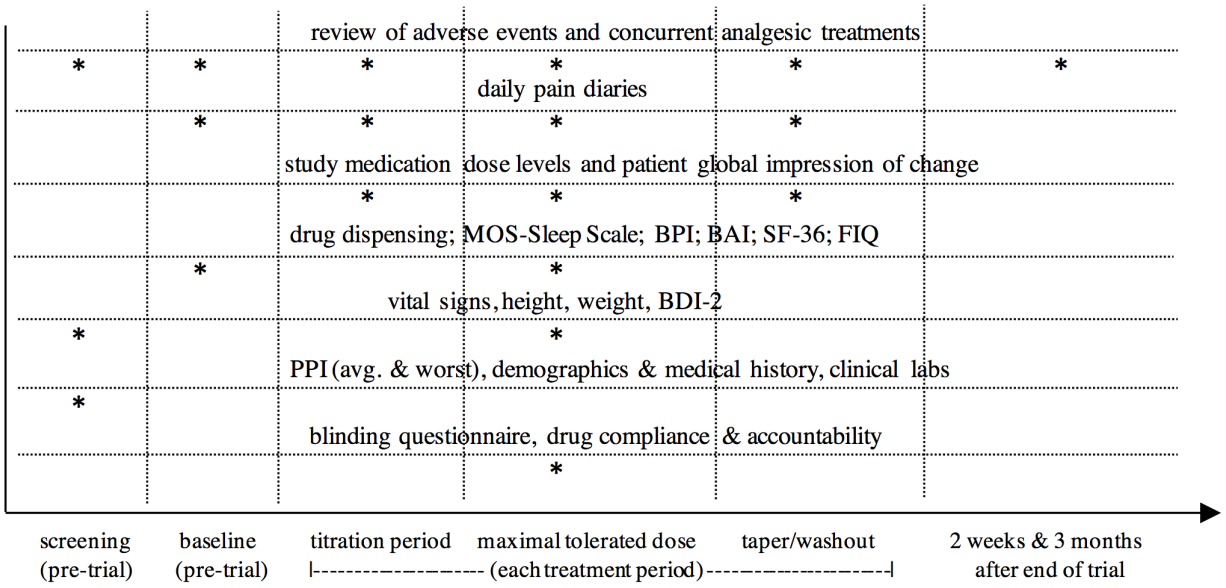
For each trial period, patients receive blue "A" (ALA 300 or placebo), and grey "P" (PGB 75mg or placebo), capsules. Each step above indicates the number of each capsules taken before breakfast & in the evening. The schedule above indicates uppermost dose ceilings at each timepoint. Titration towards individual maximally tolerated dose (MTD) will be guided by safety determined by weekly AE monitoring. Thus, titration may be slower and MTD may be lower than shown above.

Outcome Measures and Safety Assessment

The primary outcome is mean daily pain (0 to 10 numerical rating scale with 0=no pain and 10=worst pain imaginable), rated 3 times daily (8 AM, 4 PM, and 8 PM) and averaged over the MTD fixed dose week (days 39 to 45) of each period. Secondary outcomes include daily pain at other timepoints, the MTDs of PGB and ALA, frequency and severity of AEs and patient global impression of change [28], short form McGill Pain Questionnaire [29], Fibromyalgia Impact Questionnaire [30], Brief Pain Inventory [31], Beck Depression Inventory-II [32], Beck Anxiety Inventory [33], Short Form-36 health survey

[34], blinding questionnaires, and acetaminophen consumption. Timing of outcome assessments is described in Figure 4. Patient safety will be ensured by vigilant AE assessment and judicious drug titration. Any occurrences of major AEs will be tracked as secondary outcomes and also reported to the Queen's Ethics Board, Health Canada. Assessment and reporting of AEs will adhere to Consolidated Standards for Reporting Trials recommendations [35].

Figure 4. Schedule of assessments.



PPI – present pain intensity (0-10 numerical scale)
MOS – Medical Outcomes Study
SF-MPQ – short form McGill Pain Questionnaire;
BPI – Brief Pain Inventory;
BDI-2 – Beck Depression Inventory - 2;
BAI – Beck Anxiety Inventory;
SF-36 – the MOS 36-item short-form health survey
FIQ – Fibromyalgia Impact Questionnaire

Sample Size

Based on previous estimates of within-patient variation, *s*=2.45, from a previous study in fibromyalgia [36], we calculate that a sample of 55 trial completers would provide an 80% chance of detecting (alpha of .05) a mean treatment difference of 1 point (0 to 10 scale). For a sample size divisible by 6, the number of treatment sequences, we adjusted the sample size to 54 patients.

Statistical Analyses

Analyses for this trial are based on the null hypothesis of no difference between PGB, ALA, and PGB+ALA, and the alternative hypothesis that at least 2 treatments are different. When a patient contributes data from only 1 period, sensitivity analyses including all patients will also be performed by assuming some reasonable but extreme values for the remaining periods. All patients receiving at least 1 dose of a drug will be included in the safety analyses.

Primary Outcome Analysis

The primary outcome—mean daily pain from the last 7 days (at MTD) of each treatment period—will be calculated as an average of pain scores as recorded in the pain diary if more than 50% of the information (ie, at least 4 days) is not missing. Otherwise, mean daily pain will be treated as missing. This is based on the half rule often used to summarize repeated responses, which has proven unlikely to introduce bias to trial results [37]. Sensitivity analyses based on the average of all available pain scores will also be performed to confirm the results of the primary analysis. Although carryover effects are unlikely, we recognize this possibility. Therefore, a linear mixed model with sequence, period, treatment, and the first order carryover term as fixed effects and patient as a random effect [38] will be used to test for differences among the 3 treatments and to estimate the least square mean of the mean daily pain intensity for each treatment, adjusting for carryover as well as period effects (ie, stability of pain levels). The following 3 pair-wise comparisons will be performed based on the least square means and standard deviations from the linear mixed model: combination versus ALA alone, combination versus PGB alone, and ALA alone versus PGB alone. Sensitivity analyses will be performed using a pattern-mixture model [39] based on patterns of missing data so as to check the robustness of results in the case that data may not be missing at random. A Fisher's least significant difference [40] procedure will be used to adjust the *P* values for these 3 comparisons.

Secondary Analyses

Secondary outcomes will be analyzed similarly except that only one measurement is analyzed in the last week for the singular measures (ie, final week questionnaires) and the scoring

algorithms developed for the Brief Pain Inventory, Beck Depression Inventory-II, and Short Form-36 will be first used to derive the subscales or domains within these instruments, and the scores on these subscales or domains will be used as response variables in the linear mixed model analysis.

As with many other analgesic trials that allow concomitant medications, treatment group comparisons are made with the assumption in the setting of randomization of an equal distribution of concomitant medications across participants. Nevertheless, our trial analysis will further conduct exploratory analyses, as we have done on our previous trials, to investigate the possibility that concomitant analgesic medications have an important effect on the trial results.

Results

Participant recruitment is expected to begin in August 2017. This trial was awarded external peer-reviewed funding by the Canadian Institutes of Health Research–Strategy on Patient-Oriented Research–Canadian Pain Network in August 2016.

Discussion

Fibromyalgia continues to be difficult to manage, and current treatments provide only partial relief often at the risk of disabling AEs. To the best of our knowledge, this proposed trial is the first to compare the combination of an anticonvulsant with an antioxidant to treat fibromyalgia. Because ALA and PGB have different AE profiles, we expect their combination to provide superior analgesic efficacy in fibromyalgia without increasing AEs.

Possible threats to trial completion include challenges with participant recruitment, noncompliance, protocol violations, and early dropouts. However, we are confident that the proposed trial design and our experience with recent and previous RCTs will minimize these concerns. Noncompliance, protocol violations, and early dropouts will be minimized by the crossover design as well as thorough patient teaching and careful follow-up of trial participants.

Given the urgent need for improved fibromyalgia treatments that provide better pain relief with better safety and tolerability, this trial will provide rigorous evidence for a potentially improved treatment strategy for fibromyalgia.

Acknowledgments

The authors thank the Canadian Institutes of Health Research and the Queen's University Department of Anesthesiology and Perioperative Medicine.

Authors' Contributions

All authors read and approved the manuscript. IG led the writing of this manuscript and the development of this protocol. DT led the development of the statistical analysis plan and contributed to writing of this manuscript and development of the protocol. RRH participated in the writing of this manuscript and initial protocol development, including selection of mood and quality of life measures. TT participated in the initial protocol development. EV participated in the writing of this manuscript and initial protocol development. RM participated in the writing of this manuscript and initial protocol development. All authors will be involved in data analysis and interpretation and manuscript preparation.

Conflicts of Interest

IG has received support from Biogen, Adynxx, TARIS Biomedical, AstraZeneca, Pfizer, and Johnson and Johnson and has received grants from the Canadian Institutes of Health Research, Physicians' Services Incorporated Foundation, and Queen's University. RRH has received research funding from the Canadian Institutes of Health Research, the Social Sciences and Humanities Research Council of Canada, the American Foundation for Suicide Prevention, and Queen's University. The remaining authors have no conflicts of interest to declare.

Multimedia Appendix 1

Peer reviewer comments (CIHR-reviews-CADENCE-funded).

[PDF File (Adobe PDF File), 133KB - resprot_v6i8e154_app1.pdf]

References

1. Reitsma ML, Tranmer JE, Buchanan DM, Vandenkerkhof EG. The prevalence of chronic pain and pain-related interference in the Canadian population from 1994 to 2008. Chronic Dis Inj Can 2011 Sep;31(4):157-164 [FREE Full text] [Medline: 21978639]
2. St Sauver JL, Warner DO, Yawn BP, Jacobson DJ, McGree ME, Pankratz JJ, et al. Why patients visit their doctors: assessing the most prevalent conditions in a defined American population. Mayo Clin Proc 2013 Jan;88(1):56-67 [FREE Full text] [doi: 10.1016/j.mayocp.2012.08.020] [Medline: 23274019]

3. Institute of Medicine. Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research. Washington: National Academies Press; 2011.
4. Wolfe F. Fibromyalgia wars. J Rheumatol 2009 Apr;36(4):671-678. [doi: 10.3899/jrheum.081180] [Medline: 19342721]
5. Gelonch O, Garolera M, Valls J, Rosselló L, Pifarré J. Cognitive complaints in women with fibromyalgia: are they due to depression or to objective cognitive dysfunction? J Clin Exp Neuropsychol 2017 Mar 17:1-13. [doi: 10.1080/13803395.2017.1301391] [Medline: 28301977]
6. Clauw DJ. Fibromyalgia: a clinical review. JAMA 2014 Apr 16;311(15):1547-1555. [doi: 10.1001/jama.2014.3266] [Medline: 24737367]
7. Gran JT. The epidemiology of chronic generalized musculoskeletal pain. Best Pract Res Clin Rheumatol 2003 Aug;17(4):547-561. [Medline: 12849711]
8. Fitzcharles M, Ste-Marie PA, Goldenberg DL, Pereira JX, Abbey S, Choinière M, National Fibromyalgia Guideline Advisory Panel. 2012 Canadian Guidelines for the diagnosis and management of fibromyalgia syndrome: executive summary. Pain Res Manag 2013;18(3):119-126 [FREE Full text] [Medline: 23748251]
9. Thieme K, Mathys M, Turk DC. Evidenced-based guidelines on the treatment of fibromyalgia patients: are they consistent and if not, why Not? Have Effective Psychological Treatments Been Overlooked? J Pain 2016 Dec 27. [doi: 10.1016/j.jpain.2016.12.006] [Medline: 28034828]
10. Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Häuser W, Fluß E, et al. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis 2017 Feb;76(2):318-328. [doi: 10.1136/annrheumdis-2016-209724] [Medline: 27377815]
11. Mease PJ, Seymour K. Fibromyalgia: should the treatment paradigm be monotherapy or combination pharmacotherapy? Curr Pain Headache Rep 2008 Dec;12(6):399-405. [Medline: 18973731]
12. Gilron I, Chaparro LE, Tu D, Holden RR, Milev R, Towheed T, et al. Combination of pregabalin with duloxetine for fibromyalgia: a randomized controlled trial. Pain 2016 Dec;157(7):1532-1540. [doi: 10.1097/j.pain.0000000000000558] [Medline: 26982602]
13. Ziegler D, Nowak H, Kempler P, Vargha P, Low PA. Treatment of symptomatic diabetic polyneuropathy with the antioxidant alpha-lipoic acid: a meta-analysis. Diabet Med 2004 Feb;21(2):114-121. [Medline: 14984445]
14. Finnerup NB, Attal N, Haroutounian S, McNicol E, Baron R, Dworkin RH, et al. Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. Lancet Neurol 2015 Feb;14(2):162-173 [FREE Full text] [doi: 10.1016/S1474-4422(14)70251-0] [Medline: 25575710]
15. Gilron I, Tu D, Holden R, Towheed T, Ziegler D, Wang L, et al. Innovations in the management of musculoskeletal pain with alpha-lipoic acid (IMPALA trial): study protocol for a double-blind, randomized, placebo-controlled crossover trial of alpha-lipoic acid for the treatment of fibromyalgia pain. JMIR Res Protoc 2017 Mar 28;6(3):e41 [FREE Full text] [doi: 10.2196/resprot.7198] [Medline: 28351829]
16. Luo ZD, Chaplan SR, Higuera ES, Sorkin LS, Stauderman KA, Williams ME, et al. Upregulation of dorsal root ganglion (alpha)2(delta) calcium channel subunit and its correlation with allodynia in spinal nerve-injured rats. J Neurosci 2001 Mar 15;21(6):1868-1875 [FREE Full text] [Medline: 11245671]
17. Taylor CP, Angelotti T, Fauman E. Pharmacology and mechanism of action of pregabalin: the calcium channel alpha2-delta (alpha2-delta) subunit as a target for antiepileptic drug discovery. Epilepsy Res 2007 Feb;73(2):137-150. [doi: 10.1016/j.eplepsyres.2006.09.008] [Medline: 17126531]
18. Nagamatsu M, Nickander KK, Schmelzer JD, Raya A, Wittrock DA, Tritschler H, et al. Lipoic acid improves nerve blood flow, reduces oxidative stress, and improves distal nerve conduction in experimental diabetic neuropathy. Diabetes Care 1995 Aug;18(8):1160-1167. [Medline: 7587852]
19. Lee WY, Orestes P, Latham J, Naik AK, Nelson MT, Vitko I, et al. Molecular mechanisms of lipoic acid modulation of T-type calcium channels in pain pathway. J Neurosci 2009 Jul 29;29(30):9500-9509 [FREE Full text] [doi: 10.1523/JNEUROSCI.5803-08.2009] [Medline: 19641113]
20. Han T, Bai J, Liu W, Hu Y. A systematic review and meta-analysis of alpha-lipoic acid in the treatment of diabetic peripheral neuropathy. Eur J Endocrinol 2012 Oct;167(4):465-471 [FREE Full text] [doi: 10.1530/EJE-12-0555] [Medline: 22837391]
21. Ziegler D, Low PA, Litchy WJ, Boulton AJM, Vinik AI, Freeman R, et al. Efficacy and safety of antioxidant treatment with alpha-lipoic acid over 4 years in diabetic polyneuropathy: the NATHAN 1 trial. Diabetes Care 2011 Sep;34(9):2054-2060 [FREE Full text] [doi: 10.2337/dc11-0503] [Medline: 21775755]
22. López-D'alessandro E, Escovich L. Combination of alpha lipoic acid and gabapentin, its efficacy in the treatment of Burning Mouth Syndrome: a randomized, double-blind, placebo controlled trial. Med Oral Patol Oral Cir Bucal 2011 Aug 01;16(5):e635-e640 [FREE Full text] [Medline: 20711135]
23. Crofford LJ, Mease PJ, Simpson SL, Young JP, Martin SA, Haig GM, et al. Fibromyalgia relapse evaluation and efficacy for durability of meaningful relief (FREEDOM): a 6-month, double-blind, placebo-controlled trial with pregabalin. Pain 2008 Jun;136(3):419-431. [doi: 10.1016/j.pain.2008.02.027] [Medline: 18400400]
24. Felson DT, Goldenberg DL. The natural history of fibromyalgia. Arthritis Rheum 1986 Dec;29(12):1522-1526. [Medline: 3801074]

25. Forseth KO, Førre O, Gran JT. A 5.5 year prospective study of self-reported musculoskeletal pain and of fibromyalgia in a female population: significance and natural history. Clin Rheumatol 1999;18(2):114-121. [Medline: 10357115]
26. Kennedy M, Felson DT. A prospective long-term study of fibromyalgia syndrome. Arthritis Rheum 1996 Apr;39(4):682-685. [Medline: 8630121]
27. Wolfe F, Clauw DJ, Fitzcharles M, Goldenberg DL, Häuser W, Katz RL, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016 Dec;46(3):319-329. [doi: 10.1016/j.semarthrit.2016.08.012] [Medline: 27916278]
28. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001 Nov;94(2):149-158. [Medline: 11690728]
29. Melzack R. The short-form McGill Pain Questionnaire. Pain 1987 Aug;30(2):191-197. [Medline: 3670870]
30. Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. J Rheumatol 1991 May;18(5):728-733. [Medline: 1865419]
31. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singapore 1994 Mar;23(2):129-138. [Medline: 8080219]
32. Beck A, Steer R, Brown G. BDI-II: Beck Depression Inventory Manual, 2nd Edition. San Antonio: Harcourt Brace and Company; 1996.
33. Steer RA, Kumar G, Ranieri WF, Beck AT. Use of the Beck Anxiety Inventory with adolescent psychiatric outpatients. Psychol Rep 1995 Apr;76(2):459-465. [doi: 10.2466/pr0.1995.76.2.459] [Medline: 7667457]
34. Ware J, Snow K, Kosinski MB. SF-36 Health Survey Manual and Interpretation Guide. Boston: The Health Institute; 1993.
35. Ioannidis JPA, Evans SJW, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004 Nov 16;141(10):781-788. [Medline: 15545678]
36. Russell IJ, Mease PJ, Smith TR, Kajdasz DK, Wohlreich MM, Detke MJ, et al. Efficacy and safety of duloxetine for treatment of fibromyalgia in patients with or without major depressive disorder: Results from a 6-month, randomized, double-blind, placebo-controlled, fixed-dose trial. Pain 2008 Jun;136(3):432-444. [doi: 10.1016/j.pain.2008.02.024] [Medline: 18395345]
37. Fairclough D. Analyzing studies with missing data. In: Fairclough D, Fayers PM, Hays RD, editors. Assessing Quality of Life in Clinical Trials, 2nd Edition. Oxford: Oxford University Press; 2005:182.
38. Jones B, Kenward MG. Design and Analysis of Cross-Over Trials, 3rd Edition. Boca Raton: CRC Press; 2015.
39. Molenberghs G, Kenward MG. Missing Data in Clinical Studies. Hoboken: Wiley; 2007.
40. Hochberg Y, Tamhane A. Multiple Comparison Procedures. New York: Wiley; 1987.

Abbreviations

AE: adverse effect
ALA: alpha-lipoic acid
MTD: maximal tolerated dose
PGB: pregabalin
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 08.05.17; peer-reviewed by H Shanthana; comments to author 14.07.17; revised version received 19.07.17; accepted 21.07.17; published 04.08.17.

Please cite as:
Gilron I, Tu D, Holden R, Towheed T, Vandenkerkhof E, Milev R
Combination Analgesic Development for Enhanced Clinical Efficacy (CADENCE Trial): Study Protocol for a Double-Blind, Randomized, Placebo-Controlled Crossover Trial of an Alpha-Lipoic Acid – Pregabalin Combination for the Treatment of Fibromyalgia Pain
JMIR Res Protoc 2017;6(8):e154
URL: http://www.researchprotocols.org/2017/8/e154/
doi:10.2196/resprot.8001
PMID:28778847

Protocol

Impact of a Tutored Theoretical-Practical Training to Develop Undergraduate Students' Skills for the Detection of Caries Lesions: Study Protocol for a Multicenter Controlled Randomized Study

Mariana Minatel Braga[1,2*], DDS, PhD; Tathiane Larissa Lenzi[1,3,4*], DDS, MSc PhD; Fernanda Rosche Ferreira[1], MSc, DDS; Fausto Medeiros Mendes[1], DDS, MSc PhD; Daniela Prócida Raggio[1], MSc, DDS, PhD; José Carlos Imparato[1,5], MSc, DDS, PhD; Marcelo Bonecker[1], MSc, DDS, PhD; Ana Carolina Magalhães[6], DDS, MSc PhD; Linda Wang[6], MSc, DDS, PhD; Daniela Rios[6], MSc, DDS, PhD; Juliano Pelim Pessan[7], DDS, MSc PhD; Cristiane Duque[7], DDS, PhD; Maria Augusta Bessa Rebelo[8], DDS, MSc PhD; Ary Oliveira Alves Filho[8], MSc, DDS, PhD; Marina De Deus Moura Lima[9], MSc, DDS, PhD; Marcoeli Silva Moura[9], MSc, DDS, PhD; Alessandro Diogo De Carli[10], MSc, DDS, PhD; Mariane Emi Sanabe[11], DDS, MSc PhD; Maximiliano Sergio Cenci[12], DDS, MSc PhD; Elenara Ferreira Oliveira[12], DDS, MSc PhD; Marcos Britto Correa[12], MSc, DDS, PhD; Rachel Oliveira Rocha[3], MSc, DDS, PhD; Julio Eduardo Zenkner[3], MSc, DDS, PhD; Pedroza Uribe Murisí[1,13], MSc, DDS; Stefania Martignon[14,15], DDS, PhD; Juan Sebastian Lara[1,16,17], DDS, MSc PhD; Fatima Gabriela Aquino[18], Msc Candidate, DDS; Alfredo Carrillo[19], Msc Candidate, DDS; Chun Hung Chu[20], MSc, DDS, PhD; Chris Deery[21], MSc, BDS, PhD; David Ricketts[22], BDS, MSc PhD; Paulo Melo[23], PhD; José Leopoldo Ferreira Antunes[24], MSc PhD; Kim Rud Ekstrand[25], DDS, PhD; IuSTC Group[2*]

[1]Dental School, Pediatric Dentistry Departament, University of São Paulo, Sao Paulo, Brazil

[2]Initiatives for undergraduate Students' Training in Cariology (IuSTC Group), Multi-institutional group, São Paulo, Brazil

[3]Dental School, Federal University of Santa Maria, Santa Maria, Brazil

[4]Dental School, Federal University of Minas Gerais, Belo Horizonte, Brazil

[5]CPO Sao Leopoldo Mandic, Campinas, SP, Brazil

[6]School of Dentistry, University of São Paulo, Bauru, Brazil

[7]School of Dentistry, São Paulo State University (Unesp), Araçatuba, Brazil

[8]School of Dentistry, Federal University of Amazonas, Manaus, Brazil

[9]Postgraduate Programme in Dentistry, Federal University of Piaui, Teresina, PI, Brazil

[10]School of Dentistry, Federal University of Mato Grosso do Sul, Campo Grande, MS, Brazil

[11]School of Dentistry "Prof Albino Coimbra Filho", Federal University of Mato Grosso do Sul, Campo Grande, MS, Brazil

[12]Graduate Program in Dentistry, Federal University of Pelotas, Pelotas, RS, Brazil

[13]Facultad de Odontologia, Universidad de Guadalajara, Guadalajara, Mexico

[14]UNICA - Caries Research Unit, Research Vice-rectory,, Universidad El Bosque, Bogotá, Colombia

[15]Dental Innovation and Translation Centre, King's College of London, London, United Kingdom

[16]UNICA - Caries Research Unit, Research Vice-rectory, Universidad El Bosque, Bogotá, Colombia

[17]The Dental Health Unit, The University of Manchester, Manchester, United Kingdom

[18]Carrera de Odontología, Universidad Católica, Asunción, Paraguay

[19]Facultad Autonoma de Asunción, Asunción, Paraguay

[20]Faculty of Dentistry, The University of Hong Kong, Hong Kong, China (Hong Kong)

[21]University of Sheffield, Sheffield, United Kingdom

[22]Dundee Dental Hospital and School, University of Dundee, Dundee, United Kingdom

[23]Faculty of Dental Medicine, EpiUnit, University of Porto, Porto, Portugal

[24]School of Public Health, University of São Paulo, São Paulo, Brazil

[25]Department of Odontology, University of Copenhagen, Copenhagen, Denmark

[*]these authors contributed equally

Corresponding Author:
Mariana Minatel Braga, DDS, PhD
Dental School
Pediatric Dentistry Departament

University of São Paulo
Av. Professor Lineu Prestes, 2227
Sao Paulo, 05508-000
Brazil
Phone: 55 1130917835
Fax: 55 1130917854
Email: mmbraga@usp.br

Abstract

Background: Tutored laboratorial activities could be a manner of improving the competency development of students. However, its impact over conventional theoretical classes has not yet been tested. Additionally, different university contexts could influence this issue and should be explored.

Objective: To assess the impact of a tutored theoretical-practical training for teaching undergraduate students to detect caries lesions as compared with theoretical teaching activities. The impact of these teaching/learning activities will be assessed in terms of efficacy, cost/benefit, retention of knowledge/acquired competences, and student acceptability.

Methods: Sixteen centers (7 centers from Brazil and 9 centers from other countries throughout the world) are involved in the inclusion of subjects for this protocol. A randomized controlled study with parallel groups will be conducted. One group (control) will be exposed to a 60- to 90-minute conventional theoretical class and the other group (test) will be exposed to the same theoretical class and also a 90-minute laboratory class, including exercises and discussions based on the evaluation of a pool of images and extracted teeth. The mentioned outcomes will be evaluated immediately after the teaching activities and also in medium- and long-term analyses. To compare the long-term outcomes, students who enrolled in the university before the participating students will be interviewed for data collection and these data will be used as a control and compared with the trained group. This stage will be a nonrandomized phase of this study, nested in the main study. Appropriate statistical analysis will be performed according to the aims of this study. Variables related to the centers will also be analyzed and used to model adjustment as possible sources of variability among results.

Results: This ongoing study is funded by a Brazilian national funding agency (CNPq- 400736/2014-4). We expect that the tutored theoretical-practical training will improve the undergraduate students' performance in the detection of caries lesions and subsequent treatment decisions, mainly in terms of long-term retention of knowledge. Our hypothesis is that tutored theoretical-practical training is a more cost-effective option for teaching undergraduate students to detect caries lesions.

Conclusions: If our hypothesis is confirmed, the use of laboratory training in conjunction with theoretical classes could be used as an educational strategy in Cariology to improve the development of undergraduate students' skills in the detection of caries lesions and clinical decision-making.

(JMIR Res Protoc 2017;6(8):e155) doi:10.2196/resprot.7414

KEYWORDS

active learning; cariology; dental education; laboratory training

Introduction

Background

There is an imminent demand to prepare dental professionals to not only possess knowledge, but also have developed skills to detect dental caries and make decisions about caries management based upon the current best scientific evidence [1]. Thus, an improved education in Cariology for dental undergraduate students as future practitioners is fundamental for adequate caries diagnoses [2].

A recent systematic review has pointed out that caries diagnosis is the topic within Cariology that has received the most concern in terms of developing training methodologies and/or education for undergraduate students [3]. This suggests not only the importance of the subject within the dental professional training, but also the difficulty involved in successfully preparing professionals to be able to perform this step in the dental clinic.

Although it may appear to be a simple step, the detection of caries lesions, as a part of the diagnosis process, involves several aspects that require not only knowledge about the subject, but also theoretical, clinical, and interpersonal skills. In this sense, practical training may be essential for the development of these specific skills [4]. The laboratory activities systematically provide opportunities for undergraduate students to experience different simulated situations covering a variety of circumstances that they should be able to solve in clinical practice. Moreover, practical training allows more interactivity between students and teachers and/or tutors in the teaching-learning process [5,6], besides providing rapid feedback on their performance [7]; thus, improving the educational outcome.

Several systems for caries detection have been used, to provide important information about the disease and also to guide professionals in making treatment decisions [8]. The International Caries Detection and Assessment System (ICDAS) is one of the options proposed to standardize this process

internationally [9]. Although student performance in the detection of caries lesions using this system has already been assessed [10-12], these investigations were performed in small groups of students or in only 1 institutional center, reducing the external validity of these findings [13].

A preliminary study conducted at the University of São Paulo found that the tutored theoretical-practical training seems to improve the caries detection by undergraduate students using ICDAS [14]. However, it was not possible to assess the real impact of this active-learning methodology in the teaching of this field, which involves more time and costs, as compared with the theoretical activity alone. Furthermore, this effect may be dependent on different university contexts, mainly because caries diagnosis content has not been taught in the same way in different curricula. To the best of our knowledge, this is a pioneering multicenter study that aims to investigate the short- and long-term impact of the tutored theoretical-practical training for detecting caries lesions in undergraduate courses in comparison with theoretical activities.

Objective

The aim of the present protocol will be to evaluate the effect of a tutored theoretical-practical training for teaching undergraduate students to detect caries lesions in comparison with theoretical teaching activities. The impact of these teaching/learning activities will be assessed in terms of efficacy, cost/benefit, retention of knowledge/acquired competences, and student acceptance.

Methods

Ethical Aspects

The Research Ethics Committee of all institutions involved previously approved this protocol. The participants will receive and sign an informed consent prior to their involvement in the research. For ethical reasons, the students allocated to the theoretical group will receive laboratory training at the end the outcome assessments. Participant confidentiality will be ensured using identification code numbers. Participant identifiable information will be stored in locked filing cabinets in a secure room.

Study Design

A multicenter randomized controlled study with parallel groups will be conducted. One group (control) will be exposed to a 60- to 90-minute conventional theoretical class and the other group (test) will be exposed to the same theoretical class along with a 90-minute laboratory class, including exercises and discussions based on the evaluation of a pool of images and extracted teeth.

This study will involve 7 centers from Brazil and 9 centers from other countries throughout the world (Figure 1). The centers represent two distinct contexts: 1 coordinating center, a precursor in the development of the theoretical and laboratory activity that will be tested, while the other institutions represent those that will receive the proposal to develop the same activity with support provided by the precursor center. The inclusion of different centers aims to cover different regions and institutions with different profiles to increase the external validity of the study. As the coordinating center is in São Paulo, Brazil, 6 other institutions were invited to cover different parts of the country. Similarly, the other centers throughout the world are located in different regions, including Latin America (Colombia and Paraguay), Asia (Hong Kong), and Europe (Copenhagen, United Kingdom, and Portugal). Other centers may be optionally incorporated throughout the study, according to interest and judgment of necessity. The Initiatives for Undergraduate Students' Training in Cariology is a collaborative group made up of all people involved in this study. The detailed roles of each member and respective affiliations are described in Multimedia Appendix 1.

Training for Operationalization of the Study in Different Centers

The researchers responsible for the study in each center will participate in an initial joint meeting to clarify the methodology to be followed in the study. Training and collaboration of the participants from each center will be made in order to properly adapt the logistics needed for the study, considering the characteristics of each center. The research team will detect all particularities of each center, and a proposal for the implementation of the activity will be jointly developed by the coordinating center and each participating institution. In addition, the principal investigator will train the staff members that will apply the activity at each center.

Preparation of the Material for Evaluations

The coordinating center will be responsible for organizing and/or assisting in the organization of the material to be used in the theoretical-practical training, including preparation of a pool of images and extracted teeth contemplating different stages of caries severity. The pool of images will be prepared and used in all centers. The teeth will be obtained from the Bank of Human Teeth located at each institution or similar donators of material for this purpose. A tutorial was developed to guide investigators from each center to prepare their samples aided by an investigator from the coordinating center.

Figure 1. Mapping of the institutions involved in the study.

Experimental Groups

Undergraduate students at the beginning (up to second or third year) and the end of the Dentistry course (last 2 years) will be included. All students will be invited to participate, but only those who consent to participate will be included. The classes will be randomly divided into 2 groups. The randomization will be made by a central allocation located in the coordinating center, by sending the list of students to be invited. All students will be randomized, even those that will not participate. These losses will be computed and analyzed later.

One group (control) will be exposed to a 60- to 90-minute conventional theoretical class and the other group (test) will be exposed to the same theoretical class and a 90-minute laboratory class, including exercises and discussions based on the evaluation of a pool of images and extracted teeth. All students will attend the class together before the allocation. For the test group, the participants will be assigned in small groups (7-10 students) mediated by graduate tutors and/or teachers. Students will assess 30 clinical images and 10 teeth for severity (ICDAS scores) and activity status [14]. Tutors/teachers will work on correcting exercises and also discussion of questions raised by the group. In case of disagreement, tooth revaluation will be possible with the help of tutors/teachers so that students understand potential mistakes. For the control group, they will begin the outcomes assessment after having finished the theoretical class.

Despite the allocation, both groups will begin their activities (for training or outcomes assessment) at the same time, avoiding demotivation of any group. Students and teachers/tutors will know the allocated group only when they will be prepared for the activities after the theoretical class, guaranteeing the allocation concealment.

Outcomes Assessment

Cost Efficacy

To evaluate the efficacy of the teaching and learning methods both the theoretical knowledge, and the practical skills and clinical decision-making capacity of the students will be evaluated after the theoretical class (control group) and after the theoretical laboratory training (test group).

To evaluate the theoretical knowledge, the students will answer questions about caries lesions detection prepared by professors with expertise in this field. Furthermore, the participants must understand the impact of the correct decision or error when using the ICDAS in clinical practice. The practical assessment will be made through the evaluation of extracted teeth for 1 minute each. Students will evaluate selected tooth surfaces according to the ICDAS criteria with the aid of a light reflector, a plane buccal mirror, and a ballpoint probe. To simulate the distance to observe teeth in the mouth, teeth will be positioned on a plane surface approximately 30 cm from the examiners' eyes. The teeth will first be examined wet, and then they will be dried for 5 seconds with compressed air. For this evaluation, the students' responses regarding caries severity will be compared with a template made by teachers after visual inspection. To evaluate the capacity of decision making, the undergraduate students should propose the treatment option based on caries lesion detection for 5 clinical cases. Cases were prepared to simulate usual clinical conditions. To assess the cost efficacy of the methods, the efficacy will be considered the number of true answers registered by the students independently on the theoretical and practical tests, as well as in the decision-making query. Direct and indirect costs will be calculated, both individually (the activity itself) and institutionally (to maintain the necessary structure and provide time for the activity). Cost estimation should include planning, preparation of the material, staff training, and implementation of the activity. Lecture and practical activity costing data will be directly collected at all study sites considering institutional information itself. In addition, national databases will be

searched to estimate costs of maintenance of universities in each location. Data from these 2 sources will be collated to produce tables of costs for the activity, center, and country.

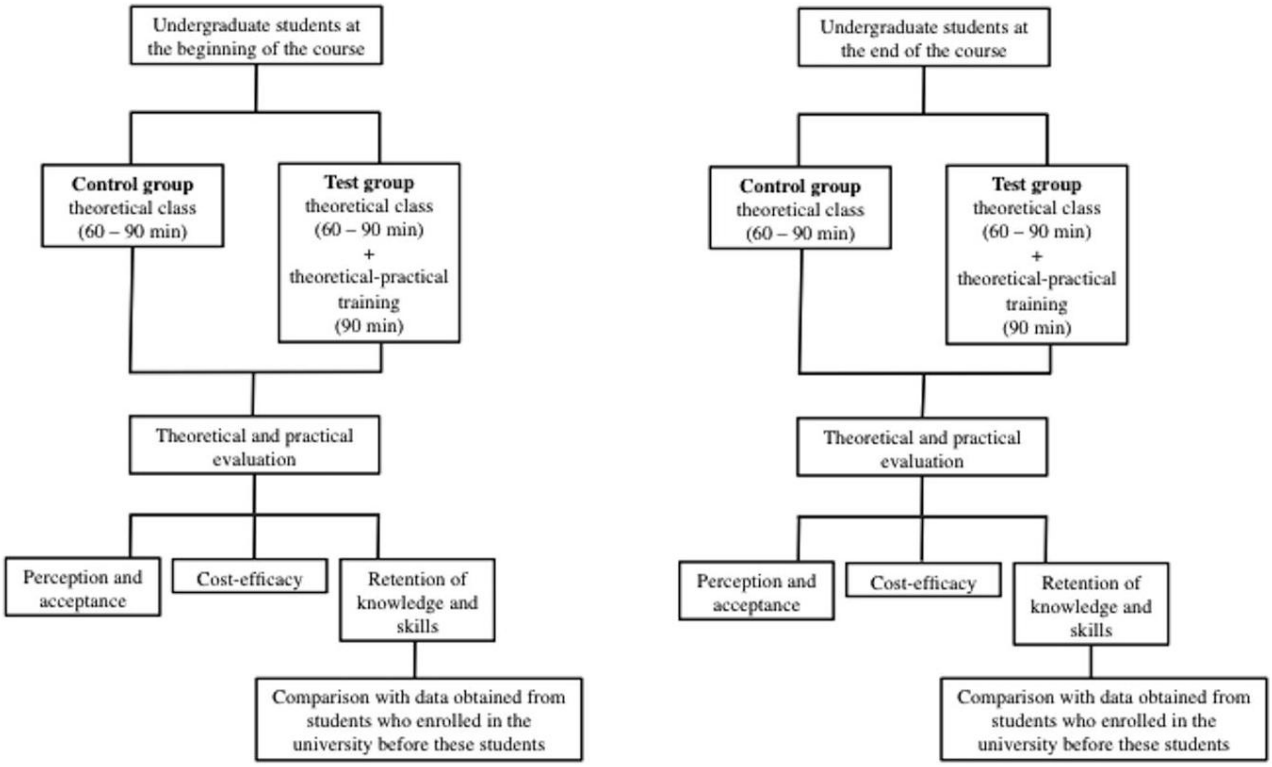
Perception and Acceptance

The students' perception of the activity and their degree of satisfaction with the knowledge achieved will be assessed. Questions will evaluate the preparation level, nervousness, satisfaction, and acceptance of students regarding the proposed activities. These outcomes will be evaluated based on the State Trait Personality Inventory Scale [15]. Moreover, the goal of this outcome will verify if the students who participated in the laboratory training felt more prepared for the caries lesions detection in a clinical situation.

Retention of Knowledge and Skills

To evaluate the retention of knowledge and skills acquired after the tutored theoretical laboratory training, students who entered the university before the participating students will be interviewed for data collection, and these data will be used as control and compared with the trained group. This stage will be a nonrandomized phase of the study, nested in the main study. To compare the long-term outcomes, a theoretical and practical assessment will be applied to participating students after 1 or 2 years following the first training and compared with a class in the same year of the undergraduate program, but who had not been exposed to this methodology. This stage of study will be adjusted considering the curriculum of each center in accordance with the course subjects offered in the periods of interest. The flowchart of the study design is summarized in Figure 2.

Figure 2. Study design and phases to be conducted.



Center/Student Characteristics

Some characteristics related to each center will be collected during the planning and execution/data collection stages. In the planning phase, a structured questionnaire will be applied to the center coordinators in order to identify relevant aspects of center size/potential, curricula, human resources/faculty preparation, infrastructure/facilities, and expected difficulties for participant centers. Afterward, in the execution phase, we will first map each student's background and willingness to learn Cariology and caries detection by a form they should fill out before the theoretical class. The student's background will show both the student's knowledge at baseline, and also the possible contents taught at his/her center. Second, qualitative research on topics and strategies used in teaching/learning of Cariology will be investigated among students and lecturers in different centers. These quantitative-qualitative variables will be used as independent variables related to the center in further analysis.

Statistical Analysis

For the theoretical evaluation, total number of correct responses for test and control groups will be compared at both times of evaluation (immediate and long term). Data from the number of correct responses per question will be submitted to analysis, also considering the question as a factor.

For the practical assessment, performance-related parameters (sensitivity, specificity, accuracy) when using the visual method for the caries detection and evaluation of lesions activity will be calculated. These values will be compared between groups

using multilevel analysis, considering both experimental groups and different centers.

The scores given by students for acceptance and perception of the activities carried out will be compared between groups. Regression analysis will be used in each experimental group to verify the association between student performance in caries lesions detection versus theoretical knowledge and student perception in relation to the activity. Baseline student knowledge and willingness to learn caries detection will be considered during these analyses.

The association between the results of immediate theoretical and practical assessments with knowledge and skills retained in the long term will be assessed.

For all analyses, a multilevel approach will be used to consider different levels explored. The assessment, the student, the center, and the country could be considered as possible levels for these analyses considering the possible clustering observed for centers and, eventually, countries. For all the parameters tested, the influence of the different centers and the level in the course (first or end years) will be investigated by regression analysis. The significance level will be set at 5%.

Results

This protocol refers to an ongoing study mainly funded by a Brazilian national funding agency (CNPq- 400736/2014-4). The standardized protocol has been implemented in different centers and data collected throughout them. On these occasions, different situations have been experienced and all peculiarities have been recorded in order to help in explaining possible differences among the studied centers.

We expect that the tutored theoretical-practical training will improve the undergraduate students' performance in the detection of caries lesions and subsequent treatment decisions, mainly in terms of long-term retention of knowledge. Our hypothesis is that tutored theoretical-practical training is a more cost-effective option for teaching undergraduate students to detect caries lesions.

Discussion

Motivation and Design

In aiming to formalize the curriculum in Cariology at the 1st Consensus Workshop on the Development of the European Curriculum in Cariology [1], 5 distinct areas were highlighted [2], including the caries diagnosis, which involves risk assessment and detection of caries lesions [16]. However, theoretical classes, often adopted for teaching Cariology, may not be enough to enable those undergraduate students to perform procedures, such as the detection of caries lesions in clinical practice. In this regard, we expect this study to provide the best scientific evidence for defining the best teaching/learning strategy for detection of caries.

As in other areas of dentistry, practical training may be essential for the development of these specific skills that are expected of a future professional [4]. A preliminary study conducted by our research group showed that laboratory activity could help even in solving theoretical questions that the students may have related to caries diagnosis [14]. This is a before and after type of study, in which the same students are evaluated at different times of the laboratory activities. Nevertheless, no study to date has compared this new methodology with theoretical classes. Although the practical activities tend to reinforce the theoretical content, the real cost-efficacy of using this methodology is unknown.

One of the motivations of this multicenter study was checking the impact of the same training activity considering different contexts. The inclusion of several university centers with different students profiles may increase the actual external validity of the proposed activity [13]. Therefore, influence of center dimension, differences in curricula, workflow, faculty formation/willingness to participate in the project, available human resources, and infrastructure/facilities can be tested using this study design. Moreover, distinct universities resources could interfere in the educational outcome or demand an extra investment for implementing the activity.

Most studies that assessed the efficacy or acceptance of students on measures associated with the theoretical classes [12,17,18], but the comparison of these measures with the performance achieved only after theoretical classes is rare [19]. Furthermore, they did not estimate the additional cost for implementation of this alternative, pointing out an innovative approach of our protocol. Once again, the multicenter study design that will permit to compare cost implementation in the different tested contexts [20].

Expectations

Among positive aspects, active methodologies tend to stimulate higher participation of students in the learning-teaching process [21]. Also, we speculate that by the end of the course undergraduate students could benefit more from this approach than those at the beginning of the course. The cumulative knowledge of the faculty associated with practical training may result in a better performance of dental undergraduate students for application of this integrated knowledge while making decisions related to prevention and management of caries [22].

A recent systematic review [8] has pointed out that the visual caries detection method has good overall performance, and the use of indices improves the accuracy of the method. With the expected results, we aim to achieve the inclusion of the teaching/learning of caries detection using the ICDAS into the curriculum of undergraduate courses. In many clinical disciplines, the content taught in laboratories may have been reduced due to a need to gain time in the clinical environment. On the other hand, considering that laboratory training may help the students solve doubts before clinical examinations and optimize chair time, we believe that tutored theoretical-practical training is a cost-effective option for teaching undergraduate students to detect caries lesions. If the results confirm our hypotheses, this study will contribute significantly to the reformulation of Cariology curriculum guidelines for undergraduate students.

Acknowledgments

This study is funded by the National Council for Scientific and Technological Development (CNPq-400736/2014-4) and The São Paulo Research Foundation (FAPESP), Brazilian funding agencies. The Principal Investigator and other Center Coordinators also receive scholarships as awards for Research Productivity in Brazil from the CNPq.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Initiatives for Undergraduate Students' Training in Cariology collaborative (IuSTC) group.

[PDF File (Adobe PDF File), 419KB - resprot_v6i8e155_app1.pdf]

Multimedia Appendix 2

Funding agency report.

[PDF File (Adobe PDF File), 196KB - resprot_v6i8e155_app2.pdf]

Multimedia Appendix 3

Funding agency report 2.

[PDF File (Adobe PDF File), 185KB - resprot_v6i8e155_app3.pdf]

References

1. Schulte A, Pitts N, Huysmans M, Splieth C, Buchalla W. European core curriculum in cariology for undergraduate dental students. Caries Res 2011;45:345. [doi: 10.1159/000330006]
2. Schulte A, Buchalla W, Huysmans M, Amaechi B, Sampaio F, Vougiouklakis G. A survey on education in cariology for undergraduate dental students in Europe. Eur J Dent Educ 2011;15(Suppl 1):3-8. [doi: 10.1111/j.1600-0579.2011.00708.x]
3. Braga MM, Floriano I, Lara JS, Haddad AE, Mattos-Silveira J, Mendes FM. PROSPERO records. 2014. Educational and teaching techniques in cariology: a systematic review URL: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014010343 [accessed 2017-07-28] [WebCite Cache ID 6sJ16Y72I]
4. Cowpe J, Plasschaert A, Harzer W, Vinkka-Puhakka H, Walmsley A. Profile and competences for the graduating European dentist - update 2009. Eur J Dent Educ 2010;14:193-202. [doi: 10.1111/j.1600-0579.2009.00609.x]
5. Tsai C, Aida M, Wan Zah Wan A, AbRahim B. Utilization of intelligent tutoring system in mathematics learning. Int J Educ Dev using Inf Commun Technol 2008;4:50-63.
6. Pennaforte T, Moussa A, Loye N, Charlin B, Audétat M. Exploring a new simulation approach to improve clinical reasoning teaching and assessment: randomized trial protocol. JMIR Res Protoc 2016;5:e26 [FREE Full text] [doi: 10.2196/resprot.4938] [Medline: 26888076]
7. Divaris K, Barlow P, Chendea S, Cheong W, Dounis A, Dragan I. The academic environment: the students? perspective. Eur J Dent Educ 2008;12:120-130. [doi: 10.1111/j.1600-0579.2007.00494.x]
8. Gimenez T, Piovesan C, Braga M, Raggio D, Deery C, Ricketts D. Visual inspection for caries detection: a systematic review and meta-analysis. J Dent Res 2015;94:895-904. [doi: 10.1177/0022034515586763]
9. Pitts N. ICDAS --an international system for caries detection and assessment being developed to facilitate caries epidemiology, research and appropriate clinical management. Community Dent Health 2004;21(3):193-198. [Medline: 15470828]
10. Foley J. Dental students consistency in applying the ICDAS system within paediatric dentistry. Eur Arch Paediatr Dent 2012;13:319-322. [Medline: 23235133]
11. Diniz M, Lima L, Santos-Pinto L, Eckert G, Zandoná A, de Cássia Loiola Cordeiro R. Influence of the ICDAS e-learning program for occlusal caries detection on dental students. J Dent Educ 2010;74:862-868. [Medline: 20679455]
12. Zandona A, Al-Shiha S, Eggertsson H, Eckert G. Student versus faculty performance using a new visual criteria for the detection of caries on occlusal surfaces: an in vitro examination with histological validation. Oper Dent 2009;34:598-604. [doi: 10.2341/08-082-L]
13. Reed D, Cook D, Beckman T, Levine R, Kern D, Wright S. Association between funding and quality of published medical education research. JAMA 2007;298(9):1002-1009. [doi: 10.1001/jama.298.9.1002]
14. Tibério B, Floriano I, Novaes T, Matos R, Camargo L, Bonecker M. Human-based interactive laboratorial training on ICDAS as teaching strategy to undergraduate students. In: Caries Res. 2013 Presented at: 60th ORCA Congress; July 3-6, 2013; Liverpool, UK p. 2013 URL: https://www.karger.com/Article/Abstract/350670 [doi: 10.1159/000350670]

15. Marteau T, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). Br J Clin Psychol 1992;31:301-306. [Medline: 1393159]
16. Pitts N, Melo P, Martignon S, Ekstrand K, Ismail A. Caries risk assessment, diagnosis and synthesis in the context of a European Core Curriculum in Cariology. Eur J Dent Educ 2011;15(Suppl 1):23-31. [doi: 10.1111/j.1600-0579.2011.00711.x]
17. Diniz M, Lima L, Santos-Pinto L, Eckert G, Zandoná A, de Cássia Loiola Cordeiro R. Influence of the ICDAS e-learning program for occlusal caries detection on dental students. J Dent Educ 2010;74:862-868. [Medline: 20679455]
18. Luz P, Stringhini C, Otto B, Port A, Zaleski V, Oliveira R. Performance of undergraduate dental students on ICDAS clinical caries detection after different learning strategies. Eur J Dent Educ 2015;19:235-241. [doi: 10.1111/eje.12131]
19. Jablonski-Momeni A, Busche J, Struwe C, Lange J, Heinzel-Gutenbrunner M, Frankenberger R. Use of the international caries detection and assessment system two-digit coding method by predoctoral dental students at Philipps University of Marburg, Germany. J Dent Educ 2012;76:1657-1666.
20. Mugford M, Hutton G, Fox-Rushby J. Methods for economic evaluation alongside a multicentre trial in developing countries: a case study from the WHO Antenatal Care Randomised Controlled Trial. J Paediatr Perinat Epidemiol 1998;12(Suppl 2):75-97. [Medline: 9805724]
21. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H. Active learning increases student performance in science, engineering, and mathematics. Proc Natl Acad Sci U S A 2014;111:8410-8415. [doi: 10.1073/pnas.1319030111]
22. Hendricson W, Anderson E, Andrieu S, Chadwick D, Cole J, George M. Does faculty development enhance teaching effectiveness? J Dent Educ 2007;71:1513-1533. [Medline: 18096877]

Abbreviations

ICDAS: International Caries Detection and Assessment System

Edited by G Eysenbach; submitted 08.02.17; peer-reviewed by MC Skelton, K Yin; comments to author 13.04.17; revised version received 08.06.17; accepted 29.06.17; published 16.08.17.

Please cite as:
Braga MM, Lenzi TL, Ferreira FR, Mendes FM, Raggio DP, Imparato JC, Bonecker M, Magalhães AC, Wang L, Rios D, Pessan JP, Duque C, Rebelo MAB, Alves Filho AO, Lima MDDM, Moura MS, De Carli AD, Sanabe ME, Cenci MS, Oliveira EF, Correa MB, Rocha RO, Zenkner JE, Murisí PU, Martignon S, Lara JS, Aquino FG, Carrillo A, Chu CH, Deery C, Ricketts D, Melo P, Antunes JLF, Ekstrand KR, IuSTC Group
Impact of a Tutored Theoretical-Practical Training to Develop Undergraduate Students' Skills for the Detection of Caries Lesions: Study Protocol for a Multicenter Controlled Randomized Study
JMIR Res Protoc 2017;6(8):e155
URL: http://www.researchprotocols.org/2017/8/e155/
doi:10.2196/resprot.7414
PMID:28814381

Protocol

Systematic Review Protocol to Assess the Effectiveness of Usability Questionnaires in mHealth App Studies

Leming Zhou[1], ScD, PhD; Jie Bao[1], MS; Bambang Parmanto[1], PhD

University of Pittsburgh, Department of Health Information Management, Pittsburgh, PA, United States

Corresponding Author:
Leming Zhou, ScD, PhD
University of Pittsburgh
Department of Health Information Management
6021 Forbes Tower
3600 Forbes Ave at Atwood St
Pittsburgh, PA, 15260
United States
Phone: 1 412 383 6653
Fax: 1 412 383 6655
Email: lmzhou@gmail.com

Abstract

Background: Usability questionnaires have a wide use in mobile health (mHealth) app usability studies. However, no systematic review has been conducted for assessing the effectiveness of these questionnaires.

Objective: This paper describes a protocol for conducting a systematic review of published questionnaire-based mHealth app usability studies.

Methods: In this systematic review, we will select recently published (2008-2017) articles from peer-reviewed journals and conferences that describe mHealth app usability studies and implement at least one usability questionnaire. The search strategy will include terms such as “mobile app” and “usability.” Multiple databases such as PubMed, CINAHL, IEEE Xplore, ACM Digital Library, and INSPEC will be searched. There will be 2 independent reviewers in charge of screening titles and abstracts as well as determining those articles that should be included for a full-text review. The third reviewer will act as a mediator between the other 2 reviewers. Moreover, a data extraction form will be created and used during the full article data analysis. Notably, the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) guidelines will be followed in reporting this protocol.

Results: A preliminary search produced 1271 articles, 40 of which are duplicate records. The inclusion-exclusion criteria are being strictly followed in performing the ongoing study selection.

Conclusions: Usability questionnaires are an important tool in mHealth app usability studies. This review will summarize the usability questionnaires used in published research articles while assessing the efficacy of these questionnaires in determining the usability of mHealth apps.

(JMIR Res Protoc 2017;6(8):e151) doi:10.2196/resprot.7826

KEYWORDS

usability; questionnaire; mobile health app; systematic review

Introduction

In recent years, a large number of mobile health (mHealth) apps have been created to augment various personal health regimens including weight loss, smoking cessation, chronic disease management, virtual clinical visits, and medical education. These apps have been evaluated in various usability studies—a critical step in determining the quality of the apps.

According to the International Organization for Standardization, usability marks “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [1]. This definition provides one usability model: effectiveness, efficiency, and satisfaction. Another widely cited usability study model included five alternative components: efficiency, satisfaction, learnability, memorability, and errors [2]. Additionally, there are existing efforts for creating new usability

models specifically for mobile apps by considering new usability challenges (eg, mobility, connectivity, and additional cognitive load) introduced by mobile devices [3]. However, without a thorough evaluation, it is hard to determine which usability model is the best for mHealth app usability studies.

There are many usability study methods. These methods can be categorized into two major types according to the usability study participants. In the first method style, participants are the members of a research/development team or a group of usability study experts. In these roles, they perform cognitive walkthroughs or heuristic evaluations of the mHealth app to determine the app's usability [4]. In the second method style, participants are selected from a pool of the app's actual users. Typically, these participants are required to use the app to finish a number of tasks and then provide their feedback. Here, not only are the participants' performances and activities logged and analyzed, but their opinions of the product are also collected using study questionnaires, focus groups, or interviews. These participants may be encouraged to speak out when working on their tasks to tell researchers their ideas or comments (think aloud).

There are currently a number of validated and reliable usability questionnaires such as the Post-Study System Usability Questionnaire (PSSUQ) [5], System Usability Scale (SUS) [6], Usefulness, Satisfaction, and Ease of Use Questionnaire [7], Telehealth Usability Questionnaire [8], Questionnaire for User Interface Satisfaction [9], Perceived Usefulness and Ease of Use [10], and the Health Information Technology Usability Evaluation Scale [11], to name just a few. These questionnaires (especially SUS and PSSUQ) have been used in a number of mHealth app usability studies because of their previous wide use in IT system usability studies. None of these validated and widely used questionnaires, however, were specifically designed for evaluating the usability of individual mHealth apps, and there is no specific study on the consistency of usability study results from these questionnaires and other usability study methods. Note that there are a number of studies on mHealth app usability evaluation for a group of apps, with the purpose of selecting the best app for some specific tasks such as chronic disease management, headache diary, weight loss, and smoking cessation [12-15]. Because of the significant difference in the research purposes, these studies are substantially different from this study. However, the results obtained in these studies may be useful in this usability questionnaire effectiveness study; therefore, they will be assessed in this study.

mHealth apps have many inherent characteristics limiting their usability. For instance, mHealth apps typically run on mobile devices such as phones and tablets. Mobile phones have a small screen size, tiny fonts, and soft keyboards, which may limit the interactions between the user and the app [3]. For health care purposes, mHealth apps need to be on and available at all times, otherwise users could find themselves in a life-threatening situation. Furthermore, since mHealth apps often need to handle real-time and highly sensitive communications between patients and health care providers, existing usability questionnaires used in the mHealth app usability studies must be sufficiently evaluated so as to determine which ones are the most effective in assessing mHealth app's usability.

There are several review articles [3,16,17] published regarding usability studies in mobile apps. Most of these reviews, however, did not focus specifically on the evaluation of the usability questionnaires used in mHealth app usability studies or they did not compare the result consistency between the usability questionnaires and other usability study methods. Therefore, this systematic review protocol aims at describing the procedure underlying the design of a systematic review for evaluating the effectiveness of usability questionnaires implemented in published mHealth apps. Here, the "effectiveness" indicates that the usability questionnaires can obtain similar results in terms of the app's usability as obtained from other usability study methods. In other words, high effectiveness indicates that the results from the usability questionnaires are highly consistent with the results from other usability study methods on usability aspects they both can measure. The results of this systematic review may lead to the creation of usability questionnaire selection guidelines for researchers seeking to evaluate the usability of mHealth apps using questionnaires.

Methods

The construction of this systematic protocol followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) guidelines [18], which recommend a number of essential components in systematic reviews. Those components most relevant to this review study will be described in detail later.

Search Strategies

The search strategy consists of keywords appropriate to the objective of the review. More specifically, in the initial literature query, our search strategy will simply entail "mobile app" AND "usability." Notably, this initial query contains no keywords relating to health, since there are many different ways of describing health-related mobile apps. Searches will first be performed in bibliographic databases PubMed, CINAHL, INSPEC, ACM Digital Library, and IEEE Xplore digital library. These databases include a huge number of journal articles and conference proceedings. From here, the obtained studies will be selected by following the inclusion and exclusion criteria described below. All reference lists of selected articles will be evaluated, and the studies meeting the inclusion criteria will be added into the review list. Searches on other suitable resources will also be performed if that the search in the five previously listed databases does not produce a sufficient number of studies. Other suitable resources include conference proceedings and reports that are not indexed by the five databases. If necessary, study authors will be contacted if they have mentioned other relevant studies without delineation.

Inclusion and Exclusion Criteria

In general, studies on mobile app usability will be included in the initial results if they were both published in English and in peer-reviewed journals or conferences. More specific inclusion and exclusion criteria related to time frame, participants, types of studies, interventions, and outcomes are described below.

All studies published between January 2008 and March 2017 will be included in the review. It is believed that the database

search will produce only a small number of studies published before 2008. This can be explained by considering the history of mobile phone app development. Prior to 2008, although there were some mobile health apps, they were typically text messaging based apps. The major concentration of those studies was on the content of messages and timing or frequency of delivering those messages. Therefore, the usability of those mobile apps was typically not extensively evaluated or reported.

The participants of usability studies should include targeted users of the designed app in the study instead of relying solely on the research team or usability evaluation experts. The selected studies should include usability evaluation with at least two methods—cognitive walkthrough, think aloud, questionnaires, log and usage file analysis, and heuristic evaluation done by usability experts. One of the methods must be a usability questionnaire answered by a group of targeted app users.

Studies that examined the usability of individual mHealth apps using one or more usability questionnaires will be included in this review. Such studies might include randomized and nonrandomized controlled trials, pre/posttest designs, nonexperimental observation (cross-sectional, case series, case studies), and qualitative studies. Studies will be excluded if they are not about mobile apps, not related to health, not usability studies, or did not use any questionnaires in the usability study. Additionally, studies will be excluded if they are editorials, letters to the editor, interviews, study protocols, reviews, position papers, and opinion papers. Furthermore, if the number of participants in a given usability study is extremely small (<4) or the questionnaire is very brief (<4 questions in the entire questionnaire, including subquestions), the corresponding studies will also be excluded from this review.

There is no restriction on the intervention but the intervention must be delivered by or through the mobile apps, for instance, generating reminders, providing education materials, collecting data from users actively, and assisting with the communications between users and other parties (eg, patients vs health care providers, patients vs caregivers, users and their friends). In other words, the users had interactions with the mHealth apps via the interface of the apps. Studies will be excluded if the described mHealth apps collect data from the users only silently and the users never need to interact with the app.

The outcomes will be collected from those studies in this systematic review that contain usability questionnaires. Note that in this review we will not evaluate the quality of those apps, but rather we will evaluate the consistency between the usability study results from the chosen questionnaires and the results from other usability study methods. Both results should be reported in the selected studies. If a particular study does not yield either of these outcomes, it will be removed from this systematic review.

Study Record Management

All search results will be exported into an EndNote library (Clarivate Analytics). The citations from all searches will be evaluated, and duplicate records will be removed using EndNote. The portable document format (PDF) files of all reviewed articles are to be stored in a shared Box folder (Box, Inc.). Additionally, study selection results, research team discussion notes, and data extraction forms will be stored in this shared Box folder as well.

In certain cases, duplicated publications (with different titles and authors) may be encountered during the study selection. The authors and those studies will be carefully compared and evaluated. If they show no significant difference, the duplicated studies will be removed from the study records. When needed, study authors may be contacted to clarify whether the same study participants, questionnaires, and mobile apps were used in multiple studies.

Study Selection

In the first round of the study selection, 2 reviewers (LZ and JB) will independently select studies by reading the titles and abstracts of all obtained studies and determining their eligibility based on the inclusion/exclusion criteria. A third reviewer (BP) will resolve any disagreements between LZ and JB. All three reviewers will reach an agreement on all selected studies.

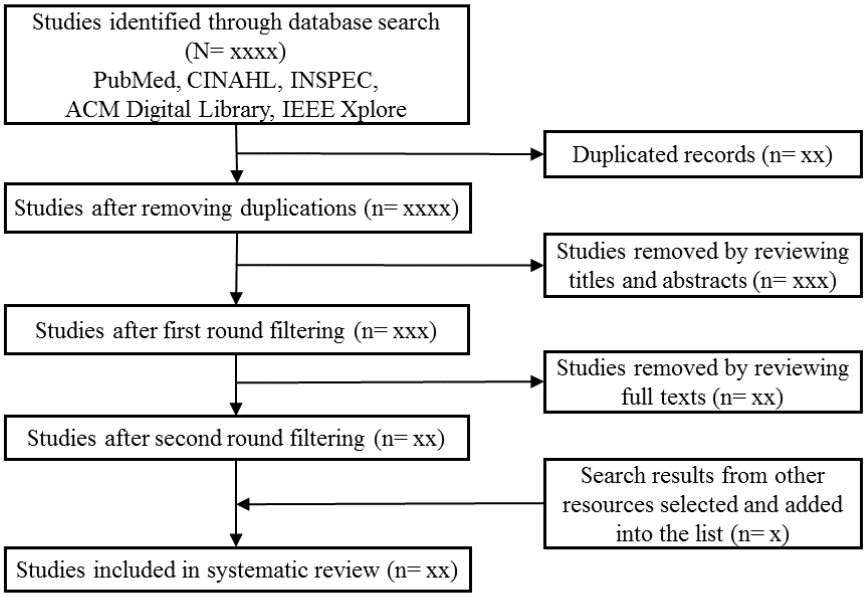
After both reviewers (LZ and JB) finish their study selections, interrater reliability will be measured using Cohen's kappa [19]. To achieve a high kappa score, reviewers should foster extensive discussion regarding the inclusion/exclusion criteria before performing the study selection. Reviewers may also make study selections on a small set of the database search results so as to determine whether or not they apply the inclusion/exclusion criteria in the same way.

In the second round of the study selection, full-text articles of the studies selected in the first round will be downloaded and screened according to the inclusion/exclusion criteria. Each reviewer will work on one third of these studies and the selection results will be combined for further discussion among the 3 reviewers. The 3 reviewers should reach consensus on all selected studies in the second round since these studies will be included in the systematic review. The consensus will be reached by having extensive discussions among 3 reviewers.

Furthermore, the 3 reviewers will screen the articles in the reference lists of the selected studies from the second round using an identical procedure. Some articles in the reference list may be selected and added into the systematic review.

If the number of studies identified is small after the completion of the previous step (<10), other databases and websites may be searched to identify highly relevant studies that are not indexed by the five databases. In most cases, this step is not necessary. The entire study selection procedure is demonstrated in Figure 1.

Figure 1. Flow diagram illustrating the study selection process.



Data Extraction

Each study selected in the previous step will be reviewed and the information about the study will be extracted and documented using a data extraction form by the search team (LZ, JB, and BP). In most cases, the full text of the study and the supplementary materials are sufficient for the purpose of data extraction. Sometimes study authors may need to be contacted if they describe the usability study results but do not provide the usability questionnaire itself.

The data extraction form captures the following data items: paper information, app information, descriptions of usability study methods, study participants, questionnaires, and usability study results. The study participants' characteristics, their performance in the usability study, their answers to questions in the questionnaire, and their comments made during the study will be collected. If needed, further details will be requested from authors of some studies.

After the data extraction has been completed, the quality of studies and publication bias will be evaluated. Publication bias can be determined by using the Egger's test paired with a funnel plot [20]. The quality of the study will be assessed using appropriate tools such as the Cochrane Collaboration's Risk of Bias tool for randomized controlled trials [21,22] and the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system [23,24]. Other observational studies will be assessed using the Newcastle-Ottawa quality assessment scale [25].

Data Synthesis

First, the study results will be narratively summarized and synthesized. This narrative synthesis will include the types of mHealth apps, the summary of each questionnaire, the usability models applied in each study, the specific measured properties (eg, usefulness, learnability, satisfaction), and other usability study methods. Then, the consistency between the results from questionnaire-based usability studies and the ones from other usability study methods will be determined. For instance, in a usability study, study participants are usually asked whether or not the app is easy to use. Moreover, all their activities on the app can be recorded, including individual button clicks, entered data, and finger movement. These two pieces of data can be compared to determine the consistency. Consider the case where one study participant chose "strongly agree" on the ease of use statement in a usability questionnaire but actually took a long time or a lot of efforts to have one simple task done. The consistency between these two usability study methods would not be high. Notably, the comparison and analysis will be qualitative since usability study results are highly subjective and are often presented as opinions or comments. Summary measurements may include descriptive statistics such as frequencies, percentages, measures of central tendency, and variation.

If subgroups are available in the studies—that is, the same type of apps and similar usability studies—more in-depth comparisons may be performed within these subgroups. This comparison result will remove the contribution from mHealth app types and the usability study methods and therefore better reflect the result differences from different usability study methods. Moreover, comparative content analysis may be employed to determine themes across qualitative data by using NVivo software (QSR International).

Results

This study aims to determine the effectiveness of questionnaire-based usability studies in evaluating mHealth app usability as a comparison to other usability study methods

commonly used in mobile app evaluation. Currently, searches in the five selected databases have been performed and 1271 studies have been identified. We have removed 40 duplicated records in these citations. The first round study selection is ongoing. It is expected that the project will be completed in 2017. The results of this study will be used to determine which questionnaire is the most effective in determining the usability of mHealth apps and whether it is necessary to create a new usability questionnaire specifically for mHealth apps.

It is anticipated that some selected articles will have reported only an overall summary of their usability studies, or complementary usability components from different study methods. In this case, authors will be directly contacted to provide more detailed data. If these authors do not respond to the request, these articles shall either be removed from this review study or only used to conduct a comparison for available usability study results.

Discussion

Principal Considerations

Questionnaires have been widely used in mHealth app usability studies. There is no specific guideline, however, on the selection of questionnaires in mHealth app usability studies. Often, researchers simply choose a well-known usability questionnaire such as PSSUQ and SUS [5,6] or decide to create their own questionnaires. Questionnaires such as PSSUQ and SUS were not designed for mHealth apps and many newly created usability questionnaires were for only one specific app or were not validated at all. Even so, it is believed that these questionnaires were still useful in determining the usability of mHealth apps.

This systematic review will collect these studies and perform a systematic evaluation on the effectiveness of each questionnaire in usability studies. This in turn may help researchers to make proper selections when choosing a usability questionnaire for their mHealth studies.

Conclusion

This systematic review protocol is used to demonstrate the transparency of this review study and also provide other investigators with a methodology to conduct systematic reviews.

Acknowledgments

BP and LZ are supported by grants from the National Institute on Disability, Independent Living, and Rehabilitation Research (grants #90RE5018 RERC ICT Access, and #90DP0064 Promoting Self-management and Independence using mHealth).

Authors' Contributions

All authors conceptualized and designed the review protocol, performed the systematic literature searches, selected studies, and extracted and interpreted the data. They were involved in all aspects of drafting, revising, and finalizing this protocol.

Conflicts of Interest

None declared.

References

1. The International Organization for Standardization. Part 11: Guidance on Usability. 1998. ISO 9241, Ergonomic requirements for office work with visual display terminals (VDTs).
2. Nielsen J. Usability engineering. San Francisco, CA: Morgan Kaufmann Publishers; 1993.
3. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1. [doi: 10.1186/2194-0827-1-1]
4. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
5. Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. International Journal of Human-Computer Interaction 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]
6. 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 & Francis; 1996.
7. Lund A. STC Usability SIG Newsletter. 2001. Measuring Usability with the USE Questionnaire URL: https://www.researchgate.net/publication/230786746_Measuring_usability_with_the_USE_questionnaire [accessed 2017-07-28] [WebCite Cache ID 6sIimatm1]
8. Parmanto B, Lewis AN, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehabil 2016;8(1):3-10 [FREE Full text] [doi: 10.5195/ijt.2016.6196] [Medline: 27563386]
9. Chin J, Diehl V, Norman K. Development of an instrument measuring user satisfaction of the human-computer interface. In: Proceedings of CHI'88. New York, NY: ACM; 1988 Presented at: The SIGCHI Conference on Human Factors in Computing Systems; May 15-19; Washington, DC p. 213-218 URL: http://dl.acm.org/citation.cfm?id=57203 [doi: 10.1145/57167.57203]

10. Davis F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
11. Yen P, Wantland D, Bakken S. Development of a Customizable Health IT Usability Evaluation Scale. AMIA Annu Symp Proc 2010 Nov 13:917-921 [FREE Full text] [Medline: 21347112]
12. Anderson K, Burford O, Emmerton L. App Chronic Disease Checklist: Protocol to Evaluate Mobile Apps for Chronic Disease Self-Management. JMIR Res Protoc 2016 Nov 04;5(4):e204 [FREE Full text] [doi: 10.2196/resprot.6194] [Medline: 27815233]
13. Hundert A, Huguet A, McGrath P, Stinson J, Wheaton M. Commercially available mobile phone headache diary apps: a systematic review. JMIR Mhealth Uhealth 2014 Aug 19;2(3):e36 [FREE Full text] [doi: 10.2196/mhealth.3452] [Medline: 25138438]
14. Patel R, Sulzberger L, Li G, Mair J, Morley H, Shing MN, et al. Smartphone apps for weight loss and smoking cessation: Quality ranking of 120 apps. N Z Med J 2015 Sep 04;128(1421):73-76. [Medline: 26370762]
15. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
16. Coursaris C, Kim D. A Meta-Analytical Review of Empirical Mobile Usability Studies. Journal of Usability Studies 2011;6(3):117-171 [FREE Full text]
17. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
18. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
19. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
20. Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [Medline: 9310563]
21. 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:d5928 [FREE Full text] [Medline: 22008217]
22. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions. Chichester, England: Wiley-Blackwell; 2008.
23. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011 Apr;64(4):401-406. [doi: 10.1016/j.jclinepi.2010.07.015] [Medline: 21208779]
24. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008 Apr 26;336(7650):924-926 [FREE Full text] [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]
25. Wells G, Shea B, O'Connell D. The Newcastel-Ottawa Scale (NOS) for assessing the qualtiy of nonrandomised studies in meta-analysis. 2014. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2017-07-27] [WebCite Cache ID 6sH6IrJMb]

Abbreviations

mHealth: mobile health
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
PSSUQ: Post-Study System Usability Questionnaire
SUS: System Usability Scale

Edited by G Eysenbach; submitted 05.04.17; peer-reviewed by M Ashford, K Anderson; comments to author 29.06.17; revised version received 11.07.17; accepted 14.07.17; published 01.08.17.

Please cite as:
Zhou L, Bao J, Parmanto B
Systematic Review Protocol to Assess the Effectiveness of Usability Questionnaires in mHealth App Studies
JMIR Res Protoc 2017;6(8):e151
URL: http://www.researchprotocols.org/2017/8/e151/
doi:10.2196/resprot.7826
PMID:28765101

Protocol

mHealth Technologies for Palliative Care Patients at the Interface of In-Patient to Outpatient Care: Protocol of Feasibility Study Aiming to Early Predict Deterioration of Patient's Health Status

Gudrun Theile[1], MPH, MD; Vanessa Klaas[2]; Gerhard Tröster[2], PhD; Matthias Guckenberger[1], PhD, MD

[1]Clinic of Radiation-Oncology, Competence Center Palliative Care, University Hospital Zurich, Zurich, Switzerland

[2]Wearable Computing Laboratory, Swiss Federal Institute of Technology, Zurich, Switzerland

Corresponding Author:
Gudrun Theile, MPH, MD
Clinic of Radiation-Oncology
Competence Center Palliative Care
University Hospital Zurich
100 Raemistrasse
Zurich, 8091
Switzerland
Phone: 41 44 2555495
Fax: 41 44 2554127
Email: gudrun.theile@usz.ch

Abstract

Background: Palliative care patients are a particularly vulnerable population and one of the critical phases in patients' trajectories is discharge from specialized in-patient palliative care into outpatient care, where availability of a palliative care infrastructure is highly variable. A relevant number of potentially avoidable readmissions and emergency visits of palliative patients is observed due to rapid exacerbation of symptoms indicating the need for a closer patient monitoring. In the last years, different mHealth technology applications have been evaluated in many different patient groups.

Objective: The aim of our study is to test feasibility of a remote physical and social tracking system in palliative care patients.

Methods: A feasibility study with explorative, descriptive study design, comprised of 3 work packages. From the wards of the Clinic of Radiation-Oncology at the University Hospital Zurich, including the specialized palliative care ward, 30 patients will be recruited and will receive a mobile phone and a tracking bracelet before discharge. The aim of work package A is to evaluate if severely ill patients accept to be equipped with a tracking bracelet and a mobile phone (by semiquantitative questionnaires and guideline interviews). Work package B evaluates the technical feasibility and quality of the acquired electronic health data. Work package C will demonstrate whether physical activity parameters, such as step count, sleep duration, social activity patterns like making calls, and vital signs (eg, heart rate) do correlate with subjective health data and can serve as indicator to early detect and predict changes in patients' health status. Activity parameters will be extracted from the mobile phone's and wristband's sensor data using signal processing methods. Subjective health data is captured via electronic version of visual analog scale and Distress Thermometer as well as the European Organization for Research and Treatment of Cancer – Quality of Life Questionnaire C30 in paper version.

Results: Enrollment began in February 2017. First study results will be reported in the middle of 2018.

Conclusions: Our project will deliver relevant data on patients' acceptance of activity and social tracking and test the correlation between subjective symptom assessment and objective activity in the vulnerable population of palliative care patients. The proposed study is meant to be preparatory work for an intervention study to test the effect of wireless monitoring of palliative care patients on symptom control and quality of life.

(JMIR Res Protoc 2017;6(8):e142) doi:10.2196/resprot.7676

KEYWORDS

mobile apps; palliative care; pain; symptom assessment; hospitalization; aged

Introduction

Severely ill patients and their caregivers face many challenges. Suffering from a life-threatening disease not only confronts people with deeply existential fears that demand complex coping strategies. It also means to be subdued to numerous diagnostic and therapeutic interventions, which sometimes themselves include relevant risks of side effects and complications. In the last months of their lives many palliative care patients are afflicted with different significant symptoms, such as pain, nausea, dyspnea, fatigue, and sleeplessness, just to name the most frequently occurring complaints [1-3].

One of the particular critical phases in patients' trajectories is discharge from hospital care. After structured and multiprofessional care within the hospital environment, maybe even at a specialized Palliative Care Unit, the quality of care at home depends on many local factors that are difficult or impossible to be changed: patients` environment, including housing, social, and family situation, as well as availability and training of local professional health services. Due to different initiatives on the federal and cantonal level, the availability of ambulant specialized palliative care services in Switzerland is on the rise. Still, most patients discharged from hospital are lost to follow-ups by a palliative care specialist. As a consequence, many patients present for unplanned readmissions at hospitals and emergency departments [4]. Depending on different settings, between 17% and 50% of these (re)admissions and emergency visits (EV) of cancer patients are deemed avoidable [5-7]. Men, lung cancer patients, patients with low continuity of care, and those not attending (ambulant) palliative care have been shown to be particularly vulnerable [8,9]. Main causes for EV in cancer patients are uncontrolled symptoms, predominantly pain, nausea, vomiting, constipation, and dyspnea followed by complications associated with cancer treatment [10-12]. Currently, disease trajectories leading to EV are not well understood and this is not only true for cancer patients, but also applying to patients with life-limiting chronic diseases (eg, heart failure); continuity of care seems to also be one of the key factors in both disease spectra [13,14]. Thus, better coordination and communication between services, establishing transitional palliative care (based in hospital, collaborating with primary care physicians) are repeatedly named as potentially effective tools to diminish unplanned readmissions and EVs [2,4,9,15].

Due to a dramatic development of mHealth the use of apps and texting for patients' support, but also wireless tracking of physical activity using mobile phones and wearable sensors has been expeditiously adopted in western societies in the last few years. In 2015, 63% of the US adult population have owned a mobile phone device and even in the age group >65-years old, mobile phones were owned by 27% [16]. A study conducted in patients from a department of Internal Medicine in Los Angeles in 2014 to 2015 showed that 91% of those patients owned a mobile phone, with 76% of these reporting having Internet capability. Mobile apps were used by 75% of patients and 32% of these even used health apps [17]. The use of mobile phones and mobile sensor devices in medical research is also growing fast: almost US $1 billion in grant money was invested in mHealth research in 2013 resulting in approximately 1000 web-of-science publications in 2013 [18]. Reasons for this tremendous hope on mHealth are their possibilities to obtain various objective patient-related parameters in a noninvasive, passive, continuous, and real-time fashion.

Meanwhile mHealth has been explored in many diseases and patient groups, for example, in cardiovascular diseases [19], mental disorders [20,21], obese people [22], hemodialysis patients [23], chronic pain [24], children with cerebral palsy [25], pulmonary rehabilitation [26], Parkinson's disease [27], and stressed persons [28]. Also, the collection of patient subjective reported outcome using mHealth is already established in oncology and yet has been proven feasible in palliative care patients of different age groups [29-33]. In a study with 162 cancer patients in different disease stages (early to advanced) plus 20 healthy controls, Ferrioli et al. [34] could demonstrate a strong correlation between daily physical activity monitored by an accelerometer sensor and disease stage, functional status, fatigue, and quality of life (QoL) in cancer patients [34]. Patients were equipped with a mobile sensor that was attached to the anterior thigh using adhesive dressings (thus, being much more uncomfortable than current bracelet devices); nevertheless, 98% of patients judged the sensor as acceptable and wore it 7 days as proposed. Arguments based on the assumption the population of elderly people could not adapt to the technical progress have been successfully refuted by several studies [35-37]. Consequently, the international scientific palliative care community promotes the use of mHealth for quite a while, but up to now there has not been much experience gained and published [38].

The overall aim of our feasibility study is to advance transitional care and continuity of care of seriously ill patients by application of mHealth technologies. We assume that by monitoring patients` activity using mHealth technology, we will be able to predict a decline of the patients` health status in time: this could then trigger intensified care by palliative care professionals. The future goal of our research is a clinical interventional trial, which aims at reducing unplanned hospital readmissions and EV of ambulant palliative care patients by the use of mHealth. The feasibility study presented in this paper will evaluate and optimize patients' acceptance of mHealth technologies in the palliative care setting of patients discharged from in-patient care. We aim to analyze the correlation between patient's behavioral patterns and changes in crucial symptoms as well as elaborate algorithms for trend detection.

Methods

Aims of the Study

The first aim of this study is to evaluate palliative care patients' acceptance with regard to the supply with a commercially available wireless physical activity tracking bracelet and a mobile phone in order to monitor objective health and behavioral data and to capture subjective symptom assessment. The second aim is the evaluation of correlations between the patient-specific activity patterns (physical, social activity, and vital signs) and the subjective patients' ratings of pain, distress, and QoL in order to early detect changes in the patient-specific pattern of

physical and social behavior, vital sign patterns, and the named symptom ratings.

Characteristics of Participants

Patients will be consecutively recruited at the 2 wards of the clinic of radiation-oncology of the University Hospital Zurich (USZ). One ward is part of the Competence Center Palliative Care USZ and offers specialized palliative care for patients with life-threatening diseases from all clinics of the university hospital. Most of these patients do suffer from cancer, a minority from incurable heart failure, severe pulmonary affections, or neurologic diseases. Inclusion criteria are: patients with established diagnosis of metastatic cancer or other severe illness with limited life-expectancy (physicians guess <12 months, >8 weeks), Karnofsky Index ≥50%, scale of performance status developed by the Eastern Cooperative Oncology Group (ECOG) ≤2, aged >18 years. Eligible patients with interest of participation will have to pass a practical minitest concerning the handling of the devices (ie, simple operations on a mobile phone and handling of activity bracelet). Exclusion criteria are relevant cognitive impairment and insufficient knowledge of German language. Similar to related work [20,39] a number of 30 participants is intended to ensure relevant findings by capturing a representative cohort regarding sex, age, health status (Karnofsky Index/ECOG), life-limiting diseases, and different grades of pre-existing experience in the use of mobile phones or other electronic devices. Patients interested in the study will receive a leaflet with sound information on the aim of the study, a rough sketch of study procedure, and requirements with regard to participants (eg, wearing a tracking bracelet). Attached at this information leaflet is the informed consent. Moreover, we hand out a second leaflet with detailed information on the handling of the tracking bracelet, the app, and characteristics of data gathered by device and app.

Study Design

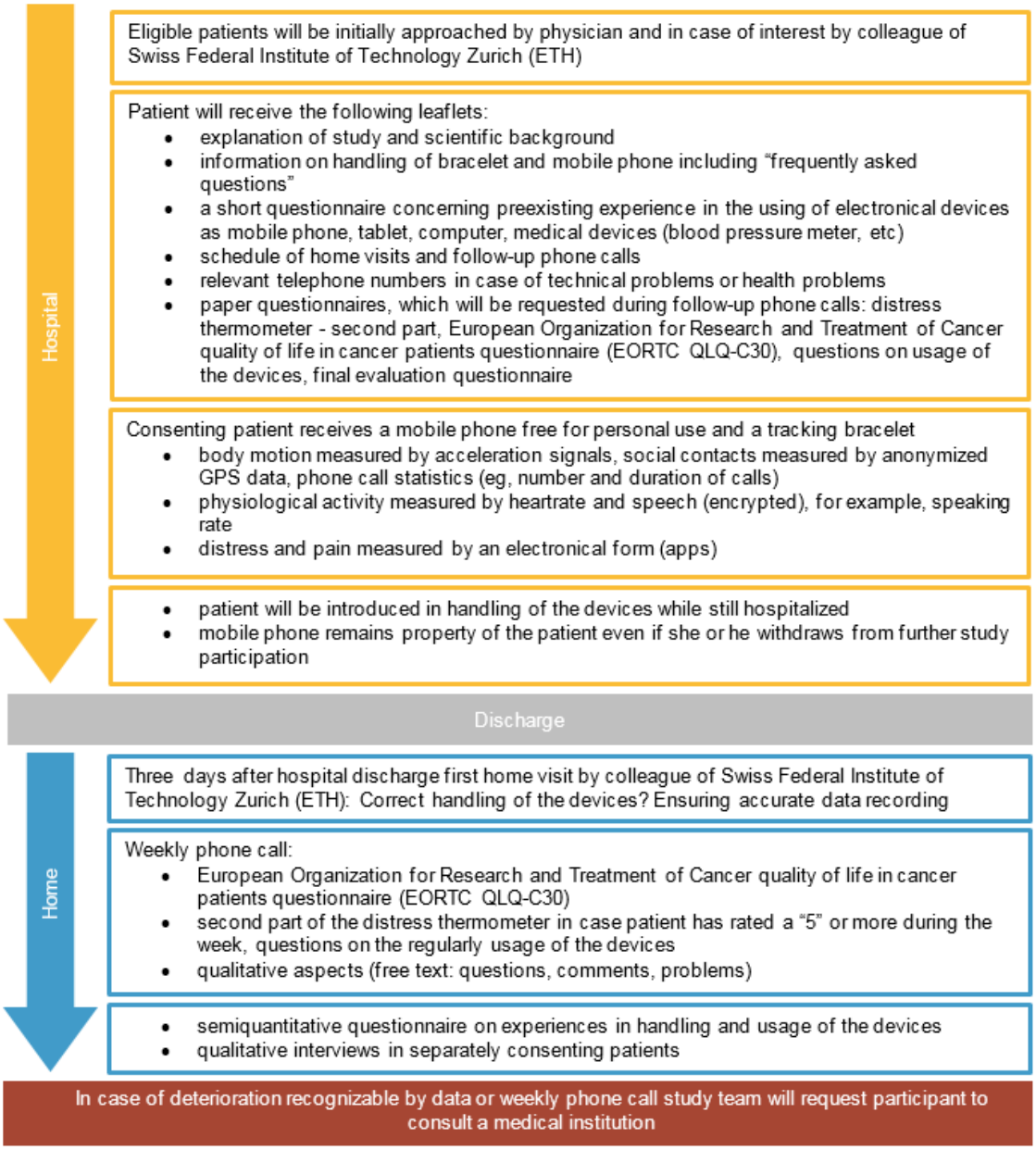
For this feasibility study, we chose an explorative descriptive design observing palliative care patients discharged from hospital and equipped with physical activity tracker and mobile phone technology. The underlying assumption is that wireless tracking of physical and social activity as well as of vital sign data in the palliative care setting is accepted by patients, is feasible, and its pattern analysis allows the generation of objective information on the current health status in the vulnerable phase of leaving inpatient care. For specific study question, see Textbox 1.

Patients leaving inpatient care of the Clinic of Radiation-Oncology will be equipped with a mobile phone and tracking bracelets. We will use the Samsung Galaxy S5 mobile phone and it will capture body motion data (by acceleration signals), location (anonymized global positioning system [GPS] data), encrypted speech, and phone call statistics. Furthermore, in this study, a mobile phone questionnaire app will be installed, a visual pain scale, and the Distress Thermometer. Patients will be requested to answer these simple to handle questionnaire app twice a day. The tracking bracelet used in this feasibility study is provided by Biovotion company, it is worn at the upper arm, easy to put on, and it will capture heart rate, heart rate variability, blood oxygen, blood perfusion, skin temperature, stress (by galvanic skin response), and body motion data (by acceleration data). Figure 1 gives a detailed overview on study procedures.

Textbox 1. Specific questions of the feasibility study.

- Are severely ill patients willing and able to join the study?
- Are severely ill and often elderly patients able to handle mHealth devices and mobile apps presenting visual scales?
- How can we support correct handling of the electronic devices to achieve good data quality?
- What is the patient selection criteria for the use of mHealth?
- Feasibility, acceptance, and technical functionality in the home care setting.
- Patient acceptance and feasibility over a longer follow-up period (maximum 12 months).
- How accurate is the data collected and transferred by the devices to the study center over the follow-up period?
- How accurate is data of physical and social activity trackers correlated with subjective reported outcomes by patients?
- Is it possible to detect behavioral changes, especially within person, with gathered data?
- Is it possible to extend a patient-specific model with a general model?

Figure 1. Study flow and explanation of different procedures.



Textbox 2. Work package A: patient acceptance of mHealth.

- Quantitative (descriptive statistics) and qualitative evaluation of usage, acceptance, and potential problems during follow-up (maximum 12 weeks) within home visits and/or weekly telephone interviews.
- Definition of criteria for successful application (patients' selection, instructive approaches).
- Final evaluation of usage and potential problems by semiquantitative questionnaire at the end of follow-up, guideline interviews of some of the patients, and of the involved personnel.
- Descriptive evaluation of problems indicated in the second part of the Distress Thermometer (if applicable).
- As nonresponders will be defined those patients, who are approached to participate but who decline from participation. We will capture demographic date of nonresponders, including main diagnosis if these patients gave general informed consent to the University Hospital for usage of their data from electronic records for research purposes. Additionally, we will ask nonresponders to document their reason for decline in a short paper form; they have to consent to this procedure by giving as a signature. As drop-outs will be defined those patients, who gave consent for participation, but drop out during follow-up for other reason than death.

Planned Data Collection and Analysis

With regard to analysis, the study is composed of 3 work packages: A, B, and C. Work packages A and B enclose interims analysis; that way preliminary evaluation results can be implemented during patient recruitment phase aiming continuously to improve feasibility and data quality. In Textboxes 2-4, we describe the different work packages.

Work Package B – Technological Feasibility of Behavior Tracking and Monitoring of Health Status

Work package B evaluates whether physical activity parameters, such as step count, sleep duration, social activity patterns like making calls, going out to meet people, and vital signs (eg, heart rate) can be technically measured over sufficiently long periods and whether this data correlate with changes in the health status of patients. Health status will be assessed in terms of the following parameters in Textbox 3.

Work Package C - Data Analysis, Correlation, and Trend Detection

The recorded data of the bracelet and mobile phone are uploaded to a secured webserver of the Swiss Federal Institute of Technology once per day to allow remote access to the data (only for the conductors of the study) in order to ensure accurate recording and for evaluation purposes. We chose the Biovotion bracelet, because it gives us the possibility to avoid data storage in a cloud. Additionally, it is an officially recognized medical device. Privacy sensitive data as real location (GPS) is anonymized, speech is encrypted, names and phone numbers of contacts are not recorded. All patient sensitive data is encoded and accessed only by people determined by the principal investigator. The local Ethics Committee (Kantonale Ethikkommission Zürich) has already approved the study and agreed to detailed specifications on data security we gave to them. Textbox 4 gives an overview on planned data collection and analysis.

Textbox 3. Work package B: subjective parameters.

- Pain measured by an electronic form of a visual analogue scale (rating scale 0-10)
- Distress measured by an electronic form of the National Comprehensive Cancer Network Distress Thermometer [40] part 1 (rating scale 0-10); part 2 will be administered in weekly phone calls if ratings in the distress app were 5 or higher during the previous week.
- Quality of life: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30 [41,42].

Textbox 4. Planned data analysis.

- Evaluation of the correlation (eg, the Pearson coefficient, statistical tests) between the patient-specific activity patterns (physical, social activity, and vital signs) and the ratings of the National Comprehensive Cancer Network Distress Thermometer, the pain visual analogue scale, and the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30.
- Quantitative ranking of the activity parameters regarding their relevance to describe the patient's pain and distress status using statistics and machine learning (eg, minimum redundancy maximum relevance).
- Development of models to detect the patient's current pain and distress level based on analysis of behavioral patterns and trend detection as well as knowledge-driven approaches.
- Evaluate the potential of the proposed activity tracking system to predict the health status of patients by assessing the accuracy of the prediction model using the patient questionnaire data as reference.
- Record the interventions needed based on the early detection of activity changes seen, including (emergency) hospital readmission.
- Application of adaptively train semisupervised classifiers, such as support vector machines, hidden Markov models, and clustering methods for trend detection and detection of behavioral changes, especially within person.
- Application of knowledge-driven approaches in order to compensate missing data to generate general model out of patient-specific model.

Results

Enrollment began in February 2017. Data collection will be completed in January 2018, first results will be reported in the middle of the year.

Discussion

Continuous symptom assessment and critical review of symptoms by patient and doctor have been defined as a mainstay of effective ambulant palliative care [39]. Yet, in many European countries primary care practices are rarified [43], reimbursement of home visits are insufficient [44], and despite different movements to embed palliative care in the public health care system ambulatory palliative care institutions are still lacking. Our study aims to detect early health status changes with the help of mHealth technology in order to improve health care for ambulatory palliative care patients and avoid unplanned or emergency hospital readmissions. As to our knowledge with regard to this special patient population, no data on correlation of subjective symptom burden and objective activity parameters do exist until now, this project will deliver highly novel data. Not only on aspects of feasibility and patients` acceptance, but particularly on the proof (or the rejection) of the presumed existence of a correlation between subjective and objective data. Due to the fact, we collect a broad range of activity data reaching form blood oxygen, perfusion, skin temperature and heart beat to call statistics, and voice modulation, we will be able to identify parameters that correlate best with the patient's subjective symptom burden. This is important for the future

development of effective mHealth systems supporting remote symptom management by wearables combined with proactive care in a patient group that is in constant danger of a rapid and/or unexpected deterioration in health status. With the help of advanced data analysis algorithm we will be able not only to detect subtle changes in health status sometimes initially unnoticed by the patient and his/her family, but relevant precursor of possible symptom exacerbations. This approach will furthermore allow for even predicting these sometimes very finely graduated changes in time; thus, enabling professionals with data access to initiate effective intervention in time. We are aware of the fact that equipment of palliative patients with tracking devices requires sensitivity and the goal of palliative care remains to provide the best QoL possible, not only with respect to medical, but also to psychosocial and spiritual needs. This project therefore aims to supplement and assist existing palliative care structures and resources to build-up effective home care programs, which provide comprehensive and coordinated care closely adapted to the patient individual needs.

Acknowledgments

This study is funded by the “Research in Palliative Care” program of the Swiss Academy of Medical Sciences (SAMW) and the Gottfried and Julia Bangerter-Rhyner Foundation, Bern, Switzerland.

The local Ethics Committee (Kantonale Ethikkommission Zürich) has approved the study protocol; approval number PB_2016-00895.

Authors' Contributions

GTH conceived the study design and drafted the study protocol. MG, GTR, and VK made substantial contributions to conception and design, GTR and VK especially with regard to all technical aspects as selection of electronic devices and data analysis. MG, GTR, and VK revised the manuscript critically. All authors gave final approval of the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Funding report.

[PDF File (Adobe PDF File), 115KB - resprot_v6i8e142_app1.pdf]

Multimedia Appendix 2

CONSORT eHealth Form.

[PDF File (Adobe PDF File), 670KB - resprot_v6i8e142_app2.pdf]

References

1. Chang VT, Hwang SS, Feuerman M, Kasimis BS. Symptom and quality of life survey of medical oncology patients at a veterans affairs medical center: a role for symptom assessment. Cancer 2000;88:1175-1183. [Medline: 10699909]
2. Peters L, Sellick K. Quality of life of cancer patients receiving inpatient and home-based palliative care. J Adv Nurs 2006;53:524-533. [doi: 10.1111/j.1365-2648.2006.03754.x] [Medline: 16499673]
3. LeBlanc TW, Nickolich M, Rushing CN, Samsa GP, Locke SC, Abernethy AP. What bothers lung cancer patients the most? A prospective, longitudinal electronic patient-reported outcomes study in advanced non-small cell lung cancer. Support Care Cancer 2015;23:3455-3463. [doi: 10.1007/s00520-015-2699-4] [Medline: 25791391]
4. Gibson S, McConigley R. Unplanned oncology admissions within 14 days of non-surgical discharge: a retrospective study. Support Care Cancer 2016;24:311-317. [doi: 10.1007/s00520-015-2786-6] [Medline: 26062923]
5. Wallace EM, Cooney MC, Walsh J, Conroy M, Twomey F. Why do palliative care patients present to the emergency department? Avoidable or unavoidable? Am J Hosp Palliat Care 2013;30:253-256. [doi: 10.1177/1049909112447285] [Medline: 22628898]
6. Alsirafy SA, Raheem AA, Al-Zahrani AS, Mohammed AA, Sherisher MA, El-Kashif AT, et al. Emergency department visits at the end of life of patients with terminal cancer: pattern, causes, and avoidability. Am J Hosp Palliat Care 2016;33:658-662. [doi: 10.1177/1049909115581819] [Medline: 25877944]
7. Delgado-Guay MO, Kim YJ, Shin SH, Chisholm G, Williams J, Allo J, et al. Avoidable and unavoidable visits to the emergency department among patients with advanced cancer receiving outpatient palliative care. J Pain Symptom Manage 2015;49:497-504. [doi: 10.1016/j.jpainsymman.2014.07.007] [Medline: 25131891]
8. Burge F, Lawson B, Johnston G. Family physician continuity of care and emergency department use in end-of-life cancer care. Med Care 2003;41:992-1001. [doi: 10.1097/01.MLR.0000078155.78366.5B] [Medline: 12886178]

9. Henson L, Gao W, Higginson I, Smith M, Davies J, Ellis-Smith C, et al. Emergency department attendance by patients with cancer in the last month of life: a systematic review and meta-analysis. Lancet 2015;385(Suppl 1):S41. [doi: 10.1016/S0140-6736(15)60356-7] [Medline: 26312863]
10. Delgado-Guay MO, Rodriguez-Nunez A, Shin SH, Chisholm G, Williams J, Frisbee-Hume S, et al. Characteristics and outcomes of patients with advanced cancer evaluated by a palliative care team at an emergency center. A retrospective study. Support Care Cancer 2016;24:2287-2295. [doi: 10.1007/s00520-015-3034-9] [Medline: 26590844]
11. Tsai S, Liu L, Tang S, Chen J, Chen M. Cancer pain as the presenting problem in emergency departments: incidence and related factors. Support Care Cancer 2010;18:57-65. [doi: 10.1007/s00520-009-0630-6] [Medline: 19350283]
12. Elsayem AF, Elzubeir HE, Brock PA, Todd KH. Integrating palliative care in oncologic emergency departments: challenges and opportunities. World J Clin Oncol 2016;7:227-233 [FREE Full text] [doi: 10.5306/wjco.v7.i2.227] [Medline: 27081645]
13. Brown J, Grudzen C, Kyriacou DN, Obermeyer Z, Quest T, Rivera D, et al. The emergency care of patients with cancer: setting the research agenda. Ann Emerg Med 2016;68:706-711 [FREE Full text] [doi: 10.1016/j.annemergmed.2016.01.021] [Medline: 26921969]
14. Vogt V, Koller D, Sundmacher L. Continuity of care in the ambulatory sector and hospital admissions among patients with heart failure in Germany. Eur J Public Health 2016. [doi: 10.1093/eurpub/ckw018] [Medline: 26962039]
15. Morrison J, Palumbo MV, Rambur B. Reducing preventable hospitalizations with two models of transitional care. J Nurs Scholarsh 2016;48:322-329. [doi: 10.1111/jnu.12210] [Medline: 27074394]
16. Smtih A. Chapter one: a portrait of smartphone ownership. Pew Research Center 2017 Apr 23 [FREE Full text]
17. Ramirez V, Johnson E, Gonzalez C, Ramirez V, Rubino B, Rossetti G. Assessing the use of mobile health technology by patients: an observational study in primary care clinics. JMIR Mhealth Uhealth 2016;4:e41 [FREE Full text] [doi: 10.2196/mhealth.4928] [Medline: 27095507]
18. Steinhubl S, Muse E, Topol E. The emerging field of mobile health. Sci Transl Med 2015;7:283rv3 [FREE Full text] [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
19. Urrea B, Misra S, Plante TB, Kelli HM, Misra S, Blaha MJ, et al. Mobile health initiatives to improve outcomes in primary prevention of cardiovascular disease. Curr Treat Options Cardiovasc Med 2015;17:59. [doi: 10.1007/s11936-015-0417-7] [Medline: 26474892]
20. Grünerbl A, Muaremi A, Osmani V, Bahle G, Ohler S, Tröster G, et al. Smartphone-based recognition of states and state changes in bipolar disorder patients. IEEE J Biomed Health Inform 2015;19:140-148. [doi: 10.1109/JBHI.2014.2343154] [Medline: 25073181]
21. Faurholt-Jepsen M, Vinberg M, Christensen EM, Frost M, Bardram J, Kessing LV. Daily electronic self-monitoring of subjective and objective symptoms in bipolar disorder--the MONARCA trial protocol (MONitoring, treAtment and pRediCtion of bipolAr disorder episodes): a randomised controlled single-blind trial. BMJ Open 2013;3 [FREE Full text] [doi: 10.1136/bmjopen-2013-003353] [Medline: 23883891]
22. Taylor D, Murphy J, Ahmad M, Purkayastha S, Scholtz S, Ramezani R, et al. Quantified-self for obesity: physical activity behaviour sensing to improve health outcomes. Stud Health Technol Inform 2016;220:414-416. [Medline: 27046615]
23. Marteau TM, Dormandy E. Facilitating informed choice in prenatal testing: how well are we doing? Am J Med Genet 2001;106:185-190. [doi: 10.1002/ajmg.10006] [Medline: 11778978]
24. Seiter J, Macrea L, Amft O, Feese S, Arnrich B, Tröster G. Proceedings of the 8th, Evaluating daily life activity using smartphones as novel outcome measure for surgical pain therapy, International Conference on Body Area Networks, 2013. In: https://pdfs.semanticscholar.org/8ac8/cd4442bd624de39128b6f1e2273bf3076003.pdf. 2013 Presented at: 8th International Conference on Body Area Networks (BODYNETS 2013); 30 September – 2 October 2013; Boston URL: https://pdfs.semanticscholar.org/8ac8/cd4442bd624de39128b6f1e2273bf3076003.pdf
25. Seiter J, Derungs A, Schuster-Amft C, Amft O, Tröster G. Daily life activity routine discovery in hemiparetic rehabilitation patients using topic models. Methods Inf Med 2015;54:248-255. [doi: 10.3414/ME14-01-0082] [Medline: 25658903]
26. Rizk AK, Wardini R, Chan-Thim E, Trutschnigg B, Forget A, Pepin V. Using continuous data tracking technology to study exercise adherence in pulmonary rehabilitation. J Vis Exp 2013:e50643 [FREE Full text] [doi: 10.3791/50643] [Medline: 24300076]
27. Mazilu S, Blanke U, Calatroni A, Gazit E, Hausdorff J, Tröster G. The role of wrist-mounted inertial sensors in detecting gait freeze episodes in Parkinson?s Disease. Pervasive Mob Comput 2016:33-16 [FREE Full text] [doi: 10.1016/j.pmcj.2015]
28. Muaremi A, Arnrich B, Tröster G. Towards measuring stress with smartphones and wearable devices during workday and sleep. Bionanoscience 2013;3:172-183 [FREE Full text] [doi: 10.1007/s12668-013-0089-2] [Medline: 25530929]
29. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. CA Cancer J Clin 2012;62:337-347 [FREE Full text] [doi: 10.3322/caac.21150] [Medline: 22811342]
30. Brunelli C, Kaasa S, Knudsen AK, Hjermstad MJ, Pigni A, Caraceni A. Comparisons of patient and physician assessment of pain-related domains in cancer pain classification: results from a large international multicenter study. J Pain 2014;15:59-67. [doi: 10.1016/j.jpain.2013.09.011] [Medline: 24373570]
31. Lundström S. Is it reliable to use cellular phones for symptom assessment in palliative care? Report on a study in patients with advanced cancer. J Palliat Med 2009;12:1087. [doi: 10.1089/jpm.2009.0207] [Medline: 19995287]

32. Blum D, Koeberle D, Omlin A, Walker J, Von MR, Mingrone W, et al. Feasibility and acceptance of electronic monitoring of symptoms and syndromes using a handheld computer in patients with advanced cancer in daily oncology practice. Support Care Cancer 2014;22:2425-2434. [doi: 10.1007/s00520-014-2201-8] [Medline: 24705855]
33. Egbring M, Far E, Roos M, Dietrich M, Brauchbar M, Kullak-Ublick GA, et al. A mobile app to stabilize daily functional activity of breast cancer patients in collaboration with the physician: a randomized controlled clinical trial. J Med Internet Res 2016;18:e238 [FREE Full text] [doi: 10.2196/jmir.6414] [Medline: 27601354]
34. Ferriolli E, Skipworth RJ, Hendry P, Scott A, Stensteth J, Dahele M, et al. Physical activity monitoring: a responsive and meaningful patient-centered outcome for surgery, chemotherapy, or radiotherapy? J Pain Symptom Manage 2012;43:1025-1035. [doi: 10.1016/j.jpainsymman.2011.06.013] [Medline: 22269181]
35. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older adults' experiences using a commercially available monitor to self-track their physical activity. JMIR Mhealth Uhealth 2016;4:e35 [FREE Full text] [doi: 10.2196/mhealth.5120] [Medline: 27076486]
36. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older adults' experiences using a commercially available monitor to self-track their physical activity. JMIR Mhealth Uhealth 2016;4:e35 [FREE Full text] [doi: 10.2196/mhealth.5120] [Medline: 27076486]
37. Rumer KK, Saraswathula A, Melcher ML. Prehabilitation in our most frail surgical patients: are wearable fitness devices the next frontier? Curr Opin Organ Transplant 2016;21:188-193. [doi: 10.1097/MOT.0000000000000295] [Medline: 26859220]
38. Kaasa S, Loge JH, Fayers P, Caraceni A, Strasser F, Hjermstad MJ, et al. Symptom assessment in palliative care: a need for international collaboration. J Clin Oncol 2008;26:3867-3873. [doi: 10.1200/JCO.2007.15.8881] [Medline: 18688054]
39. McCall K, Keen J, Farrer K, Maguire R, McCann L, Johnston B, et al. Perceptions of the use of a remote monitoring system in patients receiving palliative care at home. Int J Palliat Nurs 2008;14:426-431. [doi: 10.12968/ijpn.2008.14.9.31121] [Medline: 19060793]
40. Mehnert A, Müller D, Lehmann C, Koch U. Die deutsche version des NCCN distress-thermometers. Empirische Prüfung eines Screening-Instruments zur Erfassung psychosozialer Belastung von Krebspatienten. Z Psychiatr Psych Ps 2006;2006:213. [doi: 10.1024/1661-4747.54.3.213]
41. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, de Haes J C. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85:365-376. [Medline: 8433390]
42. Kaasa S, Bjordal K, Aaronson N, Moum T, Wist E, Hagen S, et al. The EORTC core quality of life questionnaire (QLQ-C30): validity and reliability when analysed with patients treated with palliative radiotherapy. Eur J Cancer 1995;31A:2260-2263. [Medline: 8652253]
43. Cerny T, Rosemann T, Tandjung R, Chmiel C. [Reasons for general practitioner shortage – a comparison between France and Switzerland]. Praxis (Bern 1994) 2016;105:619-636. [doi: 10.1024/1661-8157/a002362] [Medline: 27223415]
44. Theile G, Kruschinski C, Buck M, Müller CA, Hummers-Pradier E. Home visits - central to primary care, tradition or an obligation? A qualitative study. BMC Fam Pract 2011;12:24 [FREE Full text] [doi: 10.1186/1471-2296-12-24] [Medline: 21513534]

Abbreviations

ECOG: scale of performance status developed by the Eastern Cooperative Oncology Group
EV: emergency visits
GPS: global positioning system
QoL: quality of life
USZ: University Hospital Zurich

Edited by A McDougall; submitted 10.03.17; peer-reviewed by N Bradford; comments to author 04.04.17; revised version received 25.04.17; accepted 05.05.17; published 16.08.17.

Please cite as:
Theile G, Klaas V, Tröster G, Guckenberger M
mHealth Technologies for Palliative Care Patients at the Interface of In-Patient to Outpatient Care: Protocol of Feasibility Study Aiming to Early Predict Deterioration of Patient's Health Status
JMIR Res Protoc 2017;6(8):e142
URL: http://www.researchprotocols.org/2017/8/e142/
doi:10.2196/resprot.7676
PMID:28814378

Proposal

Automating Construction of Machine Learning Models With Clinical Big Data: Proposal Rationale and Methods

Gang Luo[1], PhD; Bryan L Stone[2], MD, MS; Michael D Johnson[2], MD; Peter Tarczy-Hornoch[1,3,4], MD; Adam B Wilcox[1], PhD; Sean D Mooney[1], PhD; Xiaoming Sheng[2], PhD; Peter J Haug[5,6], MD; Flory L Nkoy[2], MD, MS, MPH

[1]Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, United States

[2]Department of Pediatrics, University of Utah, Salt Lake City, UT, United States

[3]Division of Neonatology, Department of Pediatrics, University of Washington, Seattle, WA, United States

[4]Department of Computer Science and Engineering, University of Washington, Seattle, WA, United States

[5]Homer Warner Research Center, Intermountain Healthcare, Murray, UT, United States

[6]Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Gang Luo, PhD
Department of Biomedical Informatics and Medical Education
University of Washington
UW Medicine South Lake Union
850 Republican Street, Building C, Box 358047
Seattle, WA, 98109
United States
Phone: 1 206 221 4596
Fax: 1 206 221 2671
Email: gangluo@cs.wisc.edu

Abstract

Background: To improve health outcomes and cut health care costs, we often need to conduct prediction/classification using large clinical datasets (aka, clinical big data), for example, to identify high-risk patients for preventive interventions. Machine learning has been proposed as a key technology for doing this. Machine learning has won most data science competitions and could support many clinical activities, yet only 15% of hospitals use it for even limited purposes. Despite familiarity with data, health care researchers often lack machine learning expertise to directly use clinical big data, creating a hurdle in realizing value from their data. Health care researchers can work with data scientists with deep machine learning knowledge, but it takes time and effort for both parties to communicate effectively. Facing a shortage in the United States of data scientists and hiring competition from companies with deep pockets, health care systems have difficulty recruiting data scientists. Building and generalizing a machine learning model often requires hundreds to thousands of manual iterations by data scientists to select the following: (1) hyper-parameter values and complex algorithms that greatly affect model accuracy and (2) operators and periods for temporally aggregating clinical attributes (eg, whether a patient's weight kept rising in the past year). This process becomes infeasible with limited budgets.

Objective: This study's goal is to enable health care researchers to directly use clinical big data, make machine learning feasible with limited budgets and data scientist resources, and realize value from data.

Methods: This study will allow us to achieve the following: (1) finish developing the new software, Automated Machine Learning (Auto-ML), to automate model selection for machine learning with clinical big data and validate Auto-ML on seven benchmark modeling problems of clinical importance; (2) apply Auto-ML and novel methodology to two new modeling problems crucial for care management allocation and pilot one model with care managers; and (3) perform simulations to estimate the impact of adopting Auto-ML on US patient outcomes.

Results: We are currently writing Auto-ML's design document. We intend to finish our study by around the year 2022.

Conclusions: Auto-ML will generalize to various clinical prediction/classification problems. With minimal help from data scientists, health care researchers can use Auto-ML to quickly build high-quality models. This will boost wider use of machine learning in health care and improve patient outcomes.

(JMIR Res Protoc 2017;6(8):e175) doi:10.2196/resprot.7757

KEYWORDS

machine learning; automated temporal aggregation; automatic model selection; care management; clinical big data

Introduction

Barriers in Using Machine Learning to Realize Value From Clinical Big Data

Overview

To improve health outcomes and trim health care costs, we often need to perform predictions/classifications using large clinical datasets (aka, clinical big data), for example, to identify high-risk patients for preventive interventions. Machine learning has been proposed as a key technology for doing this. Machine learning studies computer algorithms, such as support vector machine, random forest, neural network, and decision tree, that learn from data [1]. Trials showed machine learning was used to help the following: (1) lower 30-day mortality rate (odds ratio [OR]=0.53) in emergency department (ED) patients having community-acquired pneumonia [2]; (2) increase on-target hemoglobin values by 8.5%-17% and reduce cardiovascular events by 15%, hospitalization days by 15%, blood transfusion events by 40%-60%, expensive darbepoetin consumption by 25%, and hemoglobin fluctuation by 13% in end-stage renal disease patients on dialysis [3-6]; (3) reduce ventilator use by 5.2 days and health care costs by US $1500 per patient at a hospital respiratory care center [7]; and (4) lower health care costs in Medicare patients' last 6 months of life by 4.5% [8].

Machine learning could support many clinical activities, but only 15% of hospitals use it for even limited purposes [9]. Compared to statistical methods like logistic regression, machine learning poses less strict assumptions on distribution of data, can increase prediction/classification accuracy, in certain cases doubling it [10-12], and has won most data science competitions [13]. Historically, machine learning was blamed for being a black box. A recent method can automatically explain any machine learning model's classification results with no accuracy loss [14,15]. Yet, two hurdles remain in using machine learning in health care. First, despite familiarity with data, health care researchers often lack machine learning expertise to directly use clinical big data. Data scientists take years of training to gain deep machine learning knowledge. Health care researchers can work with data scientists, but it takes time and effort for both parties to communicate effectively. Facing a shortage in the United States of data scientists estimated as high as 140,000+ by 2018 [16] and hiring competition from companies with deep pockets, health care systems have a hard time recruiting data scientists [17,18]. As detailed below, developing a machine learning model often requires data scientists to spend extensive time on model selection, which becomes infeasible with limited budgets. Second, some health care systems such as Kaiser Permanente, Intermountain Healthcare (IH), University of Washington Medicine (UWM), Columbia University Medical Center, Veterans Health Administration, and University of Utah Health have teams devoted to data cleaning. Health care researchers can obtain cleaned data from these systems' enterprise data warehouses (EDWs). In other health care systems, one needs to laboriously clean data before applying machine learning. This is often done with the help of database programmers and/or master-level statisticians, who can also help with data preprocessing and are easier to find than data scientists with deep machine learning knowledge. This study addresses the first hurdle and focuses on automating machine learning model selection and temporal aggregation, an important type of data preprocessing.

Barrier 1: Data Scientists Are Needed for Choosing Hyper-Parameter Values and Algorithms

Each learning algorithm includes two categories of parameters: hyper-parameters that a machine learning tool user manually sets prior to model training, and normal parameters automatically tuned in training the model (see Table 1). Given a modeling problem such as predicting 30-day hospital readmission, an analyst manually constructs a model as follows. First, select an algorithm from many pertinent ones like the approximately 40 algorithms for classification included in Waikato Environment for Knowledge Analysis (Weka) [19]. Second, set the values of the selected algorithm's hyper-parameters. Third, train the model to tune the normal parameters of the selected algorithm automatically. In case model accuracy is unsatisfactory, substitute the algorithm and/or hyper-parameter values and then retrain the model, while using some technique to avoid overfitting on the validation set [20-24]. This process is done over and over until the analyst runs out of time, has a model with good accuracy, or cannot improve further. If feature selection is considered, in each iteration the user also needs to choose a feature selection technique from many applicable ones and set its hyper-parameter values, making this process even more complex. Many possible combinations of hyper-parameter values and learning algorithms lead to hundreds to thousands of laborious and manual iterations to construct a model. These iterations need machine learning expertise, are typically done by a data scientist, and become a barrier [25].

Table 1. Two learning algorithms and their example normal parameters and hyper-parameters.

Learning algorithm	Example hyper-parameters	Example normal parameters
Support vector machine	Regularization constant C, kernel to use, tolerance parameter, ε for round-off error, a polynomial kernel's degree	Support vectors and their Lagrange multipliers
Random forest	Number of independent variables to examine at each inner node of a classification and regression tree, number of trees	Threshold value and input variable used at each inner node of a tree

Model accuracy is affected by choice of hyper-parameter values and learning algorithm. Thornton et al [25] demonstrated that for the 39 classification algorithms included in Weka, the impact on model accuracy averages 46% and can be up to 94%. Even considering a few popular algorithms like random forest and support vector machine, the impact is still above 20% on two-thirds of 21 benchmark datasets. The good choice changes by the particular modeling problem. Computer science researchers have investigated methods for automatically searching hyper-parameter values and algorithms [26]. Some methods can reach equal or better results compared to data scientists' manual tuning [27,28]. But in case a large number of algorithms are examined, efforts like Auto-WEKA [25,29-31], hyperopt-sklearn [28], and MLbase [32,33] cannot effectively handle large datasets in reasonable time.

A hurdle to automatic search is the amount of time needed to assess on an entire dataset a combination of hyper-parameter values and a learning algorithm. On a modern computer, it takes 2 days to train the champion ensemble model that won the Practice Fusion Diabetes Classification Challenge [34] one time on 9948 patients with 133 input or independent variables (aka, features). Even when disregarding ensembles of more than five base models, aborting long-running tests, and greatly limiting the hyper-parameter value search space (eg, allowing no more than 256 decision trees in a random forest), all impacting search result quality, more than 30 minutes are needed to test an average combination on 12,000 rows (ie, data instances) with 784 attributes [35]. To ensure search result quality, automation efforts often test more than 1000 combinations on the whole dataset [35], leading to months of search time. On a dataset with several dozen attributes and several thousand rows, a search can still take several days [25]. In reality, search time could be thousands of times longer even with a computer cluster for five reasons:

1. Model building is iterative. When a collection of clinical attributes yields low model accuracy, the analyst can include other attributes to boost accuracy. Every iteration takes a new search for hyper-parameter values and learning algorithms.

2. Frequently, ensembles of a large number of base models reach higher accuracy. The training time of an ensemble model rises proportionally to the number of base models.

3. Hyper-parameter values over a broad range are often used to achieve higher accuracy. The above champion ensemble model [34] uses 12 base models. Each random forest base model uses at least 15,000 decision trees.

4. Numerous rows, often from multiple health care systems, can reside in a dataset.

5. Numerous attributes (eg, derived from genomic or textual data) can exist in a dataset. In a hospital without genomic data, a model for readmission prediction was built using 195,901 patients and 3956 attributes already [36]. An algorithm's execution time rises proportionally to the number of attributes at a minimum and often superlinearly with the number of rows. Irrespective of whether search is done manually or automatically, a slow speed in search frequently causes a search to be terminated early, producing suboptimal model accuracy [35].

Barrier 2: Data Scientists Are Needed for Temporally Aggregating Clinical Attributes

Numerous clinical attributes are documented over time needing aggregation prior to machine learning (eg, weight at each patient visit is combined to check whether a patient's weight kept rising in the previous year). An aggregation period and operator pair (eg, increasing trend, average, count, and maximum) need to be specified for every attribute separately to compute an aggregate value. Usually, clinicians designate pairs and data scientists perform computation. Numerous pairs could be clinically meaningful. The ones that produce high accuracy change by the particular modeling problem and are usually not known in advance. Granted a modeling problem, the analyst picks one or more pairs for each attribute manually, then constructs a model. In case model accuracy is unsatisfactory, the analyst substitutes pairs for some attributes and reconstructs the model, while using some technique to avoid overfitting on the validation set [20-24]. This process between data scientists and clinicians is frequently repeated many times and becomes a barrier. No comprehensive aggregation operator list exists, demanding care to not omit effective operators.

Barrier 3: Data Scientists Are Needed for Generalizing Models

A model that is built and is accurate in a health care system often performs poorly and needs to be rebuilt for another system [37], with differing patients, practice patterns, and collected attributes impacting model selection [38,39]. This needs data scientists and is a barrier, as a system often needs many models for diverse clinical activities.

As often quoted, McKinsey estimates that proper use of clinical big data can bring more than US $300 billion in value to US health care each year [16]. The achievable value is surely less, but still significant. To realize value from data, we need new approaches to enable health care researchers to directly use clinical big data and make machine learning feasible with limited budgets and data scientist resources.

Our Proposed Software

To fill the gap, we will (1) finish developing the open source software, Automated Machine Learning (Auto-ML), to efficiently automate model selection for machine learning with clinical big data and validate Auto-ML on seven benchmark modeling problems of clinical importance, (2) apply Auto-ML and novel methodology to two new modeling problems crucial for care management allocation and pilot one model with care managers, and (3) perform simulations to estimate the impact of adopting Auto-ML on US patient outcomes. We hypothesize that adopting Auto-ML will improve outcomes. Conceptually, Auto-ML will be an automated version of Weka [19] supporting automated temporal aggregation. With minimal help from data scientists, health care researchers can use Auto-ML to quickly build high-quality models. This expands the human resource pool for clinical machine learning and aligns with the industry trend of citizen data scientists, where an organization arms its talent with tools to do deep analytics [40]. Auto-ML can greatly

reduce the time and cost required of scarce data scientists, busy clinicians, and computing resources in developing models; enable fast turnaround; and facilitate green computing. The faster a high-quality model gets built and deployed, the earlier it can bring outcome improvement. Auto-ML is not used to reach the maximum possible model accuracy in theory, which is hard to do in reasonable time. Instead, Auto-ML is used to quickly build high-quality models. If needed, data scientists and health care researchers can manually fine-tune them further.

Auto-ML will efficiently automate a selection of feature selection techniques, hyper-parameter values, learning algorithms, and temporal aggregation operators and periods. Auto-ML will continuously show, as a function of time given for model selection, forecasted model accuracy as well as expected patient outcomes of model use. If trends are not promising, the user can abort, add more clinical attributes, and restart. Auto-ML is able to operate on a cluster of computers for scalable processing.

Gaps in Patient Identification for Care Management and Our Proposed Solutions

Overview

Aim 1 involves finishing development of Auto-ML. To improve patient identification and outcomes for care management, Aim 2 involves applying Auto-ML to two new modeling problems by doing the following: (1) use a health care system's incomplete medical (ie, clinical and/or administrative) data to find future high-cost, diabetic patients and (2) use vast attributes in modern electronic medical records to find future hospital users in asthmatic patients.

Widely used for chronic diseases like asthma and diabetes, care management applies early interventions to high-risk patients to avoid high costs and health status decline [41-43]. In the United States, 7.1 million children (9.6%) and 18.7 million adults (8.0%) [44] have asthma [45,46]. Every year, asthma causes 1.8 million ED visits, 439,000 hospitalizations, US $56 billion in health care costs [47], and 3630 deaths [44]. Proper use of care management can cut down asthma exacerbations; trim costs by up to 15%; drop ED visits and hospital admissions and readmissions by up to 40%; and enhance quality of life, treatment adherence, and patient satisfaction by 30%-60% [42,48-54]. This impacts 63% of annual total asthma costs from asthma exacerbations [51,55].

For care management to be effective within resource constraints, we should only enroll patients with the worst prognosis or those anticipated to have the highest costs. Predictive modeling is widely used for care management [56] as the best method for finding high-risk patients [57], but current approaches have two gaps, as discussed below.

Scope Gap

Often, a health care system has incomplete medical data on many of its patients, as a patient's complete data may spread across several health care systems [58,59]. Typical models for predicting a patient's costs assume complete data [60-62]. A system usually does not apply models to patients on whom it possibly has incomplete data. As future high-cost patients are not found, care management is not used on them. This limits care management's scope of use to improve outcomes. UWM is seeking a way to fill the gap, notably for patients with diabetes. To do this, we will use a constraint to find patients who tend to get most of their care at UWM, use UWM's incomplete data to build a model, and apply it to them to facilitate care management.

Accuracy Gap

Existing models for predicting hospital use (ie, inpatient stay or ED visit) in asthmatic patients have low accuracy [63-68]. A typical model [65] missed 75% of future hospital users. A total of 78% of patients in the high-risk group chosen by the model did not use hospitals in the next year. Two factors degrade accuracy. First, several dozen risk factors for hospital use in asthma are known, including age, gender, race/ethnicity, asthma medication use, prior health care use, comorbidities (eg, ischemic heart disease, rhinitis, sinusitis, reflux, anxiety-depression, diabetes, cataracts, chronic bronchitis, and chronic obstructive pulmonary disease), allergies, lung function, number of asthma medication prescribers as a measure of continuity of care, health insurance type, lab test results (eg, total serum immunoglobulin E level and eosinophil count), body mass index, smoking status, secondhand smoke exposure, the ratio of controller to total asthma medications, frequency of nonasthma visits, number of procedures, number of diagnoses, number of prescription drug claims, and asthma questionnaire results (eg, frequency of asthma symptom occurrence, interference with normal activity, nighttime awakening, reliever use for symptom control, forced expiratory volume in 1 second [FEV1], peak expiratory flow rate, FEV1/forced vital capacity ratio, asthma control test score, number of exacerbations last year, controller use, asthma-related acute care, asthma trigger reduction, and asthma medication) [55,63,65,67-73]. Yet, a typical model uses fewer than 10 of these risk factors [63-67]. Existing models were built using data from either clinical trials or outdated electronic medical records gathering limited attributes [74]. No published model uses all known risk factors in modern electronic medical records gathering vast attributes [74]. Second, as with many diseases, many attributes predictive of hospital use in asthma have not been found yet. If we could enroll 5% more of future hospital users in care management, we could avoid up to 8780 hospitalizations and 36,000 ED visits for asthma each year. IH is seeking a way to fill the gap. To do this, we will use vast attributes in IH electronic medical records to build a model predicting hospital use in asthma. The attributes will cover many known risk factors for hospital use in asthma and will be used to find new predictive factors.

Innovation

Our study is innovative for multiple reasons:

1. With the new software that will be built as part of our project, for the first time, health care researchers with limited machine learning knowledge will quickly be able to build high-quality machine learning models with minimal help from data scientists. The cost and time required of data scientists and clinicians in doing machine learning will be greatly reduced. Also, it will become possible to widely use machine learning in health care to realize value from clinical big data and improve patient

outcomes. No existing software can greatly cut the long time required of data scientists in building and generalizing models.

2. We will direct care management to more patients needing it more precisely than current approaches. For patients on whom it possibly has incomplete medical data, a health care system usually does not apply predictive models to find candidates for care management. Existing models for predicting hospital use in asthmatic patients were built mainly using a small set of patients (eg, <1000) or attributes (eg, <10), creating a hurdle in finding many predictive attributes and their interactions. Many known risk factors' predictive power for hospital use in asthma is unused. In contrast, we will expand the set of diabetic adults for whom predictive models and care management can be used. We will use many asthmatic children and attributes to build new, accurate models for hospital use. The attributes will cover many known risk factors for hospital use in asthma and will be used to find new predictive factors. Our approaches to using incomplete data and vast attributes are new, with principles generalizable to many clinical applications.

3. Our software will (1) automatically choose hyper-parameter values, feature selection techniques, and algorithms for a particular machine learning problem faster than existing methods; (2) efficiently and automatically choose operators and periods for temporally aggregating clinical attributes—no such method currently exists; longitudinal data analysis [75] models the dependent variable; in contrast, our temporal aggregation can use any function of independent variables; (3) continuously show, as a function of time given for model selection, estimated patient outcomes of model use and forecasted model accuracy—for the first time, one can obtain feedback continuously throughout automatic model selection; and (4) enable fast turnaround. There is no such software at present.

4. We will systematically compile the first list of regularly used operators for temporally aggregating clinical attributes. The list can be reused for future clinical data analysis studies. Using MapReduce [76] for distributed computing, we will provide the first implementation of many aggregation operators not offered by current big data software such as Hadoop [77] and Spark [78].

5. We will estimate the impact of adopting our automated machine learning software on US patient outcomes in two scenarios; no such estimate has ever been made. Our impact estimation method is new and can be applied to other scenarios and similar software.

In summary, this study is significant in that it makes machine learning feasible with limited budgets and data scientist resources to help realize value from clinical big data and improve patient outcomes. The models that will be built for the two new modeling problems will help improve care management outcomes.

Methods

Overview

Auto-ML will be built atop current big data software, enabling it to operate on one computer or a cluster. Built atop the Hadoop distributed file system, Spark [78] is a major open source software system supporting MapReduce [76] for distributed computing. Spark has an accompanying machine learning library, MLlib [79]. Spark is able to perform machine learning more than 100 times quicker than Hadoop [80]. Auto-ML will be built using the Spark package as well as novel techniques to address the current software's limitations.

Aim 1

Overview

Our first aim is to finish developing Auto-ML to automate model selection for machine learning with clinical big data and validate Auto-ML on seven benchmark modeling problems of clinical importance.

Figure 1 compares Auto-ML's approach of constructing models to the present one. Four steps are carried out sequentially during machine learning: temporally aggregate clinical attributes; choose hyper-parameter values, feature selection techniques, and algorithms; construct models; and assess models. The temporal aggregation step is optional (eg, when no repeatedly recorded attribute exists). Auto-ML will use Spark as the basis for distributed computing. Auto-ML will be coded in Java so it can use the open source software systems Spark and Weka, which all have a Java application programming interface and/or are coded in Java. The user will specify the storage location of the dataset in Auto-ML's graphical input interface. Auto-ML will then put the dataset into Spark prior to analysis.

Figure 1. Auto-ML's approach of constructing machine learning models versus the present one.

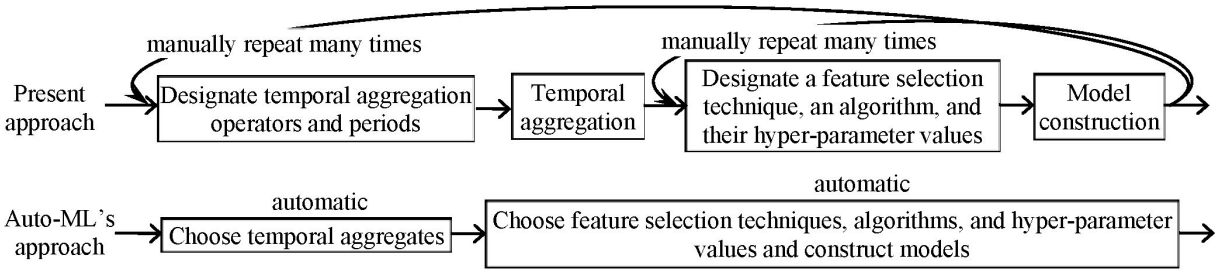
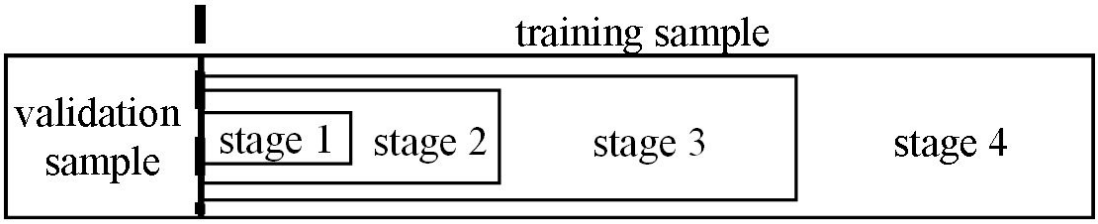


Figure 2. Progressive sampling adopted in our draft automatic model selection method.



Auto-ML's Machine Learning Functions

Auto-ML will integrate MLlib [79] and Weka's [19] machine learning functions by altering source code and/or invoking the Java application programming interfaces. As a broadly used machine learning tool kit, Weka includes many popular feature selection techniques and learning algorithms. distributedWekaSpark [81] is the distributed computing package of Weka for Spark that is able to operate on a computer cluster. MLlib is a distributed machine learning library in Spark implementing some techniques and algorithms supported by Weka. Auto-ML will support all techniques and algorithms available in Weka. Whenever possible, Auto-ML will use MLlib's code, which fuses with Spark better than distributedWekaSpark's code [81].

Weka's [19] graphical user interface covers feature selection (optional), model construction, and model assessment. In the input interface, the Weka user designates the dependent variable, independent variables, data file, learning algorithm, and the hyper-parameter values of the algorithm. After the user clicks the start button, Weka constructs a model and shows its performance measures. For machine learning, Auto-ML's graphical user interface will work similarly with two main differences. First, in Weka, the user must specify an algorithm prior to model building. Like Auto-WEKA [25], Auto-ML will use a hyper-parameter to represent the option of feature selection technique and automatically select the hyper-parameter values, technique, and algorithm. The user may override the choice of Auto-ML. Second, to facilitate the user in tracking the automatic selection's progress, Auto-ML shows a curve presenting the highest accuracy reached over time. The user can terminate the process at any moment and obtain the most accurate model built. In the following sections, we outline the main techniques that we will use to build Auto-ML.

Aim 1 (a)

Overview

For Aim 1 (a), we aim to devise a method to efficiently and automatically choose hyper-parameter values, feature selection techniques, and algorithms. Our review paper [26] showed that few automatic selection methods [25,28-31,82] have been fully implemented and can manage an arbitrary number of combinations of hyper-parameter values and many learning algorithms. All of these methods are similar to or based on the Auto-WEKA automatic selection approach [25], yet none of them can efficiently handle large datasets. To overcome the current methods' inefficiencies, we drafted a method based on Bayesian optimization for response surface to rapidly identify, for a specific modeling problem, a good combination of hyper-parameter values, a feature selection technique, and a learning algorithm when a large number of algorithms and techniques are examined [35,83]. The method represents the option of technique as a special hyper-parameter; proceeds in stages; and conducts progressive sampling [84], filtering, as well as fine-tuning to rapidly shrink the search space. We conduct fast trials on a small sample taken from the dataset to drop unpromising combinations early, reserving resources to fine-tune promising ones. A combination is promising when a model built using the combination and the sample reaches an error rate below a beginning threshold. Then, we decrease the threshold, enlarge the sample, test and adjust combinations, and cut the search space several times. At the last stage, we find an effective combination using the full dataset.

More specifically, at each stage our method uses a training sample and a validation sample. They have no overlap and contain data instances randomly chosen from the dataset. We keep the validation sample the same and expand the training sample across stages (see Figure 2). At the first stage, we start from a small training sample. For each learning algorithm, we evaluate a fixed number of random hyper-parameter value combinations, if any, as well as its default one. To evaluate a combination, we use the combination, the training sample, and algorithm to construct a model, then use the validation sample to assess the model's error rate. We identify and remove unpromising algorithms based on the test results. At each subsequent stage that is not the last one, we enlarge the training sample. For each remaining algorithm, we construct a separate regression model, use a Bayesian optimization for response surface approach to choose several new hyper-parameter value combinations, and test these combinations. We identify and remove additional unpromising algorithms based on the test results. At the last stage, we do some final tests on the full dataset to come up with the ultimate search result.

Our draft method needs further optimization for efficiency and effectiveness. To do this, we will expand the draft method to include multiple optimization techniques: the seven outlined in our design paper [24] and the six described in the following sections.

Technique 1

In Technique 1, we will use two validation samples to help avoid overfitting. At each stage except for the last one, our draft method [35,83] uses the same validation sample containing a moderate number of data instances to perform many tests. This could lead to overfitting to the validation sample [20-23] that will misguide future search. To help avoid overfitting, we will

use two validation samples of equal size with as little overlap as possible, and reduce the frequency of revealing information about the second validation sample [23]. When the dataset has enough data instances, the two validation samples will have no overlap. For a combination of hyper-parameter values and a learning algorithm, we use the combination and the training sample to construct a model and assess the model's error rate twice, once on either validation sample. Intuitively, the two error rates would be roughly the same in the absence of overfitting. If the error rate on the first validation sample is higher than a specific threshold (eg, in the top 50% of the error rates on the first validation sample of all combinations tested so far at this stage), we use it as the combination's error rate estimate. Regardless of its exact value, a high error rate estimate will guide future search to avoid the combination's neighborhood. If the threshold is not exceeded, we compare the error rate on the first validation sample with that on the second. If the former is not lower than the latter by a certain threshold (eg, 5%), we use the former as the combination's error rate estimate. Otherwise, we use the latter as the combination's error rate estimate, as overfitting to the first validation sample is likely to have occurred.

The above approach uses the same two validation samples across different stages. Alternatively, if the dataset contains many data instances, we can use a different validation sample at each stage. Each time we arrive at a new stage, we redo sampling to obtain a new validation sample. This also helps avoid overfitting to the same validation sample that is repeatedly used. We will compare the two approaches and choose the one that performs better.

Technique 2

In Technique 2, we will use multiple feature selection techniques concurrently to drop unpromising features early. Feature selection and model building time rises proportionally to the number of features at a minimum. Doing a test is slow when many features exist in the dataset. To tackle this issue, we previously proposed that before doing tests, we apply a feature selection technique to the dataset, or a large sample of it, and rapidly drop features not likely to have high predictive power [24]. Yet, like the "no free lunch" theorem [85] shows, no technique can guarantee good performance in all cases. Relying on a single technique can be risky, causing predictive features to be dropped erroneously. To reduce the risk, we will use multiple techniques concurrently. A feature is dropped only if at least a certain number of these techniques all regard it as unpromising.

Technique 3

In Technique 3, at the first stage for each learning algorithm, we will ensure a minimum number of tests conducted on every feature evaluator and feature search method. Every feature selection technique adopts a feature evaluator as well as a feature search method [25]. At the first stage for no learning algorithm, our draft method guarantees the number of tests conducted on every feature evaluator or feature search method. Without enough tests, we cannot tell how well a feature evaluator or feature search method works with the algorithm. To tackle this issue, at the first stage for each algorithm, we will check the number of tests conducted on every feature evaluator and feature search method. If the number for a feature evaluator or feature search method is smaller than a specific threshold (eg, 3), we will conduct more tests for the feature evaluator or feature search method to make up the difference. This approach can be adopted for several other components of a data analytic pipeline [86], such as handling imbalanced classes and missing values.

Technique 4

In Technique 4, we will share information on the best few results obtained so far among different learning algorithms. Our draft method conducts a separate set of tests for every algorithm. When conducting tests for an algorithm, we may find a combination of a feature selection technique and its hyper-parameter values with superior performance. Yet, the combination may not be tested together with other algorithms, as its information is not shared with them. This can degrade the ultimate search result's quality. To tackle this issue, we will share information on the best few results obtained so far among different algorithms. At the end of each stage except for the last one, we will identify a prechosen number n_1 (eg, 3) of combinations of algorithms, techniques, and hyper-parameter values that achieve the lowest error rates among all combinations examined so far. Then we will extract the corresponding n_2 combinations of techniques and their hyper-parameter values. Typically, n_2 is equal to n_1. Occasionally, n_2 can be smaller than n_1, as the same combination of a technique and its hyper-parameter values may appear in more than one of the n_1 combinations. At the next stage, for each remaining algorithm, we ensure each of the n_2 combinations of techniques and their hyper-parameter values is tested by adding additional tests, if needed.

Technique 5

In Technique 5, for a dataset with relatively few data instances, we will dynamically allocate its data instances between the training and validation samples across stages. A dataset with relatively few data instances can still be large if it contains many features. In this case, our draft method uses a fixed portion of the dataset as the validation sample, which includes a small number of data instances. Because of insufficient testing, the error rate estimates obtained on the trained models can be nonrobust, degrading the ultimate search result's quality. To tackle this issue, we will dynamically allocate the data instances in the dataset between the training and validation samples across stages. At each stage except for the last one, we give all data instances that are in the dataset, but not in the training sample, to the validation sample. With more data instances in the validation sample, the error rate estimates obtained on the trained models can be more robust. Krueger et al [87] used a similar approach to perform fast cross-validation to select a good hyper-parameter value combination for a given learning algorithm and modeling problem.

Technique 6

In Technique 6, we will consider distances between hyper-parameter value combinations when choosing randomly sampled combinations for testing. At each stage that is neither the first nor the final one, for each remaining learning algorithm,

our draft method performs one or more rounds of Bayesian optimization. In each round, several new and randomly sampled combinations are chosen out of many for testing and used to adjust the regression model. For the regression model to guide search well, the combinations chosen for testing need to have a reasonable coverage of the hyper-parameter space rather than all reside in a small region. To achieve this, we will attempt to ensure that each randomly sampled combination chosen for testing is separated from each other combination chosen for testing by at least a specific distance. The distance threshold may decrease over stages.

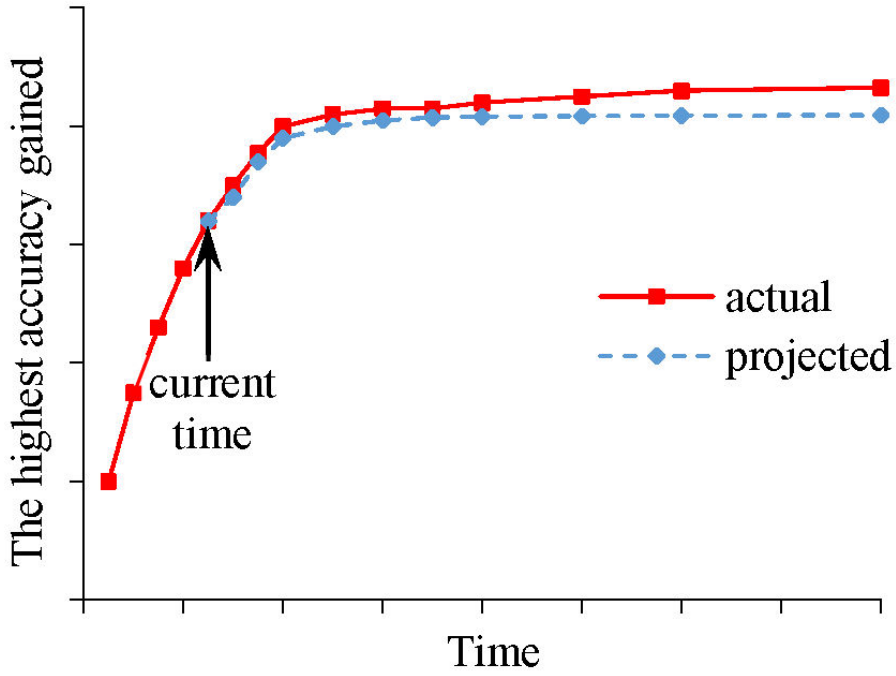
Aim 1 (b)

For Aim 1 (b), we aim to devise a method to efficiently and automatically choose operators and periods for temporally aggregating clinical attributes. Our design paper [24] outlines our method for automating the process of temporally aggregating clinical attributes. We will flesh out our method's technical details. Our automation method needs disease-specific knowledge on aggregation operators and periods compiled by clinicians and stored in Auto-ML. Various medical datasets use differing schemas, medical coding systems, and medical terminologies, forming a hurdle in applying precompiled knowledge. To tackle this, the automated temporal aggregation function of Auto-ML demands that the dataset, except for the dependent variable, complies with the Observational Medical Outcomes Partnership (OMOP) common data model [88] and its linked standardized terminologies [89]. Since OMOP standardizes administrative and clinical attributes from 10 or more large US health care systems [90,91], Auto-ML can be adopted for datasets from those systems. We intend to include support for the National Patient-Centered Clinical Research Network (PCORnet) [92] and Informatics for Integrating Biology and the Bedside (i2b2) common data models [93] in the future.

Aim 1 (c)

For Aim 1 (c), we aim to continuously show, as a function of time given for model selection, forecasted model accuracy and projected patient outcomes of model use. During automatic selection, to be more useful and user friendly, Auto-ML will show projected patient outcomes of model use and forecasted model accuracy as a function of time given for model selection (see Figure 3). Our design paper [24] outlines our method for doing this. We will flesh out our method's technical details and write a user manual for Auto-ML.

Figure 3. The highest model accuracy gained by Auto-ML over time.



Aim 1 (d)

Overview

For Aim 1 (d), we aim to validate Auto-ML on seven benchmark modeling problems. We will perform tests with health care researchers using seven modeling problems and datasets that we worked on before. Each problem uses a different dataset from a distinct health care system. We chose these problems because they are related to common diseases, are clinically important, and have readily accessible datasets. Auto-ML can be used for other clinical activities.

Subject Recruitment

Via announcements in our institution's email lists and personal contact, we will recruit 25 health care researchers from UWM, which houses approximately 2500 faculty members, most doing health care research. These health care researchers would regard their familiarity with medical data at the MD level, but would regard their machine learning knowledge as below the level

taught in a typical machine learning course for computer science undergraduates. We will conduct purposeful sampling to ensure enough variability [94]. All test participants will have fulfilled UWM's required training for information security and privacy policy. Participants will receive pseudonyms linking their responses to questions for privacy protection. After giving consent, each will get a copy of the task description, Auto-ML's user manual, and the metadata document detailing the attributes in the dataset. Upon task completion, each will receive US $200 as compensation for participation.

Computing Environment

We will perform all of our experiments on a Health Insurance Portability and Accountability Act (HIPAA)-compliant computer cluster at the University of Washington. After obtaining proper authorization, all test participants and research team members at the University of Washington will be able to access the cluster using their university computers.

Modeling Problem 1

Modeling Problem 1 will consist of predicting the diagnosis of type 2 diabetes in adult patients in the next year.

Dataset and Patient Population

The clinical and administrative dataset is deidentified and publicly available from the Practice Fusion Diabetes Classification Challenge [15,34], containing 3-year (2009-2012) records as well as the labels of 9948 adult patients from all US states in the following year. A total of 1904 of these patients had a diagnosis of type 2 diabetes in the following year. The dataset comes from an electronic medical record vendor's EDW; includes repeatedly recorded attributes; and covers patient demographics, allergies, diagnoses, immunizations, medications, smoking status, lab results, and vital signs. We will put this dataset in the OMOP common data model form with its linked standardized terminologies.

Model Information

The dependent variable is whether a patient had a diagnosis of type 2 diabetes in the following year. Two-thirds of patients will be randomly selected and put into the training set to construct models. The remaining one-third of patients will form the test set for assessing model performance. We will use the area under the receiver operating characteristic curve (AUC) [19] performance metric.

Modeling Problems 2-7

Each of the six problems from Modeling Problems 2-7 uses a distinct, deidentified, and publicly available dataset from the University of California, Irvine machine learning repository [95] to perform a task: (1) Arcene: classify mass spectrometric data into cancer versus normal patterns; (2) Arrhythmia: classify 12-lead electrocardiogram recordings into one of 16 groups about cardiac arrhythmia; (3) Cardiotocography: classify fetal cardiotocograms into one of three fetal states; (4) Diabetic Retinopathy Debrecen: use features obtained from the Messidor image set to detect whether an image includes signs of diabetic retinopathy; (5) Mammographic Mass: use Breast Imaging Reporting and Data System attributes and patient age to separate benign from malignant mammographic masses; (6) Parkinson Speech: use sound recordings to identify Parkinson's disease patients.

No dataset has repeatedly recorded attributes needing temporal aggregation. The repository [95] includes a detailed description of the problems and datasets. For each dataset, two-thirds of it will be randomly selected and put into the training set to construct models. The remaining one-third will form the test set for assessing model performance. We will use the accuracy metric suitable for multi-class classification.

Build Models

We are familiar with the literature on the seven modeling problems. For each problem, our data scientist Dr Luo (GL) will work with the clinicians in our team and manually build a machine learning model with as high accuracy as possible. This accuracy will serve as the gold standard reflecting current best practice of model building. Each of the 25 recruited health care researchers will be randomly given a problem and use Auto-ML to build models for it.

Performance Evaluation and Sample Size Justification

We will test the hypothesis that at least 60% of health care researchers can use Auto-ML to achieve model accuracy of at least 95% of the gold standard. When 60% of health care researchers can actually achieve model accuracy of at least 95% of the gold standard, a sample size of 25 health care researchers produces a one-sided 95% lower confidence limit of 42%.

User Feedback

When model construction is finished, we will use both open-ended and semistructured questions to survey the 25 health care researchers. As detailed in our design paper [83], we will obtain quantitative outcome measures covering model accuracy, time on task, self-efficacy for constructing machine learning models with clinical big data, satisfaction, trustworthiness, adequacy, and quality of documentation. The questionnaire will contain a text field for gathering comments on Auto-ML. We will refine and finalize Auto-ML by considering suggestions from those comments. We will perform a user satisfaction survey using the System Usability Scale (SUS), a widely used industry standard [96,97] on overall satisfaction ratings for products.

Analysis

We will use the accepted inductive approach endorsed by Patton et al [94,98] to do qualitative analysis. We will put the 25 health care researchers' textual comments into ATLAS.ti Version 8 (ATLAS.ti Scientific Software Development GmbH), a qualitative analysis software tool [99]. The research team will independently highlight quotations related to the issue of using Auto-ML. We will examine quotations, categorize them into precodes, and merge them into categories in multiple iterations. We will synthesize categories to find general themes. Quantitative analyses will include adding the scores in the SUS and presenting every quantitative outcome measure's descriptive statistics.

Aim 2

Overview Aim 2 involves applying Auto-ML and novel methodology to two new modeling problems crucial for care management allocation, to which our institutions are seeking solutions, and pilot one model with care managers. Both modeling problems use datasets that have repeatedly recorded attributes. We will put the datasets in the OMOP common data model form with its linked standardized terminologies. We will use the same computing environment and recruiting method mentioned in Aim 1 (d). We will recruit two health care researchers not engaged in Aim 1 (d). Each will be randomly given a problem and use Auto-ML to build models for it. Upon task completion, each will receive US $200 as compensation for participation.

Modeling Problem 8

Modeling Problem 8 involves the use of vast attributes in modern IH electronic medical records to predict hospital use in asthmatic children in the next year.

Patient Population

The patient population consists of IH pediatric patients (0-17 years of age) with asthma in 2005-2016, identified by Schatz et al's method [63,100,101] as having the following: (1) at least one diagnosis code of asthma according to the International Classification of Diseases, Ninth Revision (ICD-9) (ie, 493.xx), or the International Classification of Diseases, Tenth Revision (ICD-10) (ie, J45/J46.*); or (2) two or more "asthma-related medication dispensings (excluding oral steroids) in a one-year period, including β-agonists (excluding oral terbutaline), inhaled steroids, other inhaled anti-inflammatory drugs, and oral leukotriene modifiers."

Dataset

By running Oracle database Structured Query Language (SQL) queries, our contracted IH data analyst will extract from the IH EDW a deidentified, clinical and administrative dataset, encrypt it, and securely transfer it to a HIPAA-compliant computer cluster for secondary analysis. For each of the last 5 years, the data cover approximately 27,000 asthmatic children. The dataset is the electronic documentation of approximately 95% of pediatric care in Utah [102,103] and includes around 400 attributes partially listed in our paper [14]. These attributes cover many known risk factors for hospital use in asthma and can be used to find new predictive factors.

Model Information

The dependent variable is whether an asthmatic patient incurred hospital use—inpatient stay or ED visit—with a primary diagnosis of asthma (ie, ICD-9 493.xx or ICD-10 J45/J46.*) in the following year [14,63,64]. As outcomes need to be computed for the following year, we effectively have 11 years of IH data. We will construct models using the data in the first 10 years and acquire a model's accuracy estimate via testing on the data in the 11th year. This mirrors future use of the model in practice. We will use the AUC [19] performance metric.

Modeling Problem 9

Modeling Problem 9 involves using UWM's incomplete data to predict individual diabetic adults' costs in the next year.

Patient Population

The patient population includes UWM adult patients (18 years of age or older) with diabetes in 2012-2016, identified by the method in Neuvirth et al [104] as having one or more hemoglobin A1c test results of 6.5% or higher.

Dataset

A UWM data analyst will run SQL Server database SQL queries to extract from the UWM EDW a deidentified, clinical and administrative dataset, encrypt it, and securely transfer it to a HIPAA-compliant computer cluster for secondary analysis. The data cover approximately 28,000 diabetic adults per year. Other details of the dataset are similar to those in Modeling Problem 8.

Model Information

The dependent variable is a diabetic patient's total allowed cost to UWM in the following year [60,61]. Allowed costs are less inflated than billed costs and less subject to variation due to member cost sharing than net incurred claims [60]. We will adopt the medical consumer price index [105] to convert all costs to 2016 US dollars to handle inflation. As outcomes need to be computed for the following year, we effectively have 4 years of UWM data. We will construct models using the data in the first 3 years, and acquire a model's accuracy estimate via testing on the data in the 4th year. This mirrors future use of the model in practice. We will use the R^2 performance metric [61].

To fill the scope gap mentioned in the introduction, we will use a constraint to find patients who tend to get most of their care at UWM. Intuitively, it is easier to identify future high-cost patients among them than among others. We will use UWM's incomplete data to build a cost prediction model and apply it to them. Regardless of his/her total future cost at non-UWM facilities, a patient who will incur high cost at UWM can be a candidate for care management. By care managing future high-cost patients identified by the model, we will expand the scope of using care management to improve outcomes. The principle of our approach to using incomplete data generalizes to many other clinical applications.

Several candidate constraints exist: (1) the patient had two or more visits to UWM in the past year, (2) the patient has a UWM primary care physician and lives within 5 miles of a UWM hospital, and (3) the patient saw a primary care physician or endocrinologist at UWM in the past year and lives within 60 miles (ie, around 1 hour of driving distance) of a UWM hospital. UWM primary care physicians tend to make referrals within UWM. Endocrinologists often serve some of the same roles as primary care physicians. Usually, a patient incurs high cost because of hospital use. As patients living far away from UWM hospitals are less likely to use them, UWM tends to have less of these patients' medical data. We will refine the three candidate constraints and investigate others. To select the constraint to be used, we will use PreManage data that UWM has on all of its patients. PreManage is Collective Medical Technologies Inc's commercial product providing encounter and diagnosis data on inpatient stays and ED visits at many US hospitals [106]. PreManage data cover 105 (approximately 94%)

hospitals in Washington, including the four hospitals of UWM. Using UWM data and grouper models like the Clinical Classifications Software system to group diagnosis codes and reduce features [60], we will build two models: one for estimating an inpatient stay's allowed cost and another for estimating an ED visit's allowed cost based on patient demographics and diagnosis data. We will use UWM patient demographics data, PreManage diagnosis data, and the two models to estimate the allowed cost of each of a UWM patient's non-UWM inpatient stays and ED visits reflected by PreManage encounter data. By aggregating the estimated costs of individual non-UWM inpatient stays and ED visits, we will assess each UWM patient's portion of cost spent at non-UWM hospitals and use the portions to evaluate every candidate constraint. If a health care system does not have enough data to make the two models reasonably accurate, it can use the average costs of an inpatient stay and ED visit to assess each patient's portion of cost spent at external hospitals. If a system has an insurance plan's complete claim data on some of its patients, it can use the data similarly.

Performance Evaluation and Sample Size Justification

For each of the two new modeling problems, we will test the hypothesis that health care researchers are able to use Auto-ML to achieve higher model accuracy than existing approaches. We will regard Aim 2 as partly successful if we accept the hypothesis in only one problem, and completely successful if we accept the hypothesis in both problems.

For Modeling Problem 8, we will compare the accuracies reached by the model built by the health care researcher and the model in Schatz et al [65]. The first model is built using Auto-ML and vast attributes in modern IH electronic medical records. The second model depicting the existing approach was built using a few known risk factors for hospital use in asthma. Using vast attributes can increase prediction accuracy [107]. We will accept the hypothesis when the first model reaches a higher AUC than the second one by at least .05. Existing predictive models for hospital use in asthma usually achieve an AUC far below .8 [63-68]. Assuming these two models' prediction results have a correlation coefficient of .6 for both classes and performing a two-sided Z test at a significance level of .05, a sample size of 561 data instances per class provides 90% power to find a discrepancy of .05 between the two models' AUCs. The IH data in the 11th year include about 27,000 asthmatic children, offering enough power to test our hypothesis. Using many patients is essential for improving prediction accuracy, although only a small sample size is needed to show statistical significance.

For Modeling Problem 9, we will compare the accuracies gained by two models. The patient cohort includes those satisfying the chosen constraint. The first model is built by the health care researcher using Auto-ML and clinical and administrative data. The second model depicting the existing approach is a commercial claims-based one available at UWM achieving an R^2 less than 20%. Although the second model was not designed for such use, we will apply it to the patient cohort on whom UWM possibly has incomplete data, which is better than the normal practice of making no predictions. Adding clinical data can increase prediction accuracy [108]. We will accept the hypothesis when the first model reaches a higher R^2 than the second one by at least 5%. Using an F test at a significance level of .05 and under the assumption of the existence of 20 features from clinical data in addition to 300 or fewer features used in the second model, a sample size of 443 patients provides 90% power to identify an increase of 5% in R^2 from 20%. Using the second candidate constraint, we estimate that the patient cohort will cover approximately 22% of diabetic adult patients at UWM. The 4th year's UWM data include approximately 28,000 diabetic adults, offering enough power to test our hypothesis.

Pilot With Care Managers

We will pilot the model the health care researcher will build for Modeling Problem 9 with UWM care managers. As a UWM operational project, we are working on this modeling problem and have access to around 25 UWM care managers. Via announcing in their email lists and personal contact, we will recruit five care managers. We will conduct purposeful sampling to ensure enough variability [94]. All test participants will give consent and have fulfilled UWM's required training for information security and privacy policy. Participants will receive pseudonyms linking their responses to questions for privacy protection. Upon task completion, each will receive US $200 as compensation for participation.

We will use our previously developed method [15] to automatically explain the model's prediction results. For each care manager, we will randomly select 20 UWM diabetic adult patients, half of whom the model predicts will incur a cost of more than US $30,000. The care manager is unaware of any of these patients' outcomes in the next year. For each patient, we will first show the care manager the historical, deidentified patient attributes, then show the prediction result and automatically generated explanations, and finally survey him/her using both open-ended and semistructured questions. The questions will cover whether the prediction result and explanations will change his/her enrollment decision on the patient, their usefulness, and their trustworthiness as shown in Table 2. The questionnaire will contain a text field for gathering comments. We will analyze collected information in a similar way to Aim 1 (d).

Table 2. The dependent variable list.

Variable	Description
Impact on enrollment decision	Response to the following question: Will the prediction result and automatically generated explanations change your enrollment decision on the patient?
Usefulness of the prediction result	Response to the following question: How useful is the prediction result? Rating is on a 7-point Likert scale, ranging from "not at all" (1) to "very useful" (7).
Usefulness of the automatically generated explanations	Response to the following question: How useful are the automatically generated explanations? Rating is on a 7-point Likert scale, ranging from "not at all" (1) to "very useful" (7).
Trustworthiness of the prediction result	Response to the following question: In your opinion, how much clinical sense does the prediction result make? Rating is on a 7-point Likert scale, ranging from "not at all" (1) to "completely" (7).
Trustworthiness of the automatically generated explanations	Response to the following question: In your opinion, how much clinical sense do the automatically generated explanations make? Rating is on a 7-point Likert scale, ranging from "not at all" (1) to "completely" (7).

For Modeling Problem 8, medication order and refill information is needed for identifying asthma. The IH dataset contains this because IH has its own health insurance plan. If too much refill information is missed at IH, data from the all-payer claims database [109] will be used. For Modeling Problem 9, in our ongoing UWM operational project, we have used around 30 attributes and approximately 6000 patients to build a basic cost prediction model, which achieved an R^2 close to that of the commercial claims-based model. Since the health care researcher will use many more attributes and patients that should increase model accuracy, we expect the cost prediction model built by him/her to achieve a higher R^2 than the claims-based model.

Although using a constraint to fill the scope gap partially addresses UWM data's incompleteness, UWM still has incomplete medical data on some of its patients satisfying the constraint. For each such diabetic patient, the dependent variable of the patient's total allowed cost to UWM is only part of the patient's total allowed cost to all systems. The patient's features are computed from incomplete data. Both factors may create difficulty for significantly improving R^2. If this occurs, we will revise the dependent variable to a diabetic patient's total allowed cost to UWM or reflected by PreManage data. On average, the revised dependent variable is closer to the patient's total allowed cost to all systems than the original one. Recall that based on UWM patient demographics and PreManage diagnosis data, we will use two models to estimate the allowed cost of each of the patient's non-UWM inpatient stays and ED visits reflected by PreManage encounter data. We will supplement UWM data with PreManage data to make patient data more complete for computing patient features. This approach of using PreManage data and revising the dependent variable can be adopted to improve the accuracy of predicting future hospital use.

For either new modeling problem, if one health care researcher fails to build a reasonably accurate model, we will recruit another health care researcher.

Aim 3

Overview

Aim 3 involves performing simulations to estimate the impact of adopting Auto-ML on US patient outcomes. To determine Auto-ML's value for future clinical deployment, we will estimate the impact of adopting Auto-ML on US patient outcomes. Trials showed that machine learning helped drop the 30-day mortality rate in ED patients with community-acquired pneumonia (risk ratio≈OR=0.53, as the mortality rate is much less than 1) [2] and cut hospitalization days by 15% in end-stage renal disease patients on dialysis [3]. We will use these two scenarios to demonstrate our simulation method. Our method generalizes to other scenarios and similar software. We will use the same computing environment mentioned in Aim 1 (d). We first discuss the scenario of ED patients with community-acquired pneumonia.

Estimate Outcomes

The outcome is 30-day mortality. We will use the latest, deidentified, and publicly available Nationwide Emergency Department Sample (NEDS) database [110], including visit information from approximately 20% of US EDs. Consider the case with Auto-ML. The likelihood, L, that an ED can successfully use machine learning for this scenario is equal to $p_1 \times p_2$. p_1 is the probability that a health care researcher in the ED can build a high-quality machine learning model for this scenario using Auto-ML. p_2 is the probability that the ED can successfully deploy the model if it can be built. Using Aim 1(d)'s test results on whether health care researchers can use Auto-ML to achieve model accuracy of at least 95% of the gold standard, we will conservatively estimate p_1's minimum and maximum values (eg, by fitting a normal distribution and using its 2.5 and 97.5 percentile points). Based on his extensive experience with deploying models [2], Dr Haug (PJH) will conservatively estimate p_2's minimum and maximum values. For each of p_1 and p_2, we will adopt five levels going from the minimum to the maximum value for sensitivity analysis. The middle level is the default one and is used for hypothesis testing.

For each ED in the NEDS database, we will retrieve the annual number of patients with community-acquired pneumonia. We will simulate whether or not the ED can successfully use machine learning for this scenario based on the likelihood, L. If success/not success, for each ED patient with community-acquired pneumonia, we will simulate whether the patient will die or not based on the 30-day mortality rate reported in the paper [2] when using/not using machine learning. The overall outcome estimate combines the expected outcomes for all patients and EDs. The patients' discharge weights in the NEDS database will be used to obtain national estimates from sample data in the database. We will handle the case without Auto-ML similarly by simulating not using machine learning.

Outcome Evaluation and Sample Size Justification

Outcomes achieved with and without Auto-ML will be compared. We will test the primary hypothesis that using Auto-ML will be linked to reduced mortality. In the most conservative case assuming a proportion of discordant pairs of 10%, a sample size of 1152 patients provides 90% power to notice an OR of 0.53 [2] using a two-sided McNemar test at a significance level of .05. Each year, community-acquired pneumonia incurs 1.5 million ED patient visits [111], giving adequate power to test the hypothesis. To acquire the whole range of possible outcomes, we will do sensitivity analysis by changing the levels of the probabilities p_1 and p_2, 30-day mortality rate, and rate reduction gained by machine learning.

The scenario of end-stage renal disease patients on dialysis will be handled similarly, with the following main differences. The outcome is number of hospitalization days. The health care unit is dialysis facility. For each US dialysis facility, we will obtain its latest annual total number of hospitalization days and patient count from DialysisData.org [112] to fit a Poisson distribution. For each dialysis patient in the facility, we will simulate his/her annual number of hospitalization days using the distribution, as is often done in the literature [113]. We will test the secondary hypothesis that using Auto-ML will be linked to reduced hospitalization days. If the results from a single simulation run appear too skewed, we will conduct multiple runs and then average their results.

Ethics Approval

We have already acquired institutional review board approvals from UWM and IH for our study.

Results

Our paper [35] describes our draft method for automating machine learning model selection. The paper shows that compared to the modern Auto-WEKA automatic selection method [25], on six medical and 21 nonmedical benchmark datasets, our draft method reduced search time by 28-fold, classification error rate by 11%, and standard deviation of error rate due to randomization by 36%, on average. On each of these datasets, our draft method can finish the search process in 12 hours or less on a single computer. The results obtained on the medical datasets are similar to those obtained on the nonmedical datasets. The health care researchers in the Veterans Affairs Salt Lake City Health Care System have used our draft method successfully for a clinical research project [114]. One purpose of this study is to improve the draft method so that it can handle larger datasets more efficiently and effectively.

At present, we are writing Auto-ML's design document. We intend to finish this study by around the year 2022.

Discussion

Auto-ML will generalize to various clinical prediction/classification problems, as its design relies on no special property of a specific dataset, patient population, or disease. Auto-ML will be tested on nine modeling problems and datasets, each from a distinct health care system. By providing support for common data models (eg, OMOP [88]) and their linked standardized terminologies adopted by a large number of systems, Auto-ML can be used to construct models if attributes required to solve a problem are accessible in a structured dataset or in one of those common data models. This enables data integration and facilitates building models with data from multiple systems. To help users decide whether any data quality issues need to be handled before modeling, Auto-ML will show the numbers of attribute values outside reasonable ranges and numbers of missing values of nonrepeatedly recorded attributes.

The gaps in scope and accuracy mentioned in the introduction exist in many clinical applications. The principles of our approaches to using incomplete medical data and vast attributes generalize to many other clinical applications beyond the two on care management listed in the introduction.

In summary, our new software is designed to efficiently automate machine learning model selection and temporal aggregation of clinical attributes. By making machine learning feasible with limited budgets and data scientist resources, our new software will help realize value from clinical big data and improve patient outcomes. The models that will be built for the two new modeling problems will help improve care management outcomes.

Acknowledgments

We thank E Sally Lee, Xinran Liu, Xueqiang Zeng, and Nickolas Robison for helpful discussions.

Authors' Contributions

GL was mainly responsible for the paper; he conceptualized and designed the study, performed the literature review, and wrote the paper. BLS, MDJ, PTH, ABW, SDM, PJH, and FLN offered feedback on miscellaneous medical issues, contributed to conceptualizing the presentation, and revised the paper. XS took part in conceptualizing and writing the statistical analysis sections.

Conflicts of Interest

None declared.

References

1. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York, NY: Springer Science+Business Media; 2009.
2. Dean NC, Jones BE, Jones JP, Ferraro JP, Post HB, Aronsky D, et al. Impact of an electronic clinical decision support tool for emergency department patients with pneumonia. Ann Emerg Med 2015 Nov;66(5):511-520. [doi: 10.1016/j.annemergmed.2015.02.003] [Medline: 25725592]
3. Barbieri C, Molina M, Ponce P, Tothova M, Cattinelli I, Ion TJ, et al. An international observational study suggests that artificial intelligence for clinical decision support optimizes anemia management in hemodialysis patients. Kidney Int 2016 Aug;90(2):422-429. [doi: 10.1016/j.kint.2016.03.036] [Medline: 27262365]
4. Brier ME, Gaweda AE, Dailey A, Aronoff GR, Jacobs AA. Randomized trial of model predictive control for improved anemia management. Clin J Am Soc Nephrol 2010 May;5(5):814-820 [FREE Full text] [doi: 10.2215/CJN.07181009] [Medline: 20185598]
5. Gaweda AE, Aronoff GR, Jacobs AA, Rai SN, Brier ME. Individualized anemia management reduces hemoglobin variability in hemodialysis patients. J Am Soc Nephrol 2014 Jan;25(1):159-166 [FREE Full text] [doi: 10.1681/ASN.2013010089] [Medline: 24029429]
6. Gaweda AE, Jacobs AA, Aronoff GR, Brier ME. Model predictive control of erythropoietin administration in the anemia of ESRD. Am J Kidney Dis 2008 Jan;51(1):71-79. [doi: 10.1053/j.ajkd.2007.10.003] [Medline: 18155535]
7. Hsu J, Chen Y, Chung W, Tan T, Chen T, Chiang JY. Clinical verification of a clinical decision support system for ventilator weaning. Biomed Eng Online 2013;12 Suppl 1:S4 [FREE Full text] [doi: 10.1186/1475-925X-12-S1-S4] [Medline: 24565021]
8. Hamlet KS, Hobgood A, Hamar GB, Dobbs AC, Rula EY, Pope JE. Impact of predictive model-directed end-of-life counseling for Medicare beneficiaries. Am J Manag Care 2010 May;16(5):379-384 [FREE Full text] [Medline: 20469958]
9. CHIME. 2015 Mar 24. Jvion's latest predictive analytics in healthcare survey finds that advanced predictive modeling solutions are taking a strong foothold in the industry URL: https://chimecentral.org/jvion-releases-findings-latest-predictive-analytics-healthcare-survey/ [accessed 2017-02-18] [WebCite Cache ID 6oOiQpFqo]
10. Kuhn M, Johnson K. Applied Predictive Modeling. New York, NY: Springer Science+Business Media; 2013.
11. Axelrod RC, Vogel D. Predictive modeling in health plans. Dis Manag Health Outcomes 2003 Dec;11(12):779-787. [doi: 10.2165/00115677-200311120-00003]
12. Asadi H, Dowling R, Yan B, Mitchell P. Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. PLoS One 2014;9(2):e88225 [FREE Full text] [doi: 10.1371/journal.pone.0088225] [Medline: 24520356]
13. Kaggle. URL: https://www.kaggle.com/ [accessed 2017-02-19] [WebCite Cache ID 6oOiI4j9t]
14. Luo G, Stone BL, Sakaguchi F, Sheng X, Murtaugh MA. Using computational approaches to improve risk-stratified patient management: Rationale and methods. JMIR Res Protoc 2015;4(4):e128 [FREE Full text] [doi: 10.2196/resprot.5039] [Medline: 26503357]
15. Luo G. Automatically explaining machine learning prediction results: A demonstration on type 2 diabetes risk prediction. Health Inf Sci Syst 2016;4:2 [FREE Full text] [doi: 10.1186/s13755-016-0015-4] [Medline: 26958341]
16. Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. Big Data: The Next Frontier for Innovation, Competition, and Productivity.: McKinsey Global Institute; 2011 Jun. URL: http://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.ashx [accessed 2017-08-17] [WebCite Cache ID 6smpMfEfq]
17. Auffray C, Balling R, Barroso I, Bencze L, Benson M, Bergeron J, et al. Making sense of big data in health research: Towards an EU action plan. Genome Med 2016 Dec 23;8(1):71 [FREE Full text] [doi: 10.1186/s13073-016-0323-y] [Medline: 27338147]
18. Hoskins M. Common big data challenges and how to overcome them. Big Data 2014 Sep;2(3):142-143. [doi: 10.1089/big.2014.0030] [Medline: 27442494]
19. Witten I, Frank E, Hall M, Pal C. Data Mining: Practical Machine Learning Tools and Techniques. 4th edition. Burlington, MA: Morgan Kaufmann; Nov 17, 2016.
20. Rao RB, Fung G, Rosales R. On the dangers of cross-validation. An experimental evaluation. In: Proceedings of the SIAM International Conference on Data Mining. 2008 Presented at: SIAM International Conference on Data Mining; April 24-26, 2008; Atlanta, GA p. 588-596. [doi: 10.1137/1.9781611972788.54]
21. Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation. J Mach Learn Res 2010 Jul;11:2079-2107.
22. Reunanen J. Overfitting in making comparisons between variable selection methods. J Mach Learn Res 2003;3:1371-1382.
23. Dwork C, Feldman V, Hardt M, Pitassi T, Reingold O, Roth A. Generalization in adaptive data analysis and holdout reuse. In: Proceedings of the Annual Conference on Neural Information Processing Systems. 2015 Presented at: Annual Conference on Neural Information Processing Systems; December 7-12, 2015; Montreal, QC p. 2350-2358.
24. Luo G. PredicT-ML: A tool for automating machine learning model building with big clinical data. Health Inf Sci Syst 2016;4:5 [FREE Full text] [doi: 10.1186/s13755-016-0018-1] [Medline: 27280018]
25. Thornton C, Hutter F, Hoos H, Leyton-Brown K. Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and

Data Mining. 2013 Presented at: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 11-14, 2013; Chicago, IL p. 847-855. [doi: 10.1145/2487575.2487629]
26. Luo G. A review of automatic selection methods for machine learning algorithms and hyper-parameter values. Netw Model Anal Health Inform Bioinform 2016;5:18. [doi: 10.1007/s13721-016-0125-6]
27. Snoek J, Larochelle H, Adams R. Practical Bayesian optimization of machine learning algorithms. In: Proceedings of the Annual Conference on Neural Information Processing Systems. 2012 Presented at: Annual Conference on Neural Information Processing Systems; December 3-6, 2012; Lake Tahoe, NV p. 2960-2968.
28. Komer B, Bergstra J, Eliasmith C. Hyperopt-sklearn: Automatic hyperparameter configuration for scikit-learn. In: Proceedings of the Scientific Computing with Python Conference. 2014 Presented at: Scientific Computing with Python Conference; July 6-12, 2014; Austin, TX p. 33-39.
29. Kotthoff L, Thornton C, Hoos H, Hutter F, Leyton-Brown K. Auto-WEKA 2.0: Automatic model selection and hyperparameter optimization in WEKA. J Mach Learn Res 2017;18(25):1-5.
30. Salvador M, Budka M, Gabrys B. Towards automatic composition of multicomponent predictive systems. In: Proceedings of the International Conference on Hybrid Artificial Intelligence Systems. 2016 Presented at: International Conference on Hybrid Artificial Intelligence Systems; April 18-20, 2016; Seville, Spain p. 27-39. [doi: 10.1007/978-3-319-32034-2_3]
31. Salvador MM, Budka M, Gabrys B. arXiv. 2016 Dec 28. Automatic composition and optimisation of multicomponent predictive systems URL: https://arxiv.org/abs/1612.08789 [accessed 2017-08-13] [WebCite Cache ID 6shBDGw9I]
32. Kraska T, Talwalkar A, Duchi J, Griffith R, Franklin M, Jordan M. MLbase: S distributed machine-learning system. In: Proceedings of the Biennial Conference on Innovative Data Systems Research. 2013 Presented at: Biennial Conference on Innovative Data Systems Research; January 6-9, 2013; Asilomar, CA.
33. Sparks E, Talwalkar A, Haas D, Franklin M, Jordan M, Kraska T. Automating model search for large-scale machine learning. In: Proceedings of the ACM Symposium on Cloud Computing. 2015 Presented at: ACM Symposium on Cloud Computing; August 27-29, 2015; Kohala Coast, HI p. 368-380. [doi: 10.1145/2806777.2806945]
34. Kaggle. Practice Fusion diabetes classification URL: https://www.kaggle.com/c/pf2012-diabetes [accessed 2017-02-18] [WebCite Cache ID 6oOiaVvU5]
35. Zeng X, Luo G. Progressive Sampling-Based Bayesian Optimization for Efficient and Automatic Machine Learning Model Selection. URL: http://pages.cs.wisc.edu/~gangluo/progressive_sampling.pdf [accessed 2017-07-14] [WebCite Cache ID 6rz22eiKY]
36. Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2015 Presented at: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 10-13, 2015; Sydney, Australia p. 1721-1730. [doi: 10.1145/2783258.2788613]
37. Wiens J, Guttag J, Horvitz E. A study in transfer learning: Leveraging data from multiple hospitals to enhance hospital-specific predictions. J Am Med Inform Assoc 2014;21(4):699-706 [FREE Full text] [doi: 10.1136/amiajnl-2013-002162] [Medline: 24481703]
38. Borsi JP. Hypothesis-free search for connections between birth month and disease prevalence in large, geographically varied cohorts. AMIA Annu Symp Proc 2017 Feb 10;2016:319-325 [FREE Full text] [Medline: 28269826]
39. Wilcox A, Hripcsak G. Medical text representations for inductive learning. Proc AMIA Symp 2000:923-927 [FREE Full text] [Medline: 11080019]
40. Hickins M. Forbes. 2016 Oct 03. Citizen data scientists unite! URL: http://www.forbes.com/sites/oracle/2016/10/03/citizen-data-scientists-unite [accessed 2017-02-19] [WebCite Cache ID 6oOjBNVtO]
41. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, et al. Multiple chronic conditions: Prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007 Dec;22 Suppl 3:391-395 [FREE Full text] [doi: 10.1007/s11606-007-0322-1] [Medline: 18026807]
42. Caloyeras JP, Liu H, Exum E, Broderick M, Mattke S. Managing manifest diseases, but not health risks, saved PepsiCo money over seven years. Health Aff (Millwood) 2014 Jan;33(1):124-131. [doi: 10.1377/hlthaff.2013.0625] [Medline: 24395944]
43. Nelson L. Lessons From Medicare's Demonstration Projects on Disease Management and Care Coordination. Washington, DC: Congressional Budget Office; 2012 Jan. URL: https://www.cbo.gov/sites/default/files/112th-congress-2011-2012/workingpaper/WP2012-01_Nelson_Medicare_DMCC_Demonstrations_1.pdf [accessed 2015-08-09] [WebCite Cache ID 6agZcFxD1]
44. Centers for Disease Control and Prevention. Asthma URL: http://www.cdc.gov/nchs/fastats/asthma.htm [accessed 2015-08-09] [WebCite Cache ID 6agaQMYxr]
45. Akinbami LJ, Moorman JE, Liu X. Asthma prevalence, health care use, and mortality: United States, 2005-2009. Natl Health Stat Report 2011 Jan 12(32):1-14 [FREE Full text] [Medline: 21355352]
46. Akinbami LJ, Moorman JE, Bailey C, Zahran HS, King M, Johnson CA, et al. Trends in asthma prevalence, health care use, and mortality in the United States, 2001-2010. NCHS Data Brief 2012 May(94):1-8 [FREE Full text] [Medline: 22617340]

47. Centers for Disease Control and Prevention. 2011 May. Asthma in the US URL: https://www.cdc.gov/vitalsigns/asthma/ [accessed 2017-02-18] [WebCite Cache ID 6oOjKVf75]
48. Levine SH, Adams J, Attaway K, Dorr DA, Leung M, Popescu P, et al. Predicting the Financial Risks of Seriously Ill Patients. Oakland, CA: California Health Care Foundation; 2011 Dec. URL: http://www.chcf.org/resources/download.aspx?id={CD2B7755-9A80-4A0F-A213-E7489CE25436} [accessed 2017-02-19] [WebCite Cache ID 6sn1rEMak]
49. Rubin RJ, Dietrich KA, Hawk AD. Clinical and economic impact of implementing a comprehensive diabetes management program in managed care. J Clin Endocrinol Metab 1998 Aug;83(8):2635-2642. [doi: 10.1210/jcem.83.8.5075] [Medline: 9709924]
50. Greineder DK, Loane KC, Parks P. A randomized controlled trial of a pediatric asthma outreach program. J Allergy Clin Immunol 1999 Mar;103(3 Pt 1):436-440. [Medline: 10069877]
51. Kelly CS, Morrow AL, Shults J, Nakas N, Strope GL, Adelman RD. Outcomes evaluation of a comprehensive intervention program for asthmatic children enrolled in Medicaid. Pediatrics 2000 May;105(5):1029-1035. [Medline: 10790458]
52. Axelrod R, Zimbro K, Chetney R, Sabol J, Ainsworth V. A disease management program utilizing life coaches for children with asthma. J Clin Outcomes Manag 2001;8(6):38-42.
53. Beaulieu N, Cutler D, Ho K, Isham G, Lindquist T, Nelson A, et al. The business case for diabetes disease management for managed care organizations. Forum Health Econ Policy 2006;9(1):1-37.
54. Dorr DA, Wilcox AB, Brunker CP, Burdon RE, Donnelly SM. The effect of technology-supported, multidisease care management on the mortality and hospitalization of seniors. J Am Geriatr Soc 2008 Dec;56(12):2195-2202. [doi: 10.1111/j.1532-5415.2008.02005.x] [Medline: 19093919]
55. Forno E, Celedón JC. Predicting asthma exacerbations in children. Curr Opin Pulm Med 2012 Jan;18(1):63-69 [FREE Full text] [doi: 10.1097/MCP.0b013e32834db288] [Medline: 22081091]
56. Mays GP, Claxton G, White J. Managed care rebound? Recent changes in health plans' cost containment strategies. Health Aff (Millwood) 2004;Suppl Web Exclusives:W4-427-W4-436 [FREE Full text] [doi: 10.1377/hlthaff.w4.427] [Medline: 15451964]
57. Curry N, Billings J, Darin B, Dixon J, Williams M, Wennberg D. Predictive Risk Project: Literature Review.: National Health Service; 2005 Jun. URL: http://www.kingsfund.org.uk/sites/files/kf/field/field_document/predictive-risk-literature-review-june2005.pdf [accessed 2015-08-10] [WebCite Cache ID 6aga2XBFC]
58. Bourgeois FC, Olson KL, Mandl KD. Patients treated at multiple acute health care facilities: Quantifying information fragmentation. Arch Intern Med 2010 Dec 13;170(22):1989-1995. [doi: 10.1001/archinternmed.2010.439] [Medline: 21149756]
59. Finnell JT, Overhage JM, Grannis S. All health care is not local: An evaluation of the distribution of emergency department care delivered in Indiana. AMIA Annu Symp Proc 2011;2011:409-416 [FREE Full text] [Medline: 22195094]
60. Duncan I. Healthcare Risk Adjustment and Predictive Modeling. Winsted, CT: ACTEX Publications, Inc; 2011.
61. Ash AS, McCall N. Risk Assessment of Military Populations to Predict Health Care Cost and Utilization. Washington, DC: RTI International; 2005 Sep. URL: http://www.rti.org/pubs/tricare_riskassessment_final_report_combined.pdf [accessed 2015-08-09] [WebCite Cache ID 6aga7wYZC]
62. Iezzoni LI, editor. Risk Adjustment for Measuring Healthcare Outcomes. 4th edition. Chicago, IL: Health Administration Press; 2012.
63. Schatz M, Cook EF, Joshua A, Petitti D. Risk factors for asthma hospitalizations in a managed care organization: Development of a clinical prediction rule. Am J Manag Care 2003 Aug;9(8):538-547 [FREE Full text] [Medline: 12921231]
64. Lieu TA, Capra AM, Quesenberry CP, Mendoza GR, Mazar M. Computer-based models to identify high-risk adults with asthma: Is the glass half empty of half full? J Asthma 1999 Jun;36(4):359-370. [Medline: 10386500]
65. Schatz M, Nakahiro R, Jones CH, Roth RM, Joshua A, Petitti D. Asthma population management: Development and validation of a practical 3-level risk stratification scheme. Am J Manag Care 2004 Jan;10(1):25-32 [FREE Full text] [Medline: 14738184]
66. Lieu TA, Quesenberry CP, Sorel ME, Mendoza GR, Leong AB. Computer-based models to identify high-risk children with asthma. Am J Respir Crit Care Med 1998 Apr;157(4 Pt 1):1173-1180. [doi: 10.1164/ajrccm.157.4.9708124] [Medline: 9563736]
67. Forno E, Fuhlbrigge A, Soto-Quirós ME, Avila L, Raby BA, Brehm J, et al. Risk factors and predictive clinical scores for asthma exacerbations in childhood. Chest 2010 Nov;138(5):1156-1165 [FREE Full text] [doi: 10.1378/chest.09-2426] [Medline: 20472862]
68. Miller MK, Lee JH, Blanc PD, Pasta DJ, Gujrathi S, Barron H, et al. TENOR risk score predicts healthcare in adults with severe or difficult-to-treat asthma. Eur Respir J 2006 Dec;28(6):1145-1155 [FREE Full text] [doi: 10.1183/09031936.06.00145105] [Medline: 16870656]
69. Schatz M. Predictors of asthma control: What can we modify? Curr Opin Allergy Clin Immunol 2012 Jun;12(3):263-268. [doi: 10.1097/ACI.0b013e32835335ac] [Medline: 22517290]
70. Stanford RH, Shah MB, D'Souza AO, Schatz M. Predicting asthma outcomes in commercially insured and Medicaid populations. Am J Manag Care 2013 Jan;19(1):60-67 [FREE Full text] [Medline: 23379745]

71. Hyland ME, Whalley B, Halpin DMG, Greaves CJ, Seamark C, Blake S, et al. Frequency of non-asthma GP visits predicts asthma exacerbations: An observational study in general practice. Prim Care Respir J 2012 Dec;21(4):405-411 [FREE Full text] [doi: 10.4104/pcrj.2012.00061] [Medline: 22836742]
72. Crawford AG, Fuhr JP, Clarke J, Hubbs B. Comparative effectiveness of total population versus disease-specific neural network models in predicting medical costs. Dis Manag 2005 Oct;8(5):277-287. [doi: 10.1089/dis.2005.8.277] [Medline: 16212513]
73. Coyle YM. Predictors of acute asthma relapse: Strategies for its prevention. J Asthma 2003 May;40(3):217-224. [Medline: 12807164]
74. Evans RS. Electronic health records: Then, now, and in the future. Yearb Med Inform 2016 May 20;Suppl 1:S48-S61. [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]
75. Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis. 2nd edition. Hoboken, NJ: John Wiley & Sons; 2011.
76. Dean J, Ghemawat S. MapReduce: Simplified data processing on large clusters. In: Proceedings of the Symposium on Operating System Design and Implementation. 2004 Presented at: Symposium on Operating System Design and Implementation; December 6-8, 2004; San Francisco, CA p. 137-150.
77. White T. Hadoop: The Definitive Guide. 4th edition. Sebastopol, CA: O'Reilly Media; 2015.
78. Karau H, Konwinski A, Wendell P, Zaharia M. Learning Spark: Lightning-Fast Big Data Analysis. Sebastopol, CA: O'Reilly Media; 2015.
79. Meng X, Bradley J, Yuvaz B, Sparks E, Venkataraman S, Liu D, et al. MLlib: Machine learning in Apache Spark. J Mach Learn Res 2016;17(34):1-7.
80. Xin R, Rosen J, Zaharia M, Franklin M, Shenker S, Stoica I. Shark: SQL and rich analytics at scale. In: Proceedings of the ACM SIGMOD International Conference on Management of Data. 2013 Presented at: ACM SIGMOD International Conference on Management of Data; June 22-27, 2013; New York, NY p. 13-24. [doi: 10.1145/2463676.2465288]
81. Department of Computer Science, The University of Waikato. Mining big data using Weka 3 URL: http://www.cs.waikato.ac.nz/ml/weka/bigdata.html [accessed 2017-02-19] [WebCite Cache ID 6oOkT63lx]
82. Feurer M, Klein A, Eggensperger K, Springenberg J, Blum M, Hutter F. Efficient and robust automated machine learning. In: Proceedings of the Annual Conference on Neural Information Processing Systems. 2015 Presented at: Annual Conference on Neural Information Processing Systems; December 7-12, 2015; Montreal, QC p. 2944-2952.
83. Luo G. MLBCD: A machine learning tool for big clinical data. Health Inf Sci Syst 2015;3:3 [FREE Full text] [doi: 10.1186/s13755-015-0011-0] [Medline: 26417431]
84. Provost F, Jensen D, Oates T. Efficient progressive sampling. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 1999 Presented at: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 15-18, 1999; San Diego, CA p. 23-32. [doi: 10.1145/312129.312188]
85. Wolpert DH. The lack of a priori distinctions between learning algorithms. Neural Comput 1996 Oct;8(7):1341-1390. [doi: 10.1162/neco.1996.8.7.1341]
86. Zhang Y, Bahadori M, Su H, Sun J. FLASH: Fast Bayesian optimization for data analytic pipelines. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 Presented at: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, CA p. 2065-2074. [doi: 10.1145/2939672.2939829]
87. Krueger T, Panknin D, Braun M. Fast cross-validation via sequential testing. J Mach Learn Res 2015;16(1):1103-1155.
88. Observational Medical Outcomes Partnership (OMOP). Common data model URL: http://omop.org/CDM [accessed 2015-08-09] [WebCite Cache ID 6agamjByZ]
89. Observational Medical Outcomes Partnership (OMOP). Vocabularies URL: http://omop.org/Vocabularies [accessed 2017-02-19] [WebCite Cache ID 6oOkOnGg5]
90. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. J Am Med Inform Assoc 2012;19(1):54-60 [FREE Full text] [doi: 10.1136/amiajnl-2011-000376] [Medline: 22037893]
91. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for observational researchers. Stud Health Technol Inform 2015;216:574-578 [FREE Full text] [Medline: 26262116]
92. The National Patient-Centered Clinical Research Network (PCORnet). PCORnet Common Data Model (CDM) URL: http://www.pcornet.org/pcornet-common-data-model/ [accessed 2017-02-19] [WebCite Cache ID 6oOkB07wd]
93. i2b2 Cell Messaging Data Repository (CRC) Cell. Boston, MA: Partners HealthCare URL: https://www.i2b2.org/software/files/PDF/current/CRC_Design.pdf [accessed 2017-02-19] [WebCite Cache ID 6oOkHsUCx]
94. Patton MQ. Qualitative Research & Evaluation Methods. 3rd edition. Thousand Oaks, CA: Sage Publications; 2002.
95. University of California, Irvine, Machine Learning Repository. URL: http://archive.ics.uci.edu/ml/ [accessed 2017-02-19] [WebCite Cache ID 6oOkdrOjg]
96. Brooke J. SUS - A Quick and Dirty Usability Scale. Reading, UK: Digital Equipment Co Ltd; 1996. URL: http://hell.meiert.org/core/pdf/sus.pdf [accessed 2017-02-19] [WebCite Cache ID 6oOmoIihR]

97. Tullis T, Albert W. Measuring the User Experience: Collecting, Analyzing and Presenting Usability Metrics. 2nd edition. Waltham, MA: Morgan Kaufmann; 2013.
98. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006 Jun 01;27(2):237-246. [doi: 10.1177/1098214005283748]
99. ATLAS.ti. URL: http://atlasti.com/?q=%2Findex.html [accessed 2017-02-19] [WebCite Cache ID 6oOjqY7rj]
100. Desai JR, Wu P, Nichols GA, Lieu TA, O'Connor PJ. Diabetes and asthma case identification, validation, and representativeness when using electronic health data to construct registries for comparative effectiveness and epidemiologic research. Med Care 2012 Jul;50 Suppl:S30-S35. [doi: 10.1097/MLR.0b013e318259c011] [Medline: 22692256]
101. Wakefield DB, Cloutier MM. Modifications to HEDIS and CSTE algorithms improve case recognition of pediatric asthma. Pediatr Pulmonol 2006 Oct;41(10):962-971. [doi: 10.1002/ppul.20476] [Medline: 16871628]
102. James BC, Savitz LA. How Intermountain trimmed health care costs through robust quality improvement efforts. Health Aff (Millwood) 2011 Jun;30(6):1185-1191 [FREE Full text] [doi: 10.1377/hlthaff.2011.0358] [Medline: 21596758]
103. Byington CL, Reynolds CC, Korgenski K, Sheng X, Valentine KJ, Nelson RE, et al. Costs and infant outcomes after implementation of a care process model for febrile infants. Pediatrics 2012 Jul;130(1):e16-e24 [FREE Full text] [doi: 10.1542/peds.2012-0127] [Medline: 22732178]
104. Neuvirth H, Ozery-Flato M, Hu J, Laserson J, Kohn M, Ebadollahi S, et al. Toward personalized care management of patients at risk: The diabetes case study. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2011 Presented at: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 21-24, 2011; San Diego, CA p. 395-403. [doi: 10.1145/2020408.2020472]
105. Bureau of Labor Statistics. 2010 Apr 12. Consumer Price Index: Measuring price change for medical care in the CPI URL: http://www.bls.gov/cpi/cpifact4.htm [accessed 2015-08-10] [WebCite Cache ID 6agaajPCS]
106. Collective Medical Technologies. PreManage URL: http://collectivemedicaltech.com/what-we-do-2/premanage/ [accessed 2017-02-19] [WebCite Cache ID 6oOi8en1e]
107. Sun J, Hu J, Luo D, Markatou M, Wang F, Edabollahi S, et al. Combining knowledge and data-driven insights for identifying risk factors using electronic health records. AMIA Annu Symp Proc 2012;2012:901-910 [FREE Full text] [Medline: 23304365]
108. Selby JV, Karter AJ, Ackerson LM, Ferrara A, Liu J. Developing a prediction rule from automated clinical databases to identify high-risk patients in a large population with diabetes. Diabetes Care 2001 Sep;24(9):1547-1555. [Medline: 11522697]
109. The APCD (All-Payer Claims Database) Council. URL: http://www.apcdcouncil.org/ [accessed 2015-08-10] [WebCite Cache ID 6agaeGTXm]
110. Healthcare Cost and Utilization Project (HCUP). Rockville, MD: Agency for Healthcare Research and Quality; 2017 Mar. NEDS database documentation URL: https://www.hcup-us.ahrq.gov/db/nation/neds/nedsdbdocumentation.jsp [accessed 2017-02-19] [WebCite Cache ID 6oOi2Tqjz]
111. Kyriacou DN, Yarnold PR, Soltysik RC, Self WH, Wunderink RG, Schmitt BP, et al. Derivation of a triage algorithm for chest radiography of community-acquired pneumonia patients in the emergency department. Acad Emerg Med 2008 Jan;15(1):40-44 [FREE Full text] [doi: 10.1111/j.1553-2712.2007.00011.x] [Medline: 18211312]
112. DialysisData. URL: https://www.dialysisdata.org/ [accessed 2017-02-19] [WebCite Cache ID 6oOhv8xXR]
113. Arora P, Kausz AT, Obrador GT, Ruthazer R, Khan S, Jenuleson CS, et al. Hospital utilization among chronic dialysis patients. J Am Soc Nephrol 2000 Apr;11(4):740-746 [FREE Full text] [Medline: 10752533]
114. Divita G, Luo G, Tran L, Workman T, Gundlapalli A, Samore M. General symptom extraction from VA electronic medical notes. Stud Health Technol Inform 2017 (forthcoming).

Abbreviations

AUC: area under the receiver operating characteristic curve
Auto-ML: Automated Machine Learning
ED: emergency department
EDW: enterprise data warehouse
FEV1: forced expiratory volume in 1 second
HIPAA: Health Insurance Portability and Accountability Act
i2b2: Informatics for Integrating Biology and the Bedside
ICD-9: International Classification of Diseases, Ninth Revision
ICD-10: International Classification of Diseases, Tenth Revision
IH: Intermountain Healthcare
NEDS: Nationwide Emergency Department Sample
OMOP: Observational Medical Outcomes Partnership
OR: odds ratio
PCORnet: National Patient-Centered Clinical Research Network
SQL: Structured Query Language

SUS: System Usability Scale
UWM: University of Washington Medicine
Weka: Waikato Environment for Knowledge Analysis

Edited by G Eysenbach; submitted 26.03.17; peer-reviewed by L Zhang, L Zhou; comments to author 12.07.17; revised version received 14.07.17; accepted 15.07.17; published 29.08.17.

Please cite as:
Luo G, Stone BL, Johnson MD, Tarczy-Hornoch P, Wilcox AB, Mooney SD, Sheng X, Haug PJ, Nkoy FL
Automating Construction of Machine Learning Models With Clinical Big Data: Proposal Rationale and Methods
JMIR Res Protoc 2017;6(8):e175
URL: http://www.researchprotocols.org/2017/8/e175/
doi:10.2196/resprot.7757
PMID:28851678

Tutorial

Opening the Black Box of Electronic Health: Collecting, Analyzing, and Interpreting Log Data

Floor Sieverink[1], MSc, MA; Saskia Kelders[1,2], PhD; Mannes Poel[3], PhD; Lisette van Gemert-Pijnen[1], PhD

[1]Centre for eHealth and Wellbeing Research, Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

[2]Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa

[3]Human Media Interaction group, Department of Computer Science, University of Twente, Enschede, Netherlands

Corresponding Author:
Floor Sieverink, MSc, MA
Centre for eHealth and Wellbeing Research
Department of Psychology, Health and Technology
University of Twente
P.O. Box 217
Enschede, 7500 AE
Netherlands
Phone: 31 628554510
Fax: 31 534892895
Email: f.sieverink@utwente.nl

Related Article:

This is a corrected version. See correction statement: http://www.researchprotocols.org/2018/1/e10/

Abstract

In electronic health (eHealth) research, limited insight has been obtained on process outcomes or how the use of technology has contributed to the users' ability to have a healthier life, improved well-being, or activate new attitudes in their daily tasks. As a result, eHealth is often perceived as a black box. To open this black box of eHealth, methodologies must extend beyond the classic effect evaluations. The analyses of log data (anonymous records of real-time actions performed by each user) can provide continuous and objective insights into the actual usage of the technology. However, the possibilities of log data in eHealth research have not been exploited to their fullest extent. The aim of this paper is to describe how log data can be used to improve the evaluation and understand the use of eHealth technology with a broader approach than only descriptive statistics. This paper serves as a starting point for using log data analysis in eHealth research. Here, we describe what log data is and provide an overview of research questions to evaluate the system, the context, the users of a technology, as well as the underpinning theoretical constructs. We also explain the requirements for log data, the starting points for the data preparation, and methods for data collection. Finally, we describe methods for data analysis and draw a conclusion regarding the importance of the results for both scientific and practical applications. The analysis of log data can be of great value for opening the black box of eHealth. A deliberate log data analysis can give new insights into how the usage of the technology contributes to found effects and can thereby help to improve the persuasiveness and effectiveness of eHealth technology and the underpinning behavioral models.

(JMIR Res Protoc 2017;6(8):e156) doi:10.2196/resprot.6452

KEYWORDS

eHealth; black box; evaluation; log data analysis

Introduction

Many electronic health (eHealth) technologies, such as behavior change technologies, aim to support users in reaching certain health-related behavioral outcomes. While such technologies can be effective [1,2], one of the main problems is that users' adoption and long-term use remains lower than expected [3-5]. Moreover, eHealth research is dominated by a classic conception of medical research where randomized controlled trials (RCTs) are the golden standard for measuring outcomes [6]. Although RCTs provide valuable insight into the effectiveness of an intervention, fundamental to this methodology is to have the

technology as a fixed entity for all participants throughout the entire intervention period. In contrast, (eHealth) technology can be characterized by its constant evolution and, consequently, apps or interventions often become obsolete by the time the results of the RCT are available.

Furthermore, to conform to the complexity of behavior change, eHealth technologies often consist of multiple components that may interact in reaching a certain effect and that people can use in many different ways in terms of the elements they use as well as the frequency, time, and place of use [7,8]. However, RCTs only provide insight into outcomes at fixed time points and treat technologies as a singular entity. Therefore, no insight can be obtained on process outcomes or how the use of the different components of the technology has contributed to healthier living, improved well-being, or a user's ability to conduct daily tasks [7,9,10]. This particular lack of insight is known as the "Black Box Phenomenon" [2,10,11]. To open the black box of eHealth and to investigate why, how, and for whom a certain technology is of the most value, methodologies must extend beyond the classic evaluations of effect only. In other words, the characteristics of eHealth technology change the way evaluations are conducted. In this view, Hekler and colleagues pled for an "agile science" approach that enables early and frequent insight into the process of behavior change via technology [7].

The CeHRes Roadmap (Figure 1) adopts an agile approach in the development and evaluation process of eHealth technology. This roadmap is based on an extensive literature review of eHealth frameworks and follows a holistic and participatory research and development approach. The following phases can be distinguished in the development and evaluation of eHealth: (1) contextual inquiry, (2) value specification, (3) design, (4) operationalization, and (5) summative evaluation. The results of each phase should be the subject of formative evaluation in order to collect input for improving the product [4].

According to the roadmap, technology development and evaluation is an iterative, flexible, and dynamic process without a fixed endpoint. In this approach, continuous (formative and summative) evaluation is needed that is interwoven with all stages of technology development. The outcomes of such evaluations will be used for analyzing the process, recognizing the areas of improvement, and diving deeper into the usage (the dose) that is needed to reach certain effects (the response). Thus, technology already in early stages of development will be reshaped by its usage. In order to do so, more advanced methods are needed to understand what people do with eHealth technology and how this is related to the impact.

The analysis of log data, defined as anonymous records of real-time actions performed by each user, has the potential to provide continuous and objective insight into the actual usage (of the different components) of the technology. Such analyses are a promising approach to explain the outcomes of the more traditional methods, such as RCTs, by gaining insight into the mediating mechanisms that contribute to the found effects [8,9] and have also the potential to identify unexpected effects of a technology. Thus, the use of log data fits the aim of eHealth evaluation according to the CeHRes Roadmap by enabling early improvements of the technology in order to improve the evaluation outcomes.

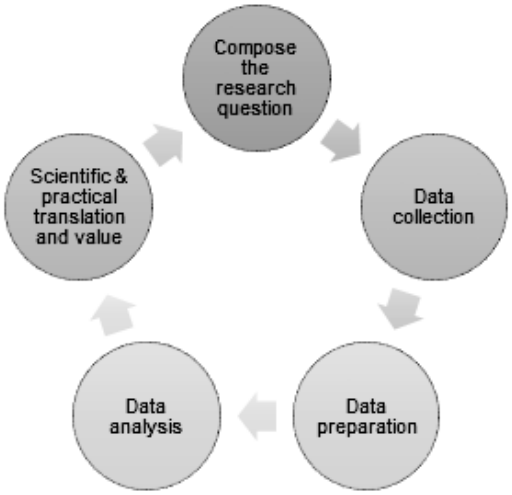
Log data analyses are currently used in diversified domains, such as education [12,13], human-computer interaction [14,15], and network security [16] where they are mainly used to analyze system performance and acceptance based on models for information retrieval. However, technology has evolved and the aim of behavior change technologies is not only to provide information, but rather to stimulate and support people in their process of behavior change [17]. Therefore, information is needed regarding how the use of technology can explain the engagement and involvement of the individual user, the found effects, and how the technology fits the user and the context.

Log data analyses in eHealth research have mainly focused on descriptive statistics, such as the number of logins, time spent, and the frequency of use of the different elements by all users as a group [18-20]. Although these statistics do provide valuable information regarding the usage of the technology, they also assume a dose-response relationship without taking the goal of the user into account. Furthermore, such analyses do not always provide insight into the actual process of technology use in relation to behavior change. For example, a longer and more frequent exposure to an eHealth technology might indicate how well the system fits the users' needs, but it can also signify unfocused and/or non-strategic use or an inefficient system [9,21]. The same applies to only counting the number of logins, since a user can log out directly without using the technology. As such, measures are needed that indicate how the actual use of the (different elements of the) technology can explain the found effects.

There is evidence that the use of log data can be of added value to the more traditional approaches of eHealth evaluation, but its possibilities have not yet been exploited to their fullest extent or have mostly been described on a conceptual level [7-9]. The aim of this paper is to give a more practical description of how log data can be of added value to understand the use of an eHealth technology with a broader approach than only descriptive statistics. Here, we describe what research questions can be answered using log data. Furthermore, we provide direction on which steps need to be taken in the collection, preparation, and analysis of log data, and how to interpret and apply the results (Figure 1).

The results and ideas presented in this paper are substantiated with examples from prior log data research. They do not provide an exhaustive overview of all research conducted in this field but are used to illustrate the possibilities of log data as a starting point for further research to open the black box of eHealth.

Figure 1. Steps in log data analysis.



Methods

What is Log Data?

Generally (transactional) log data can be seen as behavioral logs that contain information regarding the interactions between a system and the users of that system [9,14,15]. These interactions can include information regarding time of the action, content that is viewed or used, mouse clicks, browsing patterns, or saving information in the system. An important benefit of log data is that it represents the actual and continuous usage behavior and not subjective, recalled behavior.

The use of log data mainly focuses on the interaction with Web-based technologies. However, eHealth technology has evolved from Web- or telephone-based systems that required users' active involvement into unobtrusive and pervasive systems that are embedded into users' daily lives. For example, many people currently use lifestyle apps, such as "Runkeeper" or food diaries, to support their daily routines. Wearable devices like the Fitbit, Jawbone, or Apple Watch can continuously collect real-time health-related data for personalized coaching via apps for mobile phones or tablets. To be able to understand as much as possible of what people do, log data should not be limited to technologies and/or how they experience it, but to the actual usage of Web-based technologies including interactions with mobile phone apps and wearables as well.

Composing Research Questions

Before the log data can be collected and analyzed, an important step is to revisit the goals of the technology and the subsequent research questions. A variety of research questions can be addressed with log data analysis, depending on the type and the goal of the eHealth technology and on the phase of development. According to the CeHRes Roadmap, log data analyses can be of added value in both the formative and summative evaluation phases [4].

Formative evaluation is conducted within and between the different phases of the roadmap. The aim of this type of evaluation is to check whether the goals of that phase have been reached. In the early operationalization phase, for example, log data has the potential to evaluate the use of the system and to assess what people do (or not do) with a technology. Critical moments for dropping out can be identified, as well as profiles for both users and usage. The results from these evaluations can be used to improve an early release of the technology before it will be available for a bigger group of users, which fits the "agile science" approach [7]. Possible research questions for formative system evaluations are shown in Textbox 1.

Textbox 1. Questions for formative system evaluations.

Questions

- What usage patterns emerge when users navigate through the technology?
- Which (combinations of) elements of the technology are used?
- When do users drop out?
- How do users respond to behavior change strategies and persuasive triggers (strategies to support users in performing certain [usage] behaviors and/or long-term use, such as reminders)?

For example, in a previous study we found that users of a personal health record (PHR), an electronic application consisting of different tools for monitoring and coaching patients with chronic conditions to support self-management [22], followed the global menu structure when exploring the PHR for the first time. Furthermore, most users were likely to drop out when they visited the education section as a first step after the first login [23].

The uptake and impact of a technology are measured during the summative evaluation phase. Impact refers to whether the

intended goals of the technology have been realized in terms of behavioral, clinical, and organizational outcomes. Uptake refers to the implementation and usage of the technology. Log data can be used to assess the uptake of the technology. Where log data analyses in the operationalization phase mainly focus on the system performance, in this phase research questions to evaluate the whole of the system, the user, and the context can be formulated. Possible research questions to evaluate the system can be seen in Textbox 2.

Textbox 2. Questions to evaluate the system.

Questions

- How do users use the technology in order to complete an intervention or to achieve their health-related goals (in terms of frequency or combinations of elements they use)?
- How well do the users adhere to the intended usage of the technology?
  - What are predictors for adherence or dropping-out?
  - How does the use of the technology change over time?
- How did these usage patterns contribute to the (clinical, behavioral and organizational) impact? In other words, what is the dose-response relationship?

For example, studies by Kelders et al [21] and Van Gemert et al [24] showed when users of a Web-based intervention for the early treatment of depressive symptoms ("Living to the Full") were at risk of dropping out and might need additional support. In another example, Freyne et al [25] found that uploading a profile picture on a diet support site in the first week resulted in higher return rates.

Research questions to evaluate the users of the technology and possible context-related questions are shown in Textbox 3.

Textbox 3. User evaluation and context-related questions.

Questions

- User evaluation

  Who is motivated and capable of using the eHealth technology?
  - Who are the long-term users?
  - Who are the drop-outs?
- Context-related evaluation
  - How does the responsiveness of caregivers (eg, time until replying to a users' message) influence the use of eHealth technology by patients?
  - How does the technology integrate into users' daily lives?

Log data analyses can also provide answers to more fundamental research questions related to existing models and theories [8]. Behavior change theories and behavior models often form the basis for the content and the structure of eHealth technologies. For example, mental health interventions are frequently based on the principles of cognitive behavioral therapy and principles from the goal setting theory are used to support users in reaching their health-related goals. Log data can be used to check whether the incorporated (combination of) elements that represent certain theoretical concepts (eg, a chat box to facilitate social support) have been used. Or, when one of the goals of a technology is to improve self-efficacy, for example, log data can provide insight into what the most effective usage patterns are and for whom to experience any improvements in the self-efficacy. Research questions related to the evaluation of existing models and theories are shown in Textbox 4.

Textbox 4. Research questions to evaluate existing models and theories.

Questions

- To what extent did the users find and use the (combination of) elements of the technology that represent certain theoretical concepts?
- How did the use of (a combination of) these elements contributed to any improvements in the outcomes?

For example, the "Living to the Full" intervention is based on the principles of acceptance and commitment therapy (ACT). An effect evaluation indicated that a Web-based technology based on ACT might help users reduce depressive complaints [26]. However, a log data analysis revealed that many users of that intervention did not open the mindfulness exercises that are assumed to be an important element of ACT [24]. This might be an indicator that the concept of mindfulness is insufficiently operationalized in the intervention and that the found effects are an underestimation of the attainable effects.

While not complete or exhaustive, this overview serves as a starting point for composing suitable research questions for a holistic and agile evaluation of eHealth technology. The proposed questions can be adjusted based on the goal of the technology and the incorporated behavioral models and/or theories. The answers to the research questions can provide input for improving the look and feel and architecture of the technology as well as the fit between the technology, the user, and the context. In turn, this information can be used to increase the effectiveness, persuasiveness, and the long-term usage of the technology.

Data Collection

Depending on the research questions, there are different ways to collect log data. To gain rich and in-depth knowledge regarding the usage patterns of individual users server-side log data, containing information about communications with the server (requests such as opening a page, clicking a link, saving health values, or other information), can be collected. In Web-based applications, this type of logging is preferably a file where the Web addresses of the requested subpages of the system are registered. This is the most efficient way without substantially losing system performance. Another advantage is that, after updates and modifications of the system, Web addresses referring to new subpages and functions are automatically logged.

On the other side, a single user action (eg, clicking a button to add a health measurement) can lead to multiple server requests leading to multiple Web addresses in the log data file. This can make it harder to identify single user actions. It is therefore necessary to determine which (combinations of) server requests specify certain actions and to link these to a specific identification for that action, such as a code or description.

Besides server-side logging, client-side logging can also provide valuable information regarding the usage of the technology. Client-side information contains actions that do not require server requests, such as scrolling up and down the screen, moving the pointer, and clicking and filling out a text field. The research questions determine the logging method that provides the most valuable information. However, it is also important to take the possibilities and the consequences of the different logging methods into account, such as a loss of system performance.

Requirements of Log Data

Log data files are most often Comma-Separated Values (.csv) files that can easily be opened using Excel or SPSS. Information regarding the user identity, date and time of the action, and an identification of the action is essential to identify the user, logins, and the usage patterns within and between logins. An example of a fictional log data file is shown in Figure 2. Depending on the research questions, additional information can be desirable regarding the device (eg, personal computer, mobile phone, tablet, or wearable), a specification of the action (eg, measurements or other information saved in the database), the Global Positioning System (GPS) coordinates of the user, and the status of the user collected via wearables, such as stress, sleep, and activity patterns.

In order to answer the research questions, data files should be of sufficient quality, wherein the goal of the technology and the used behavior change theories and models form the basis for the data that is needed for the analyses. For example, if the research question is “What are predictors for adherence to the technology?” then the data should contain information from which the adherence can be derived, as well as the variables (such as user or usage characteristics) that might possibly predict adherence. When the focus is on exploring the dose-response relationship, there must be a possibility to link the log data to other outcomes, for instance, via the user identification number.

The amount of data needed depends on the complexity of the research question and can only be determined empirically. In general, a reasonable amount of data per user and a reasonable amount of users are needed in the dataset. For example, when analyzing 100 usage sessions, 10 usage sessions of 10 different users provide more generalizeable information than 50 usage sessions of 2 different users. Of course, the more data the better, but it needs to contain the needed information to answer the research question as well.

Importantly, the data should be available for analysis under the applicable privacy regulations, whether or not with informed consent of the individual users of the technology. Informed consent depends upon whether log data includes or needs to be combined with personal data. Currently, it remains undecided whether log data in itself is personal data, as usage data does not always, per se, contain information that can be traced back to individual users. However, as the possibilities for data analytics develop, it may become quite possible in the (near) future to trace users back to specific individuals based on their usage patterns on other technologies. Narayanan and Shmatikov, for example, were able to de-anonymize Netflix users based on reviews in the Internet Movie Database (IMDB) [27]. Thus, it may be hard to truly anonymize log data. Therefore, researchers need to always consider the applicable regulations to determine whether informed consent is needed from the users before log data can be collected and analyzed for research purposes.

Figure 2. A fictional example of log data.

User	Time Stamp	Action	Extra information
1	January 12; 01:14 p.m.	Login	
1	January 12; 01:21 p.m.	Login	
2	January 12; 01:20 p.m.	Login	
2	January 12; 01:22 p.m.	Opening monitoring	
2	January 12; 01:47 p.m.	Adding monitoring value	Blood pressure
3	January 21; 10:11 a.m.	Login	
3	January 21; 10:12 a.m.	Opening mailbox	
3	January 21; 10:13 a.m.	Opening monitoring	
3	January 21; 10:13 a.m.	Opening mailbox	
3	January 21; 10:21 a.m.	Send message	To general practitioner
1	January 23; 10:11 a.m.	Login	
1	January 23; 10:13 a.m.	Opening mailbox	
1	January 23; 10:15 a.m.	Send message	To nurse practitioner

Data Preparation

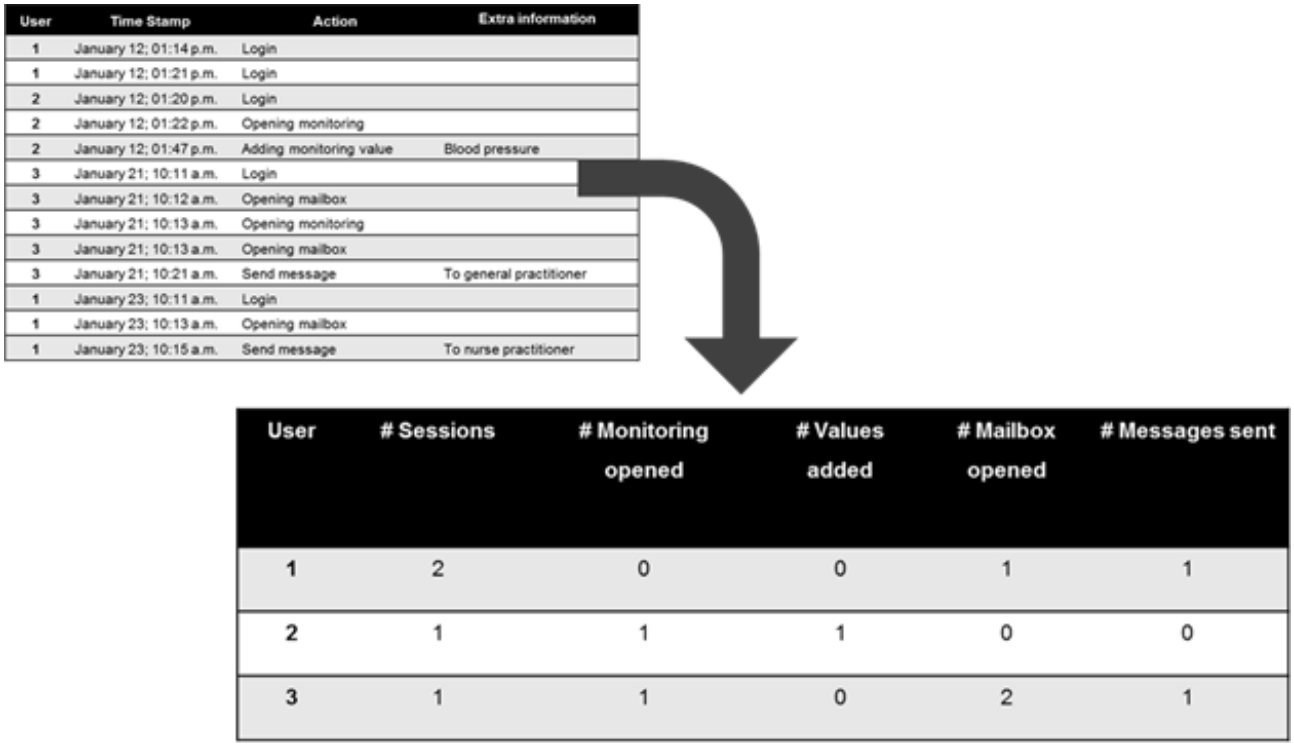
Preparing the data before analysis is vital since, for the most part, typical log data consist of tens of thousands records. Hence, these records must first be filtered for the information that is needed for the analysis (eg, Web addresses or the codes for specific actions). Then this information needs to be translated into new variables, such as the number of sessions and/or activities per user, sequence of the activities, and/or time spent per login. An example of how the raw data depicted in Figure 2 can be translated into data for analyses is shown in Figure 3.

In Figure 3, every row in the new data set represents one user. The number of sessions is defined as a period of activity ended by a period of at least 30 minutes of inactivity. In this example, this definition has consequences for User 1, having 3 logins and 2 sessions. By counting the number of logins it would seem like this user has a higher activity level then by counting the number of sessions. It is possible that a user picks up the activities where he left off when returning within 30 minutes after the last action. Furthermore, no user actions were registered during the first and the second login. Hence, counting the number of logins (only) might give a distorted image on the amount of actual use of the different elements of technology.

Second, a distinction was made in this example between visiting a certain element of the technology and actually using it (eg, opening, monitoring, and adding a value, or opening the mailbox and sending a message). Thus, 2 out of the 3 users in this example opened the function for monitoring, but only 1 actually used this function by adding a value to the database. Furthermore, 2 out of 3 users opened their mailboxes and sent a message to a caregiver, where User 3 opened the mailbox twice but sent a message only once.

These are only a few examples of the variables that can be calculated from a raw log data set. Depending on the research questions, new or other variables can be calculated as well. When the question is how the usage in the first sessions correlates to adherence, a distinction can be made between the activities (eg, opening the mailbox and sending a message) in the first, second, third, and further sessions. Also, variables can be added regarding the adherence by a user (where an adherent user is indicated with a “1” and a non-adherent user with a “0”), whether the sent message has been answered (and within what time period), the location of the user (where GPS coordinates are given for every session), the emotional status of the user, activity levels, or the time (in days or hours) between the previous and the current sessions. By combining different data sources, new variables do not have to contain, by definition, information from the same data file.

Figure 3. An example of data transformation, based on the data in Figure 2.



User	Time Stamp	Action	Extra information
1	January 12; 01:14 p.m.	Login	
1	January 12; 01:21 p.m.	Login	
2	January 12; 01:20 p.m.	Login	
2	January 12; 01:22 p.m.	Opening monitoring	
2	January 12; 01:47 p.m.	Adding monitoring value	Blood pressure
3	January 21; 10:11 a.m.	Login	
3	January 21; 10:12 a.m.	Opening mailbox	
3	January 21; 10:13 a.m.	Opening monitoring	
3	January 21; 10:13 a.m.	Opening mailbox	
3	January 21; 10:21 a.m.	Send message	To general practitioner
1	January 23; 10:11 a.m.	Login	
1	January 23; 10:13 a.m.	Opening mailbox	
1	January 23; 10:15 a.m.	Send message	To nurse practitioner

User	# Sessions	# Monitoring opened	# Values added	# Mailbox opened	# Messages sent
1	2	0	0	1	1
2	1	1	1	0	0
3	1	1	0	2	1

Data Analysis

Once the log datasets are prepared, the files are ready for analysis. The first, and up to now, most commonly used method is to describe the frequencies of use, including number of logins or sessions per user, use of the different elements of the technology, moment of use, and time spent on the technology. Although more exposure to an eHealth technology does not always lead to better health outcomes, this information might still provide a starting point for further research. Next, pathway analyses and predictive modeling provide deeper insights into the usage patterns of individual users.

A pathway analysis can provide more information regarding the different usage patterns that occur. In previous research, for example, the usage patterns of adherers and early and late non-adherers to the ”Living to the Full“ intervention were compared [21]. A 1-way analyses of variance (ANOVA) and chi-square tests showed that early non-adherers used fewer and shorter sessions than late non-adherers and adherers. Early non-adherers also used fewer sessions to complete a lesson. Furthermore, late non-adherers had a shorter total duration of sessions than the adherers. Logistic regression was used to assess the baseline characteristics of adherers and non-adherers.

We have also analyzed usage patterns of first visits to the PHR for patients with chronic conditions [23]. The results showed that users followed the structure of the system. While these analyses were conducted by hand on a subset of all users, methods for Markov modeling can be more convenient for analyzing the dominant path through the system of a more extended group of users [28,29]. This methodology can be used, for example, to analyze how people use the different elements of a technology in terms of frequency and the order they select to reach a certain goal [30].

More advanced and predictive information for pattern recognition within complex data sets can be obtained by applying machine-learning algorithms [31]. To do so, the Waikato Environment for Knowledge Analysis (Weka) tool is a relatively accessible and easy to use software package for applying machine learning algorithms for data mining tasks [32]. By using Weka, supervised learning, unsupervised learning, and market-basket methods for analysis can be applied. Supervised learning predicts adherence and effects from early use patterns, which enables early intervention for users at risk [31,33,34]. This method has the potential to answer research questions concerning when users drop out and what the predictors are for users dropping out or returning to the application. Unsupervised learning determines what usage profiles appear from the log data and if this data can be matched to a certain group of participants [33,35]. This method has the potential to answer research questions like ”What are the characteristics of (non-)users, and who are the dropouts?“ Market-basket analysis allows researchers to ask what elements of the technology are often used together [36]. Examples of such analyses can be found in the domain of online shopping, where customers get to see suggestions of products based on the products they looked at.

Although it is difficult to make predictions based on the usage data of relatively small groups of users (eg, in a RCT) and not all research questions can be answered with this data, even these analyses can provide valuable scientific and practical input for future system improvements.

Discussion

Log data analyses can be used as part of the formative, as well as the summative, evaluation of eHealth technology. As a formative evaluation, log data can provide ongoing and real-time

information on how to improve the technology and on the process in which the technology is embedded. As part of the summative evaluation, log data can provide explanations on the uptake and the outcomes of the technology, which can be both scientifically and practically valuable.

Scientific Translation

Log data analyses provide input for opening the black box of eHealth. Log data analyses not only provide insight into the effects of the single elements of a technology, but can also stipulate essential information about the effects of combinations of elements. In this way, log data can provide input to better understand the results of experimental research designs, such as RCTs or the multiphase optimization strategy (MOST) for eHealth evaluation. In a MOST, potential effective elements are selected for incorporation in an intervention (based on existing theories and/or previous research) and tested in 3 subsequent phases (screening phase, refining phase, confirming phase) [37]. Log data has the potential to validate the results of a RCT or these different phases of the MOST. For example, did the users actually find all the incorporated elements of the technology and are these elements used in the intended way? [24] And how does the use of the different elements correlate to the found effects, and for whom? Based on such insights, existing technologies can be improved and effective elements can be identified and combined into new technologies. These results are not always revealed through questionnaires, interviews, or usability tests.

Another advantage of a log data analysis is that it can reveal real-time insights into the user's response to specific persuasive triggers in different situations (eg, in terms of location, status of the user), providing new possibilities for the timing of persuasion [38]. Furthermore, several studies have demonstrated that individuals respond differently to the same persuasive strategies [39], indicating that personalization of a technology (adapting a technology to individual users) might increase a program's persuasiveness and its long-term use and effectiveness [40].

In the domain of eHealth, personalization is often limited to adapting the content of the technology to a confined user profile based on user characteristics like age, gender, and level of education [41]. However, there is evidence that such demographics (user profiles) do not predict engagement [42]. User profiles, such as early usage patterns for example, do potentially predict whether or not an individual will maintain long-term use of an application. For example, Freyne et al [25] found that individuals' use of specific elements of a technology in the first week influences their use of that technology in the second week. Based upon these results, more extended user profiles can be created that take (early) usage behavior into account, extending user profiles beyond a limited set of user characteristics.

Log data analyses also allow a timely response to flaws in the technology, a shorter evaluation cycle, and the development of more transparent technology, as stressed in the "agile science" approach [7]. Until now, modifications of technologies have often been made after an evaluation period. However, technology use is not stagnant, rather it is dynamic and changing over time, and statistics that fit this characteristic are required [43]. With a real-time analysis of log data, adaptive interfaces can be created that respond to individual (changes in) usage patterns. The effects of these adaptive interfaces on the usage can then be analyzed further and improved.

Log data can also be used to test the models and theories that are incorporated in the technology in order to improve the existing behavioral models. Patrick and colleagues made the following comparison for this approach:

> *It could be argued that today's current behavioral theories are akin to the Farmer's Almanac as they are largely descriptive, past-oriented, and simplified to a few elements. These models for understanding behavior and behavior change provide largely "on average" insights without the level of specification and prediction that could occur in behavioral science if the approach to communication, data, and iterative evaluation of computationally complex, multilevel models now common in meteorology could be replicated. (p829) [31]*

Log data has the potential to predict usage behavior and can thus be of added value in the development of complex, multilevel models for behavior change.

Furthermore, machine-learning algorithms can make predictions regarding whether and when a user might dropout from using the technology, making it possible to intervene in a timely manner and increase adherence to the technology. However, when focusing on research questions to assess adherence, it is important to substantiate this term: when is a user considered to be adherent? In research, assumptions are made about the intended usage of a technology which are not well defined or evidence-based [44-46]. As a result, it is hard to assess the results of the analyses and compare those to the outcomes of other, similar technologies. A definition of adherence does not always have to be derived from the extent to which a technology is completed (eg, a user is adherent when he/she completed 3 out of 4 lessons), but can also be extracted from other literature. For example, Kaushal and Rhodes discovered that exercising for at least 4 times a week for 6 weeks was the minimum activity to establish an exercise habit [47]. This type of evidence can be used for defining adherence to a technology, for example, a user is adherent when at least 4 usage sessions per week can be identified for a period of 6 weeks. In this example, mere login data (eg, the more logins, the better the adherence is) does not reveal adherence, but assigning a substantiated threshold value does.

An advantage of log data is that it is always available and easy to collect, without requiring any extra effort from the participants. A common problem in (eHealth) research is that participants often find it time consuming and labor intensive to complete questionnaires at different time points or to participate in an interview or focus group, resulting in dropouts from the research study. However, this result does not necessarily mean that the same participants who did not participate in the research dropped out from using the technology. By using log data in addition to questionnaires, researchers have more than one method to collect data and are no longer dependent on having

a majority of the participants complete questionnaires or participate in interviews or focus groups.

However, there are important limitations for using log data in eHealth evaluation. First, the results of the log data analyses do not always indicate why certain usage patterns occur. It is therefore important to use a mixed-methods approach to combine the analyses with additional research via interviews, usability tests, or other quantitative and qualitative research methods. For example, the log data analysis from the ”Living to the Full“ intervention showed that a fairly large group dropped out during the 6th lesson. It was revealed by counselors giving the course that this is indeed a hard lesson for participants because of the focus on observing themselves and learning new skills to accept suffering [21]. Additional research can provide more precise insight into what users experience or why they tend to drop out at certain points. Log data analysis focused on such questions can provide researchers specific areas or user groups to examine through future interviews, questionnaires, or usability tests. The results of these evaluations can then be used to improve the technology as well as to highlight the crucial moments in the treatment protocols for blended therapy. Furthermore, using log data in research might require an extra effort from researchers, developers, database managers, etc. For example, it takes time to develop a plan for data collection, management, and analysis, as well as to incorporate the possibility for data collection into the technology.

Practical Value

Besides the scientific value, the results of a log data analysis can be of added value for eHealth developers and healthcare providers. For example, the results of a pathway analysis and the identified usage profiles can be used as input for adapting and matching the system design to the users in order to make the technology more persuasive. Information regarding the elements that are often used together can also provide real-time feedback and suggestions to the users, guiding their follow-up actions in the system such as:

> *You have added a goal. Other users have added their current weight as well. Click here to add your weight*

Because log data analysis via (un)supervised learning can provide information about users that might potentially drop out from an intervention, on a practical level, healthcare providers can then make use of this information to intervene and stimulate these users to continue using the system. In addition, log data can be used to show healthcare providers how their responsiveness to client messages influences a client’s adherence to the therapy. When composing protocols for (blended) care via eHealth technologies, researchers can then take advantage of the added value of log data analyses.

Until recently, technologies have often changed after an evaluation period, but with a real-time analysis of log data, adaptive interfaces can be created that respond to individual users. The effects of the interface on the use of the technology can then be directly identified, allowing a fast response to flaws in the technology, a shorter evaluation cycle, and the development of more transparent technology.

Conclusions

The analysis of log data can be of great value for scientists and designers as well as caregivers and policy makers in their research into the black box of eHealth technology. A deliberate analysis of log data can provide insight into the usage of the technology by all users as a group as well as by individual users, helping to accelerate the persuasiveness and effectiveness of eHealth technology. Furthermore, log data can be used to assess the theories that underpin a technology. However, from the collection of log data to translating the results into valuable information, various steps need to be taken, each with their own considerations. This paper serves as a starting point for using log data analysis in eHealth research.

Conflicts of Interest

None declared.

References

1. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. CAIS 2009;24(1):28 [FREE Full text]
2. Kelders SM, Oinas-Kukkonen H, Oörni A, van Gemert-Pijnen JE. Health behavior change support systems as a research discipline; a viewpoint. Int J Med Inform 2016 Dec;96:3-10. [doi: 10.1016/j.ijmedinf.2016.06.022] [Medline: 27432207]
3. Nijland N, van Gemert-Pijnen JE, Kelders SM, Brandenburg BJ, Seydel ER. Factors influencing the use of a Web-based application for supporting the self-care of patients with type 2 diabetes: a longitudinal study. J Med Internet Res 2011 Sep 30;13(3):e71 [FREE Full text] [doi: 10.2196/jmir.1603] [Medline: 21959968]
4. 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 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
5. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 2011 Jan 18;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
6. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

7. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Jun;6(2):317-328 [FREE Full text] [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
8. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. J Behav Med 2017 Feb;40(1):85-98. [doi: 10.1007/s10865-016-9818-7] [Medline: 28058516]
9. Han JY. Transaction logfile analysis in health communication research: challenges and opportunities. Patient Educ Couns 2011 Mar;82(3):307-312. [doi: 10.1016/j.pec.2010.12.018] [Medline: 21277146]
10. Resnicow K, Strecher V, Couper M, Chua H, Little R, Nair V, et al. Methodologic and design issues in patient-centered e-health research. Am J Prev Med 2010 Jan;38(1):98-102 [FREE Full text] [doi: 10.1016/j.amepre.2009.09.034] [Medline: 20117564]
11. van Gemert-Pijnen JE. In: van Gemert-Pijnen JE, Peters O, Ossebaard HC, editors. Improving eHealth. The Hague: Eleven International Publishing; 2013.
12. Bruckman A. Analysis of log file data to understand behavior learning in an online community. In: Weiss J, Nolan J, Hunsinger J, Trifonas P, editors. International Handbook of Virtual Learning Environments. Dordrecht: Springer Netherlands; 2006:1449-1465.
13. Jamali HR, Nicholas D, Huntington P. The use and users of scholarly e‐journals: a review of log analysis studies. AP 2005 Dec;57(6):554-571. [doi: 10.1108/00012530510634271]
14. Dumais S, Jeffries R, Russell D, Tang D, Teevan J. Understanding user behavior through log data analysis. In: Olson JS, Kellogg WA, editors. Ways of Knowing in HCI. New York, NY: Springer New York; 2014:349-372.
15. Jansen BJ. Search log analysis: What it is, what's been done, how to do it. Libr Inf Sci Res 2006;28(3):407-432.
16. Choi Y, Chang S, Kim Y, Lee H, Son W, Jin S. Detecting and monitoring game bots based on large-scale user-behavior log data analysis in multiplayer online games. J Supercomput 2015 Oct 26;72(9):3572-3587. [doi: 10.1007/s11227-015-1545-2]
17. Chatterjee S, Price A. Healthy living with persuasive technologies: framework, issues, and challenges. J Am Med Inform Assoc 2009;16(2):171-178 [FREE Full text] [doi: 10.1197/jamia.M2859] [Medline: 19074300]
18. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res 2011 Jan 25;13(1):e9 [FREE Full text] [doi: 10.2196/jmir.1391] [Medline: 21371992]
19. Kim E, Stolyar A, Lober WB, Herbaugh AL, Shinstrom SE, Zierler BK, et al. Usage patterns of a personal health record by elderly and disabled users. AMIA Annu Symp Proc 2007 Oct 11:409-413 [FREE Full text] [Medline: 18693868]
20. Kuijpers W, Groen WG, Oldenburg HS, Wouters MW, Aaronson NK, van Harten WH. eHealth for breast cancer survivors: use, feasibility and impact of an interactive portal. JMIR Cancer 2016 May 10;2(1):e3 [FREE Full text] [doi: 10.2196/cancer.5456] [Medline: 28410178]
21. Kelders SM, Bohlmeijer ET, van Gemert-Pijnen JE. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013 Aug 20;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
22. Pagliari C, Detmer D, Singleton P. Potential of electronic personal health records. BMJ 2007 Aug 18;335(7615):330-333 [FREE Full text] [doi: 10.1136/bmj.39279.482963.AD] [Medline: 17703042]
23. Sieverink F, Kelders SM, Braakman-Jansen LMA, van Gemert-Pijnen JE. The added value of log file analyses of the use of a personal health record for patients with type 2 diabetes mellitus: preliminary results. J Diabetes Sci Technol 2014 Mar;8(2):247-255 [FREE Full text] [doi: 10.1177/1932296814525696] [Medline: 24876574]
24. Van Gemert-Pijnen JE, Kelders SM, Bohlmeijer ET. Understanding the usage of content in a mental health intervention for depression: an analysis of log data. J Med Internet Res 2014 Jan 31;16(1):e27 [FREE Full text] [doi: 10.2196/jmir.2991] [Medline: 24486914]
25. Freyne J, Saunders I, Brindal E, Berkovsky S, Smith G. Factors associated with persistent participation in an online diet intervention. In: CHI '12 Extended Abstracts on Human Factors in Computing Systems. 2012 Presented at: CHI Conference on Human Factors in Computing Systems; 2012 May 05-12; Austin, Texas p. 2375-2380.
26. Bohlmeijer E, Fledderus M, Rokx T, Pieterse M. Efficacy of an early intervention based on acceptance and commitment therapy for adults with depressive symptomatology: evaluation in a randomized controlled trial. Behav Res Ther 2011 Jan;49(1):62-67. [doi: 10.1016/j.brat.2010.10.003] [Medline: 21074752]
27. Narayanan A, Shmatikov V. Robust de-anonymization of large sparse datasets. 2008 Presented at: Proceedings of the 2008 IEEE Symposium on Security and Privacy; 2008 May 18-21; Oakland, California p. 111-125.
28. Seneta E. Markov and the birth of chain dependence theory. Int Stat Rev 1996 Dec;64(3):255-263. [doi: 10.2307/1403785]
29. Borges J, Levene M. Evaluating variable-length Markov chain models for analysis of user web navigation sessions. IEEE Trans Knowl Data Eng 2007 Apr;19(4):441-452. [doi: 10.1109/TKDE.2007.1012]
30. Akkersdijk SM, Kelders SM, Braakman-Jansen LMA, van Gemert-Pijnen JE. Using Markov chains to analyze paths through a personal health record. In: Persuasive Technology XII (Adjunct Proceedings). 2017 Presented at: Persuasive Technology XII; 2017 April 03-06; Amsterdam.

31. 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]
32. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software. SIGKDD Explor Newsl 2009 Nov 16;11(1):10. [doi: 10.1145/1656274.1656278]
33. Han J, Pei J, Kamber M. Data Mining: Concepts and Techniques. Burlington, Massachusetts: Morgan Kaufmann; 2011.
34. Hekler EB, Michie S, Pavel M, Rivera DE, Collins LM, Jimison HB, et al. Advancing models and theories for digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):825-832. [doi: 10.1016/j.amepre.2016.06.013] [Medline: 27745682]
35. Paliouras G, Papatheodorou C, Karkaletsis V, Spyropoulos C. Discovering user communities on the Internet using unsupervised machine learning techniques. Interact Comput 2002 Dec;14(6):761-791 [FREE Full text] [doi: 10.1016/S0953-5438(02)00015-2]
36. Anand S, Patrick A, Hughes J, Bell D. A data mining methodology for cross-sales. Knowl Based Syst 1998 May;10(7):449-461. [doi: 10.1016/S0950-7051(98)00035-5]
37. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
38. Räisänen T, Oinas-Kukkonen H, Pahnila S. Finding kairos in quitting smoking: smokers' perceptions of warning pictures. In: Persuasive Technology: Third International Conference. 2008 Presented at: Persuasive Technology 2008; 2008 Jun 04-06; Oulu, Finland p. 254-257.
39. Kaptein M, Eckles D. Heterogeneity in the effects of online persuasion. JIM 2012;26(3):176-188.
40. Kaptein M, Markopoulos P, de Ruyter B, Aarts E. Personalizing persuasive technologies: explicit and implicit personalization using persuasion profiles. Int J Hum Comput Stud 2015 May;77:38-51. [doi: 10.1016/j.ijhcs.2015.01.004]
41. Kaptein M, Lacroix J, Saini P. Individual Differences in Persuadability in the Health Promotion Domain. In: PERSUASIVE 2010. 2010 Presented at: Proceedings of the 5th International Conference on Persuasive Technology; 2010 Jun 07-10; Copenhagen, Denmark p. 94-105.
42. van Mierlo T, Li X, Hyatt D, Ching AT. Demographic and indication-specific characteristics have limited association with social network engagement: evidence from 24,954 members of four health care support groups. J Med Internet Res 2017 Feb 17;19(2):e40 [FREE Full text] [doi: 10.2196/jmir.6330] [Medline: 28213340]
43. Magni M, Susan Taylor M, Venkatesh V. 'To play or not to play': a cross-temporal investigation using hedonic and instrumental perspectives to explain user intentions to explore a technology. Int J Hum Comput Stud 2010 Sep;68(9):572-588. [doi: 10.1016/j.ijhcs.2010.03.004]
44. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
45. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
46. Su W, Chih M. Is more eHealth system use better for cancer patients and family caregivers? A literature review. 2016 Presented at: AMIA 2016 Annual Symposium; 2016 Nov 12-16; Chicago, IL. [doi: 10.13140/RG.2.1.4077.7200]
47. Kaushal N, Rhodes RE. Exercise habit formation in new gym members: a longitudinal study. J Behav Med 2015 Aug;38(4):652-663. [doi: 10.1007/s10865-015-9640-7] [Medline: 25851609]

Abbreviations

ACT: acceptance and commitment therapy
eHealth: electronic health
GPS: Global Positioning System
MOST: multiphase optimization strategy
PHR: personal health record
RCT: randomized controlled trial
Weka: Waikato Environment for Knowledge Analysis

Edited by G Eysenbach; submitted 05.08.16; peer-reviewed by P Karppinen, JY Han, YT Chien, GY Liao, WC Su, D Smith; comments to author 25.11.16; revised version received 02.03.17; accepted 12.07.17; published 07.08.17.

Please cite as:
Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L
Opening the Black Box of Electronic Health: Collecting, Analyzing, and Interpreting Log Data
JMIR Res Protoc 2017;6(8):e156
URL: http://www.researchprotocols.org/2017/8/e156/
doi:10.2196/resprot.6452
PMID:28784592

Original Paper

Barriers and Facilitators for the Use of a Medical Mobile App to Prevent Work-Related Risks in Pregnancy: A Qualitative Analysis

Adeline V Velu[1], MD; Monique DM van Beukering[1], MD; Frederieke G Schaafsma[2], MD, PhD; Monique HW Frings-Dresen[3], MD PhD; Ben WJ Mol[4], MD PhD; Joris AM van der Post[1], MD PhD; Marjolein Kok[1], MD, PhD

[1]Academic Medical Center, Department of Obstetrics and Gynecology, University of Amsterdam, Amsterdam, Netherlands

[2]University Medical Center, Department of Public and Occupational Health/Amsterdam Public Health Institute, Vrije Universiteit, Amsterdam, Netherlands

[3]Academic Medical Center, Coronel Institute of Occupational Health, University of Amsterdam, Amsterdam, Netherlands

[4]Robinson Institute, University of Adelaide, Adelaide, Australia

Corresponding Author:
Adeline V Velu, MD
Academic Medical Center
Department of Obstetrics and Gynecology
University of Amsterdam
Room H4-232
Meibergdreef 9
Amsterdam,
Netherlands
Phone: 31 20 5663453
Fax: 31 20 6963489
Email: a.v.velu@amc.uva.nl

Abstract

Background: The number of women participating in the labor market in Europe has increased over the last several decades. At the same time, there is growing evidence that certain conditions of employment during pregnancy may have a negative influence on pregnancy outcomes. In order to better inform pregnant women, we aim to develop an app to help assess the health risk as a result of personal and work-related factors and provide personal advice for these women and their health care providers.

Objective: The aim of this study was to compose a thematic overview of the perceived facilitators and barriers according to pregnant women, medical professionals, and employers for the use of a mobile app in obstetrical care to prevent occupational-related pregnancy complications.

Methods: Two multidisciplinary focus group meetings with in total 14 participants were conducted with pregnant women, occupational physicians, general practitioners, midwives, obstetricians, and representatives of trade unions and employer organizations. Transcripts were analyzed by qualitatively coding procedures and constant comparative methods.

Results: We identified 24 potential facilitators and 12 potential barriers for the use of the app in 4 categories: content of the app, the app as a mean to provide information, ease of use, and external factors. The 3 main facilitators identified were the need for a good interaction between the app and the user, apps were viewed as a more practical source of information, and the information should be understandable, according to the existing guidelines, and well-dosed. The 2 main barriers for use were extensive battery and memory use of the smartphone and sending frequent push notifications.

Conclusions: The results of this study are important considerations in the developing process of a medical app implementing a guideline or evidence-based information in practice.

(JMIR Res Protoc 2017;6(8):e163) doi:10.2196/resprot.7224

KEYWORDS

qualitative research; mobile app; smartphone; pregnancy; work; occupation; exposure; eHealth; mHealth

Introduction

Currently the employment rate among women aged 20 to 64 years is 64% in Europe [1]. Around 57% of women in the labor force in the Netherlands are of childbearing age [2]. At the same time, there is growing evidence that certain conditions in employment during pregnancy may have a negative influence on pregnancy outcomes. For instance, working long hours in a day or working night shifts, physically demanding work, stress, and chemical, pharmaceutical, or biological exposure can potentially cause preterm birth, low birthweight, spontaneous abortion, stillbirth, and fetal abnormalities [3-9].

Pregnant women are often unaware of potential work-related risks to their pregnancy [10]. Estimations are that only 25% of employed pregnant women receive adequate counseling on work-related risks during their pregnancy [11]. Furthermore, van Beukering et al [12] concluded in a literature study that around 25% of pregnant women in the Netherlands come in contact with above mentioned work-related risks.

If pregnant women would receive more information about potential risks in their work situation, this could lead to better work adjustments. In the Netherlands, occupational physicians developed a guideline for healthy working conditions during pregnancy and the postpartum period [13]. This guideline provides clear advice on necessary adjustments to potential harmful working conditions. With these work adjustments, a healthy working environment can be created and prevent negative pregnancy outcomes in certain cases.

Mobile health (mHealth) was defined by the World Health Organization as the use of mobile devices (mobile phones, patient monitoring devices, and personal digital assistants) for medical and public health practice [14]. The benefits of mHealth interventions include that they can be delivered anywhere at any time and they provide opportunities for interactivity and tailoring to specific groups [15]. Mobile apps and smartphones are increasingly used in health care by both health care workers and the general public. In 2015, about 94% of the population in the age category of 25 to 45 years owned a smartphone with Internet access in the Netherlands [16]. Although these data were not specified by gender, it is likely that the use of smartphones is comparable between men and women in this age category. The promising research results of apps in health care, combined with the fact that smartphones are widely used by many women of childbearing age, gives smartphones the potential to further improve maternity care as an addition to the traditional health care system [17].

For pregnant women, several mHealth interventions or apps were previously developed for the care of diabetes [18-21], achieving less gestational weight gain [22-24], and support of smoking cessation [25-27]. The effectiveness of these interventions showed promising results, although most of these studies did not show significant effects on health outcomes mainly due to small sample sizes [18-23,25,27]. A recent large study in the Netherlands, however, did show significant improvement of nutrition and lifestyle due to an mHealth platform in 603 pregnant women and 1275 couples contemplating pregnancy [28].

Previous research has shown that the user satisfaction of an app or mHealth intervention is high among active users, and most are viewed as helpful, useful, and convenient [25,27,29-31]. However, continued use lagged behind and drop-out rates were high [22,32,33]. For instance, in a large nationwide email-based health promotion program for pregnant women in the Netherlands, 45% ceased participating or never opened an email. Only 16% opened all emails received and were considered very active [32]. Therefore, it is important to evaluate potential facilitators and barriers for the use of an app during the development phase to achieve good and continued use of such an app.

In order to better inform pregnant women, we wanted to develop an app to help assess health risk as a result of personal and work-related factors and provide personal advice for women and their health care providers. In doing so we wished to create awareness of work-related risks and empower pregnant women to discuss necessary work adjustments with their supervisor and potentially prevent negative pregnancy outcomes. To our knowledge, there is no literature on the use of an app that provides personal advice for pregnant women addressing work-related risks and relevant work adjustments.

This study was the first phase of a 3-phase pilot study. After this phase, the prototype of the app will be tested for usability during a think-aloud study among pregnant women. The third phase will consist of a powered study comparing the app as an addition to standard care with standard care alone. These phases are based on models for developing new tools as studied by Elwyn et al [34] and Shorten et al [35].

The aim of this study was to compose a thematic overview of the perceived facilitators and barriers by pregnant women, professionals, and employers for the use of an app in obstetrical care to reduce occupational-related pregnancy complications.

Methods

Overview

We performed qualitative research by conducting 2 multidisciplinary focus group meetings with a total of 14 participants. We decided to conduct multidisciplinary focus groups to involve all the stakeholders and thereby evaluate a variety of opinions of both the end-users and professionals. The methods and results were reported according to the Consolidated Criteria for Reporting Qualitative Research [36]. The ethics board of the Academic Medical Center confirmed that the Medical Research–Involved Human Subjects Act did not apply to this study.

Participants

Participants were selected by purposive sampling of stakeholders involved in occupational health and obstetrical care and contacted by email and telephone. The inclusion criteria were that the participants were either pregnant women, occupational physicians, general practitioners, midwives, obstetricians, or representatives of trade unions and employer organizations. Participants who did not speak the Dutch language fluently were excluded. In total, we invited 30 potential candidates, between 4 and 6 from each stakeholder category, to ensure an adequate

number and variety of stakeholders in both focus groups. Invitations were only declined because of previous engagements, not because of unwillingness to participate.

Procedure

Two focus group meetings were conducted in 2015 to identify potential facilitators and barriers for the use of an app for pregnant women to prevent work-related risks during pregnancy. Prior to the meeting, confidentiality was assured and the process of the focus group was explained to the participants. All participants signed an informed consent form. Both meetings were audiotaped and fully transcribed afterwards. The focus group meetings were both facilitated by FS (female, occupational physician, senior researcher at a Dutch academic medical center, experienced in facilitating focus group). MvB (female, researcher on this project, occupational physician) and SD (female, coordinator of the regional network of birth care) took field notes. The duration of each meeting was planned for 2 hours; meetings were conducted in Dutch.

During the first part of the meeting, participants were briefly introduced to the background and aims of the project. Next, the participants were asked to respond to several questions about their knowledge and experience with pregnancy and work. We also asked about their knowledge of the Dutch guideline on pregnancy and work [13] and about experiences with health apps, mainly focused on lifestyle adjustments in general.

Subsequently, several examples of existing health apps were presented, followed by a discussion based on 5 predetermined statements (all questions, topics, and statements are shown in Multimedia Appendix 1).

Data Analysis

The transcriptions of the focus groups were structured and analyzed with MAXQDA (VERBI GmbH), a software program to assist qualitative data analysis. For the analyses of focus group transcripts, coding procedures and the constant comparative method developed by Strauss [37] were used. Coding was divided into 3 phases starting with open coding, axial coding, and selective coding. This is a frequently used inductive, bottom-up method for analyzing qualitative data without a predetermined theoretical framework [38].

First, each of the 2 researchers (AV, MvB) started with an open coding process by examining the transcripts of the focus groups in order to assign a series of codes that were then grouped into similar concepts [39]. To ensure consistency and intercoder reliability, the 2 focus group transcripts were independently coded by the 2 researchers. Discussions between the 2 researchers resulted in a consensus list of preliminary codes. In case of discussion on the interpretation of the codes, a third researcher (FS) was involved in the process. Second, according to the axial coding process, recurrent themes within the transcripts were selected, and text fragments were sorted according to the thematic framework that appeared during the axial coding process, divided in main and subcodes. All codes were analyzed for influencing the use of the app, either in a positive way by stimulating the usage of the app coded as a facilitator or in a negative way coded as a barrier. Some citations could be interpreted as both a facilitator and a barrier. Consensus meetings between researchers led to the final categorization of themes as described in the section below.

Results

Overview

Each focus group consisted of 7 participants. The basic demographics of the participants are shown in Table 1. We successfully achieved the aim that in both meetings all different stakeholders were represented.

During the focus group meetings, the participants identified 24 potential facilitators and 12 potential barriers for the use of our app which were classified into 4 main themes: content of the app, the app as a means for providing information, ease of use, and external factors of influence. The barriers and facilitators in each main theme will be discussed below.

Table 1. Basic demographics (N=14).

Characteristics	n (%)
Gender	
Male	2 (14)
Female	12 (86)
Occupation	
Midwife	2 (14)
Obstetrician	1 (7)
Occupational physician	3 (21)
General practitioner	2 (14)
Employer	2 (14)
Labor union	1 (7)
Physician at unemployment insurance agency	1 (7)
Pregnant woman	2 (14)

Theme 1: Content of the App

Overview

The facilitators and barriers regarding content of the app can be divided in 2 subcategories: the content of provided information and provided advice and the added value of the app compared to existing apps. The 2 categories given most value by the participants are the content of the provided information and advice. Both categories can also be subdivided into personal information and advice specified to the individual user based on her previous responses about her work situation and a more general information and advice which applies to every working pregnant woman.

Content of the Provided Information and Advice

Participants agreed on the fact that facilitators related to general information and advice content are keeping the advice clear and simple and to mainly indicate the urgency or importance to follow the advice instead of going into too many details and background.

All information should be easily understandable for all users, and the information and advice should be in line with the existing guideline [13].

A second strong facilitator is the ability to provide specific personal information and advice by using selective questions. This way, it is possible to determine if the user is at risk for a certain complication and synchronize the advice with the gestational age.

> *You wanted a start question, how did you call it, a selective question?...Do you work in one of the following occupations, you should do that.* [MB, insurance physician, FG2]

> *You should actually be able to turn off information that is irrelevant to you. I do not work with toxic agents, so everything about that is irrelevant to me...I tune out if there are, say, two pages about that.* [MH, employer and pregnant woman, FG1]

Informing the pregnant woman about the changes in her body and the development of the fetus will improve her understanding of the effects the pregnancy may have on her work situation and vice versa. Possible adverse outcomes of the pregnancy are also important to mention in the app. Women with high-risk pregnancies could particularly benefit from specific and more personalized advice for their situation.

A potential barrier is the risk of users interpreting the information themselves without seeking further professional advice. One participant pointed out that a risk profile based on a questionnaire in the app cannot be compared with an actual conversation between a physician and a pregnant woman because the app only works with the input of the user herself. This makes the reliability of personal advice difficult to interpret and could become a barrier related to the content of the app.

> *Refer really fast to a gynecologist or midwife or indeed the occupational physician. Otherwise you will indeed risk that the pregnant woman herself will interpret medical information or interpret risk factors.* [FM, occuptational physician, FG 1]

Added Value of the App Compared to Existing Apps

Participants considered it a facilitator if the new app had added value with respect to other existing apps. Examples mentioned in the focus groups to create added value were (1) develop an app based on medical knowledge and guidelines, (2) cover the preconception and postpartum periods in addition to the pregnancy period, and (3) make the app noncommercial.

Theme 2: App as a Means for Providing Information

In this category, the focus groups reported mainly facilitators in relation to the app, most importantly the practical aspects. Moreover, apps were viewed as faster and easier in searching for information and the information was available at every place and every time.

> *Always at hand. Since that is the power of an app. You always have it on you. You can consult it anytime.* [AR, labor union, FG 1]

The fact that pregnant women already receive a large amount of information regarding their pregnancy can be interpreted both as a facilitator and as a barrier for the app. One point of view, as reported by the participants, is that the app is more easily accessible than printed information and therefore a facilitator for use. On the other hand, a few participants mentioned that the app provides more information and there is already enough information available.

Theme 3: Ease of Use

Overview

The facilitators and barriers for the ease of use of an app can be divided into 3 subcategories: technical aspects of the app, feedback and interaction between app and user, and reaching the target users by the mode of delivering the information to the user.

Technical Aspects

Participants described as facilitators the fact that games and quizzes make an app more fun to use. Another important facilitator for the use of the app was that it is only applicable for a set period of time and you can delete it after 9 months because apps that are not used frequently will be deleted, according to our participants.

Potential barriers that should be kept in mind are that there are numerous existing operating systems, apps that use a lot of battery and memory are unpopular, and the information provided should be readable on a smartphone.

> *What kind of apps do you delete?* [FS, facilitator, FG 2]

> *[Apps that use] lots of memory, lots of power. Apps that are very active, in that case your battery goes down...* [HB, general practitioner]

Feedback and Interaction Between App and User

Overall consensus was that interaction between the user and the app strongly stimulates the use of an app. But the opinions on

interaction also showed some inconsistencies between participants, and sometimes in participants their opinions seemed to vary. Several participants emphasized that messages about the development of the fetus and changes in the female pregnant body are informative and entertaining and facilitate the use of the app. Also, reminders of specific personalized advice based on an earlier risk analysis in the app were evaluated as helpful and welcome.

On the other hand, every participant criticized push notifications defined as frequent uncalled-for messages. One participant also mentioned that these push notifications can be risky when users have an adverse pregnancy outcome. Suggested solutions to this issue were to offer the option to sign out of the app in case of adverse pregnancy outcome or only show new general notifications when the user opens the app itself.

> *I fully recognize that. Because I do not have an app, but I do receive emails from an organization. And then you see the changes in your body and of the baby week by week, and say, and those of the baby. So in that respect I think receiving it through an app is useful. So you see the growth, and like, we are now in week 34; this is happening with your child. And you should adjust your health to your work et cetera. So I believe that would be very good.* [DD, pregnant woman, FG 1]
>
> *But do you delete apps that for instance send very many push notifications?* [FS, facilitator, FG 1]
>
> *Yes, I always turn them off immediately.* [CvW, employer]
>
> *Those are very irritating.* [MP, gynecologist]
>
> *Yes, those are very irritating.* [AR, labor union]

Reaching Target Users by the Mode of Delivering the Information

Three main facilitators were identified related to the mode of delivering information: the content of the app must be understandable, the information should be well-ordered, and the information should be supplemented with illustrations, video fragments, and icons to improve clarification.

> *If you reduce the text and do not use extensive amounts of text and work with icons that already helps.* [CdG, pregnant woman, FG 2]

The comprehensiveness of the app was considered an important facilitator and as such subject of long debate. Several suggestions were given to achieve a comprehensive app on pregnancy and work. For instance, participants felt that offering the app in multiple languages (Dutch, English, Spanish, and Polish were named as important; Moroccan and Turkish were questioned if they were still necessary), using plain language, and having a text-to-speech function can improve accessibility of the app for all users.

Providing too much information was viewed as a barrier by the participants, risking less usage of the app. Options to avoid this barrier could be to create a hyperlink in the app for further information and give users the option to read more if desired. On the other hand, the app should not be needlessly complicated with too many hyperlinks.

Theme 4: External Factors

Obstetrical Caregivers

Obstetrical caregivers such as gynecologists and midwives are facilitators by supporting the app, according to the participants. They work according to the occupational physician practice guideline [13] and believe in screening for work-related risks as part of the standard care. For it to become standard care, this knowledge should be implemented in the education for midwives. A second option could be to actively involve the obstetrical caregivers or create an extra app for the caregivers to use.

Employer, Supervisor, or Company

The employers can potentially be very strong facilitators for the use of the app. Unfortunately the participants in the focus groups mainly identified barriers for the app. The participants thought that employers may have a negative prejudice about work adjustments for pregnant women. Work adjustments can be seen as more bothersome than sick leave, and the entire organization, including colleagues, might not understand fully the need for adjustments. Employers have little knowledge about work-related risk factors for pregnant women, and many may not see that it is in their own best interest to implement well-timed work adjustment that could lower the risk of sick leave. Therefore, they may not stimulate the use of the app. Participants also pointed out that the app might cause a disturbance in the relationship between a pregnant woman and her employer. To prevent occupational conflicts, the advice in the app should be formulated cautiously and should emphasize stimulating a constructive dialogue.

> *Yes, I have experienced that myself, so to speak. That I basically did not dare to step up to my employer, when the last 2 weeks were quite heavy. I was aware that I was entitled to extra breaks, et cetera, but somehow I was afraid to speak up. So I do understand the story you just told, that when an employee shows up with solely an app, and the employer is not informed that this situation might give some, well, disruption, so to speak.* [CvW, employer, FG 1]

A significant factor in preventing these barriers is informing and involving the employers and organizations. If employers see the usefulness of the app itself and the importance of sustainable work during pregnancy, they may become more involved as facilitators for the use of the app by their employees.

The fact that there is a large variety in type and size of employers and companies is neither a facilitator nor a barrier for the use of the app. A footnote was placed by some participants that the app should be developed irrespective of the willingness of employers to participate. It cannot be expected that an app will change the entire culture of companies.

Government

One participant suggested that a television commercial from the ministry of health might facilitate the use of the app.

Table 2. Development of a mobile app: thematic overview of facilitators and barriers.

Theme and subcategory	Facilitators	Barriers
Theme 1: Content of the app		
Content of the provided information and advice		
	Understandable information (general)	
	Information and advice according to the existing guidelines (general)	
	Keeping advice clear and simple (general)	
	Showing only relevant and personal information to the user (personal)	Reliable personal advice is difficult when the risk profile is based only on a questionnaire (personal)
	Providing information on the changes in the pregnant body and development of the baby to better understand the impact on her work situation (personal)	It's important to provide some general advice to every user; the app shouldn't become too personal (personal)
	Using a selective question to determine if the user is at risk for a certain risk factor (personal)	
	Synchronizing the advice with the gestational age (personal)	
	Providing specific advice in case of adverse pregnancy outcome (personal)	
Added value compared to existing mobile apps		
	App should be based on medical knowledge and the guideline	
	Cover the preconception and postpartum periods in addition to the pregnancy period	
	Make the app noncommercial	
Theme 2: App as a means for providing information		
Practical aspects		
	App is easier and faster for searching for information and is always available	
Pregnant women already receive a lot of information regarding their pregnancy		
	App is more accessible than printed information	App provides even more information when there is already enough
Theme 3: Ease of use		
Technical aspects		
	Games and quizzes make the app more fun to use	There are several different operating system
	The app is only useful for 9 months and can be deleted afterwards	Creating an app that uses a lot of battery and memory of the smartphone
		Information should be readable in a smartphone; no pdf documents
Feedback and interaction between app and user		
	Providing messages about the development of the fetus and the pregnant body	Push notifications are unwanted
	Providing reminder messages of specific and personal advice based on earlier risk analyses	Push notifications are risky in case of adverse pregnancy outcomes
Reaching the target users by mode of delivering the information		
	Content of the app should be understandable for every user	Providing too much information
	Information should be well-ordered	
	Illustration, icons, and videos can provide clarification	

Theme and subcategory	Facilitators	Barriers
	Offer the possibility of linking to more information if desired	
Theme 4: External factors		
Obstetrical caregivers		
	Obstetrical caregivers support the app	
Employer, supervisor, or company		
	Employers are important for the app to succeed	Employers have little knowledge about work-related risk factors for pregnant women and don't see the benefit for themselves
		Employers might have a negative prejudice about work adjustments for pregnant women
		App can cause disturbance
Government		
	A television commercial might stimulate the app	

Discussion

Principal Findings

In this study we aimed to compose a thematic overview of the perceived facilitators and barriers for pregnant women, professionals, and employers for the use of a mobile app in obstetrical care to prevent occupational-related pregnancy complications.

We identified 24 facilitators and 12 barriers within 4 categorical themes, of which we identified 3 main facilitators and 2 main barriers to the successful implementation of our app in obstetrical care to reduce occupational-related pregnancy complications. The most important facilitator, in the opinion of our participants, is the need for good interaction between the app and the user to make the app personal to the user. The second facilitator is the fact that apps are viewed as a more practical source of information compared to traditional printed information. The third main facilitator is that the information should be understandable, according to the existing guidelines, and well-dosed. Additional information should be hyperlinked.

As barriers, several technical aspects may have negative influences on the use of the app according to our participants. Extensive battery and memory use of the smartphone are considered barriers. The second important barrier mentioned by the participants was sending frequent push notifications.

Comparison With Prior Work

Previous qualitative health studies on mHealth and eHealth in obstetrical health care mainly investigated (personalized) text messages [27,40-42] or Internet-based programs [35,43]. Most of our findings are comparable to these studies, especially the interactive and personalized aspects; our participants emphasize that a personalized tool which provides only relevant and specific information for the user is a very strong facilitator for the use of the tool [35,40-43]. Tripp et al [17] also showed that apps with interaction between the app and the user were shown to be the most popular kind of apps in obstetrical care. Furthermore, findings from the qualitative research of Naughton et al [27] on attitudes toward text message smoking cessation support suggests that maximizing personalization and personal relevance can increase the value of text message support and reduce the risk of disengagement.

Since the main purpose of our app is to provide detailed information and advice on work adjustments in certain specific work-related risks in pregnancy, the personal and interactive aspects of the app could be of strong positive influence for our app.

The fact that apps are viewed as faster and easier in searching for information and the information is always available at every place and every time has been pointed out in previous research as well [41,44].

Feedback from the app to the user is a complex outcome in our results since it can potentially be a strong facilitator; however, there exists a delicate balance between important stimulating reminders of advice and frequent uncalled-for push notifications, which can be experienced as a barrier. This delicate balance has also been recognized in previous studies. Two studies reported that text messages could stimulate positive behavioral changes [40,42], and one study reported that even more frequent messages would be appreciated [30]. On the other hand, our participants expressed frequent messages as a point of concern. These concerns are in line with the results in the study by Dennison et al [44]. Reminders of advice are well accepted and considered useful, which is also supported by other studies [30,40,42,44]. The mixed method qualitative study of Knight-Agarwall [30] showed participants using an app to monitor gestational weight gain wanted pop-up messages as a reminder to undertake certain activities [30].

Strengths and Limitations

A strength of our study is the proper qualitative health method we used for the focus group meetings and the analysis of the data. Furthermore, both focus group meetings were facilitated by the same experienced facilitator. Both meetings were audiotaped and fully transcribed and were independently coded by 2 researchers, resulting in negligible intercoder variance. In

case of discussion on the interpretation of the codes, a third researcher was involved in the process.

Other strengths of our study are that in both focus group meetings all different stakeholders were represented, which created the aimed interaction between stakeholders. Such group dynamic and diversity stimulated a broad view on the topics discussed and did not prevent reaching consensus on important issues [45].

In line with previous literature on qualitative health research, our number of participants is considered sufficient [46]. Besides the sufficient number of participants, the results in both meetings were comparable and we therefore believe we have achieved data saturation.

A disadvantage of conducting focus group meetings with different stakeholders together may be reluctance to be completely honest because of possible hierarchy between the different participants. Therefore, this method might lead to potential loss of important information. The fact that in our study we chose to mix professionals and pregnant women could be considered a limitation to our study. Since the discussed subject in our study, the development of an app, is not a delicate matter and is in the best interest of all the participants, we decided that this risk was small and therefore acceptable. The active participation of all participants during the meetings also indicated no reluctance of participants to share their opinions and experiences.

Conclusion

We have identified clear facilitators and barriers for the use of an app in obstetrical care. The correct content and dosage of interaction with the end user is a complex aspect to consider in the development of an app. These outcomes will contribute to the further developmental phases of an app. The results of this study are especially of interest to medical professionals in several areas who aim to develop an app implementing a guideline or evidence-based information in practice.

In future research we aim to evaluate the usability of the app in a think-aloud study among pregnant women. Subsequently we aim to evaluate the effectiveness of the app in a controlled trial.

Acknowledgments

We would like to thank all the stakeholders who participated in the focus group meetings and the regional network of birth care. This pilot study has received funding from ZonMW, the Netherlands Organization for Health Research and Development.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Topics and statements of the focus group meetings.

[PDF File (Adobe PDF File), 32KB - resprot_v6i8e163_app1.pdf]

References

1. Eurostat employment rates for selected population groups, 2004-2015. URL: http://eceuropa.eu/eurostat/statistics-explained/images/3/37/Employment_rates_for_selected_population_groups%2C_2005%E2%80%932015_%28%25%29_YB16_III.png [accessed 2017-08-16] [WebCite Cache ID 6slx36JOu]
2. Central Statistics Office, the Netherlands. Employment participation. URL: http://statlinecbs.nl/StatWeb/publication/?VW=T&DM=SLnl&PA=71738ned&LA=nl [accessed 2016-11-01]
3. Salihu HM, Myers J, August EM. Pregnancy in the workplace. Occup Med (Lond) 2012 Mar;62(2):88-97. [doi: 10.1093/occmed/kqr198] [Medline: 22355087]
4. Katz VL. Work and work-related stress in pregnancy. Clin Obstet Gynecol 2012 Sep;55(3):765-773. [doi: 10.1097/GRF.0b013e318253b192] [Medline: 22828109]
5. Runge SB, Pedersen JK, Svendsen SW, Juhl M, Bonde JP, Nybo AA. Occupational lifting of heavy loads and preterm birth: a study within the Danish National Birth Cohort. Occup Environ Med 2013 Nov;70(11):782-788. [doi: 10.1136/oemed-2012-101173] [Medline: 23839660]
6. Bonde JP, Jørgensen KT, Bonzini M, Palmer KT. Miscarriage and occupational activity: a systematic review and meta-analysis regarding shift work, working hours, lifting, standing, and physical workload. Scand J Work Environ Health 2013 Jul;39(4):325-334 [FREE Full text] [doi: 10.5271/sjweh.3337] [Medline: 23235838]
7. Vrijkotte TGM, van der Wal MF, van Eijsden M, Bonsel GJ. First-trimester working conditions and birthweight: a prospective cohort study. Am J Public Health 2009 Aug;99(8):1409-1416. [doi: 10.2105/AJPH.2008.138412] [Medline: 19542045]
8. Snijder CA, Brand T, Jaddoe V, Hofman A, Mackenbach JP, Steegers EAP, et al. Physically demanding work, fetal growth and the risk of adverse birth outcomes. The Generation R Study. Occup Environ Med 2012 Aug;69(8):543-550. [doi: 10.1136/oemed-2011-100615] [Medline: 22744766]

9. van Beukering MDM, van Melick MJGJ, Mol BW, Frings-Dresen MHW, Hulshof CTJ. Physically demanding work and preterm delivery: a systematic review and meta-analysis. Int Arch Occup Environ Health 2014 Nov;87(8):809-834. [doi: 10.1007/s00420-013-0924-3] [Medline: 24390632]
10. van Beukering M. Cooperation document Occupational Physicians—Obstetricians. URL: https://www.nvab-online.nl/sites/default/files/bestanden-webpaginas/NVAB_-_Samenvatting_Samenwerking_bedrijfsartsen-verloskundigen_07_febr_2012.pdf [accessed 2017-08-16] [WebCite Cache ID 6slxIJXD1]
11. Hooftman WVDBS, van den Bossche SNJ. Pregnancy and work: counseling, measures and sick leave. URL: http://www.monitorarbeidtno.nl/dynamics/modules/SPUB0102/viewphp?pub_Id=100111&att_Id=4911 [accessed 2016-11-01]
12. van Beukering M, Brand T. [Gezond werken tijdens de zwangerschap: Wat leveren preventieve maatregelen op? In het kader van arbocuratieve samenwerking tijdens de zwangerschap]. Literature research ZonMw 2014.
13. van Beukering M. Practice guideline: pregnancy, postpartum period, and work: advice and guidance by the occupational physician. URL: https://www.nvab-online.nl/sites/default/files/bestanden-webpaginas/Guideline_Pregnancy_Postpartum_Period_and_Work.pdf [accessed 2017-08-16] [WebCite Cache ID 6slxSRaqM]
14. World Health Organization. New horizons for health through mobile technologies. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2017-08-16] [WebCite Cache ID 6slxWgC7y]
15. Thomairy NA, Mummaneni M, Alsalamah S, Moussa N, Coustasse A. Use of smartphones in hospitals. Health Care Manag (Frederick) 2015;34(4):297-307. [doi: 10.1097/HCM.0000000000000080] [Medline: 26506291]
16. General Statistics Office, the Netherlands. Internet access, use, and facilities. URL: http://statlinecbs.nl/Statweb/publication/?DM=SLNL&PA=83429ned&D1=5&D2=0,3-6&D3=0&D4=a&VW=T [accessed 2016-10-01]
17. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
18. Carral F, Ayala MDC, Fernández JJ, González C, Piñero A, García G, et al. Web-based telemedicine system is useful for monitoring glucose control in pregnant women with diabetes. Diabetes Technol Ther 2015 May;17(5):349-354. [doi: 10.1089/dia.2014.0223] [Medline: 25629547]
19. Homko CJ, Deeb LC, Rohrbacher K, Mulla W, Mastrogiannis D, Gaughan J, et al. Impact of a telemedicine system with automated reminders on outcomes in women with gestational diabetes mellitus. Diabetes Technol Ther 2012 Jul;14(7):624-669 [FREE Full text] [doi: 10.1089/dia.2012.0010] [Medline: 22512287]
20. Homko CJ, Santamore WP, Whiteman V, Bower M, Berger P, Geifman-Holtzman O, et al. Use of an internet-based telemedicine system to manage underserved women with gestational diabetes mellitus. Diabetes Technol Ther 2007 Jun;9(3):297-306. [doi: 10.1089/dia.2006.0034] [Medline: 17561800]
21. Pérez-Ferre N, Galindo M, Fernández MD, Velasco V, Runkle I, Martín RP, et al. The outcomes of gestational diabetes mellitus after a telecare approach are not inferior to traditional outpatient clinic visits. Int J Endocrinol 2010 [FREE Full text] [doi: 10.1155/2010/386941] [Medline: 20628517]
22. Pollak KI, Alexander SC, Bennett G, Lyna P, Coffman CJ, Bilheimer A, et al. Weight-related SMS texts promoting appropriate pregnancy weight gain: a pilot study. Patient Educ Couns 2014 Nov;97(2):256-260 [FREE Full text] [doi: 10.1016/j.pec.2014.07.030] [Medline: 25153313]
23. Soltani H, Duxbury AMS, Arden MA, Dearden A, Furness PJ, Garland C. Maternal obesity management using mobile technology: a feasibility study to evaluate a text messaging based complex intervention during pregnancy. J Obes 2015;2015:814-830 [FREE Full text] [doi: 10.1155/2015/814830] [Medline: 25960889]
24. Herring SJ, Cruice JF, Bennett GG, Rose MZ, Davey A, Foster GD. Preventing excessive gestational weight gain among African American women: A randomized clinical trial. Obesity (Silver Spring) 2016 Jan;24(1):30-36 [FREE Full text] [doi: 10.1002/oby.21240] [Medline: 26592857]
25. Pollak KI, Lyna P, Bilheimer A, Farrell D, Gao X, Swamy GK, et al. A pilot study testing SMS text delivered scheduled gradual reduction to pregnant smokers. Nicotine Tob Res 2013 Oct;15(10):1773-1776 [FREE Full text] [doi: 10.1093/ntr/ntt045] [Medline: 23569007]
26. Fujioka N, Kobayashi T, Turale S. Short-term behavioral changes in pregnant women after a quit-smoking program via e-learning: a descriptive study from Japan. Nurs Health Sci 2012 Sep;14(3):304-311. [doi: 10.1111/j.1442-2018.2012.00702.x] [Medline: 22950611]
27. Naughton F, Prevost AT, Gilbert H, Sutton S. Randomized controlled trial evaluation of a tailored leaflet and SMS text message self-help intervention for pregnant smokers (MiQuit). Nicotine Tob Res 2012 May;14(5):569-577. [doi: 10.1093/ntr/ntr254] [Medline: 22311960]
28. Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth platform for pregnancy on nutrition and lifestyle of the reproductive population: a survey. JMIR Mhealth Uhealth 2016;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
29. Abroms LC, Johnson PR, Heminger CL, Leavitt LE, Schindler-Ruwisch JM, Bushar JA. Quit4baby: results from a pilot test of a mobile smoking cessation program for pregnant women. JMIR Mhealth Uhealth 2015;3(1):e10 [FREE Full text] [doi: 10.2196/mhealth.3846] [Medline: 25650765]

30. Knight-Agarwal C, Davis DL, Williams L, Davey R, Cox R, Clarke A. Development and pilot testing of the Eating4two mobile phone app to monitor gestational weight gain. JMIR Mhealth Uhealth 2015;3(2):e44 [FREE Full text] [doi: 10.2196/mhealth.4071] [Medline: 26048313]
31. Hirst JE, Mackillop L, Loerup L, Kevat DA, Bartlett K, Gibson O, et al. Acceptability and user satisfaction of a smartphone-based, interactive blood glucose management system in women with gestational diabetes mellitus. J Diabetes Sci Technol 2015 Jan;9(1):111-115 [FREE Full text] [doi: 10.1177/1932296814556506] [Medline: 25361643]
32. Bot M, Milder IEJ, Bemelmans WJE. Nationwide implementation of Hello World: a Dutch email-based health promotion program for pregnant women. J Med Internet Res 2009;11(3):e24 [FREE Full text] [doi: 10.2196/jmir.1183] [Medline: 19674957]
33. Choi J, Lee JH, Vittinghoff E, Fukuoka Y. mHealth physical activity intervention: a randomized pilot study in physically inactive pregnant women. Matern Child Health J 2016 May;20(5):1091-1101. [doi: 10.1007/s10995-015-1895-7] [Medline: 26649879]
34. Elwyn G, Kreuwel I, Durand MA, Sivell S, Joseph-Williams N, Evans R, et al. How to develop Web-based decision support interventions for patients: a process map. Patient Educ Couns 2011 Feb;82(2):260-265. [doi: 10.1016/j.pec.2010.04.034] [Medline: 20627644]
35. Shorten A, Fagerlin A, Illuzzi J, Kennedy HP, Lakehomer H, Pettker CM, et al. Developing an Internet-based decision aid for women choosing between vaginal birth after cesarean and planned repeat cesarean. J Midwifery Womens Health 2015;60(4):390-400. [doi: 10.1111/jmwh.12298] [Medline: 26059075]
36. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
37. Strauss AL. Qualitative analysis for social scientists. Cambridge: Cambridge University Press; 1987.
38. Boeije H. A purposeful approach to the constant comparative method in the analysis of qualitative interviews. Qual Quant 2002;36(4):391-409 [FREE Full text]
39. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 8;320(7227):114-116 [FREE Full text] [Medline: 10625273]
40. Soltani H, Furness PJ, Arden MA, McSeveny K, Garland C, Sustar H, et al. Women's and midwives' perspectives on the design of a text messaging support for maternal obesity services: an exploratory study. J Obes 2012;2012:835464 [FREE Full text] [doi: 10.1155/2012/835464] [Medline: 22900153]
41. Naughton F, Jamison J, Sutton S. Attitudes towards SMS text message smoking cessation support: a qualitative study of pregnant smokers. Health Educ Res 2013 Oct;28(5):911-922 [FREE Full text] [doi: 10.1093/her/cyt057] [Medline: 23640985]
42. Kharbanda EO, Vargas CY, Castaño PM, Lara M, Andres R, Stockwell MS. Exploring pregnant women's views on influenza vaccination and educational text messages. Prev Med 2011 Jan;52(1):75-77. [doi: 10.1016/j.ypmed.2010.10.009] [Medline: 21047526]
43. Berg M, Adolfsson A, Ranerup A, Sparud-Lundin C, University of Gothenburg Centre for Person-Centred Care. Person-centered Web support to women with type 1 diabetes in pregnancy and early motherhood--the development process. Diabetes Technol Ther 2013 Jan;15(1):20-25. [doi: 10.1089/dia.2012.0217] [Medline: 23297670]
44. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
45. Mack N, Woodsong C, MacQueen KM, Guest G, Namely E. Qualitative research methods: a data collector's field guide.: Family Health International URL: https://www.fhi360.org/sites/default/files/media/documents/Qualitative%20Research%20Methods%20-%20A%20Data%20Collector's%20Field%20Guide.pdf [accessed 2017-08-16] [WebCite Cache ID 6saS5X1gy]
46. Pope C, Mays N. Qualitative Research in Health Care, 3rd Edition. Oxford: Blackwell Publishing; 2006.

Edited by G Eysenbach; submitted 27.12.16; peer-reviewed by T Toro-Ramos, J Brinkel; comments to author 16.06.17; revised version received 01.08.17; accepted 06.08.17; published 22.08.17.

Please cite as:
Velu AV, van Beukering MDM, Schaafsma FG, Frings-Dresen MHW, Mol BWJ, van der Post JAM, Kok M
Barriers and Facilitators for the Use of a Medical Mobile App to Prevent Work-Related Risks in Pregnancy: A Qualitative Analysis
JMIR Res Protoc 2017;6(8):e163
URL: http://www.researchprotocols.org/2017/8/e163/
doi:10.2196/resprot.7224
PMID:28830851

Short Paper

Engaging Adolescents to Inform the Development of a Mobile Gaming App to Incentivize Physical Activity

Lizzy Pope[1*], RD, PhD; Bernice Garnett[1*], MPH, ScD; Marguerite Dibble[2]

[1]University of Vermont, Burlington, VT, United States

[2]Game Theory, Burlington, VT, United States

[*]these authors contributed equally

Corresponding Author:
Bernice Garnett, MPH, ScD
University of Vermont
533 Waterman Building
85 South Prospect St
Burlington, VT, 05405
United States
Phone: 1 802 656 2187
Email: bgarnett@uvm.edu

Abstract

Background: Involving youth in the development of a mobile game designed to increase physical activity may increase relevancy and adoption.

Objective: To share the development process used to create a gaming app aimed at incentivizing physical activity in high school students.

Methods: Five focus groups were conducted with high school students (N=50) to understand gaming behaviors. A subset of students from the focus groups chose to complete a Web-based survey (N=10). Four different versions of gaming artwork and concept design based on student input were pilot tested (N=35), and group consensus building determined the direction of the game. The 4 game versions differed in their artwork style and gaming concept with some requiring competition versus cooperation, or being more individual versus team based. Group consensus building meant that all artwork and game concept options were displayed at the front of a classroom. Students could then vote for their top artwork and concept choices by putting stickers on the top 1 or 2 artwork and concept options that they liked best. Once all votes were cast, investigators discussed the voting results with students, and brainstormed ways to incorporate popular aspects of the 3 "losing" artwork and game concepts into the winning ideas.

Results: Focus group transcripts were analyzed for common themes. Artwork and gaming concept-voting data was tallied at the time of voting to share with students in real time. Focus groups and survey results revealed important themes for a successful gaming app: (1) competition, (2) balanced in-game rewards, (3) accessibility, and (4) aesthetic features. Consensus voting indicated the popularity of a collaborative competitive content design (35/66, 53%) and playful art (27/71, 38%).

Conclusions: To ensure saliency and effectiveness of game-based physical activity interventions, youth need to be included in design and implementation. Furthermore, the unique preferences and social constructs of high school students need to be considered during intervention development.

(JMIR Res Protoc 2017;6(8):e161) doi:10.2196/resprot.8113

KEYWORDS

adolescents; qualitative research; mHealth; physical activity

Introduction

Among American adolescents there continues to be high rates of those who are overweight and obese. This is a concern for health practitioners. According to the Centers for Disease Control and Prevention (CDC), 34.5% of youth age 12 to 19 years old were overweight or obese in 2011 to 2012 [1]. Addressing the current prevalence of overweight/obesity in adolescents is paramount as excess weight in adolescence often translates to excess weight in adulthood [2]. The positive effects

of physical activity on health outcomes, learning outcomes, and social outcomes have been elucidated in a variety of previous research studies [3]. However, physical activity declines steeply from childhood to adolescence, with 27.1% of high school students nationally meeting the CDC guideline of 60 minutes of moderate/vigorous physical activity (MVPA) per day in 2013 [4]. In one recent study, adolescents only obtained an average of 39.4 minutes of MVPA per day [5].

Previous research has indicated that not having enough time and wanting to do other things are two of the most common barriers to physical activity for high school students [6]. Monetary incentives may be very effective at encouraging physical activity in both college and elementary school students as well as adults, and therefore help address the time and attention barrier to physical activity cited by adolescents [7-9]. Very few interventions have examined how to increase physical activity school wide for high school students, or ways that the behavioral theories behind incentives can be leveraged without the use of monetary incentives. Incentives inherent in a mobile game may be a feasible way to encourage high school students to be more active, and incur the positive benefits of physical activity on health, academic, and social outcomes. Using a gamification model where real-world physical activity translates into immediate “rewards” in a mobile game may successfully incentivize physical activity in high school students. Because games involve accumulating points, moving up levels, or leaderboard social recognition, they inherently provide an incentive to continue doing whatever activity results in more points, more progress, or more prowess [10]. As many high school students are familiar with and enjoy games on either their mobile devices, tablets, or gaming systems, using a gaming context to promote physical activity may be novel, age-appropriate, and widely scalable. The objective of this paper was to provide an overview of the development process used to design a gaming app that incentivized real-world physical activity with in-game rewards.

Methods

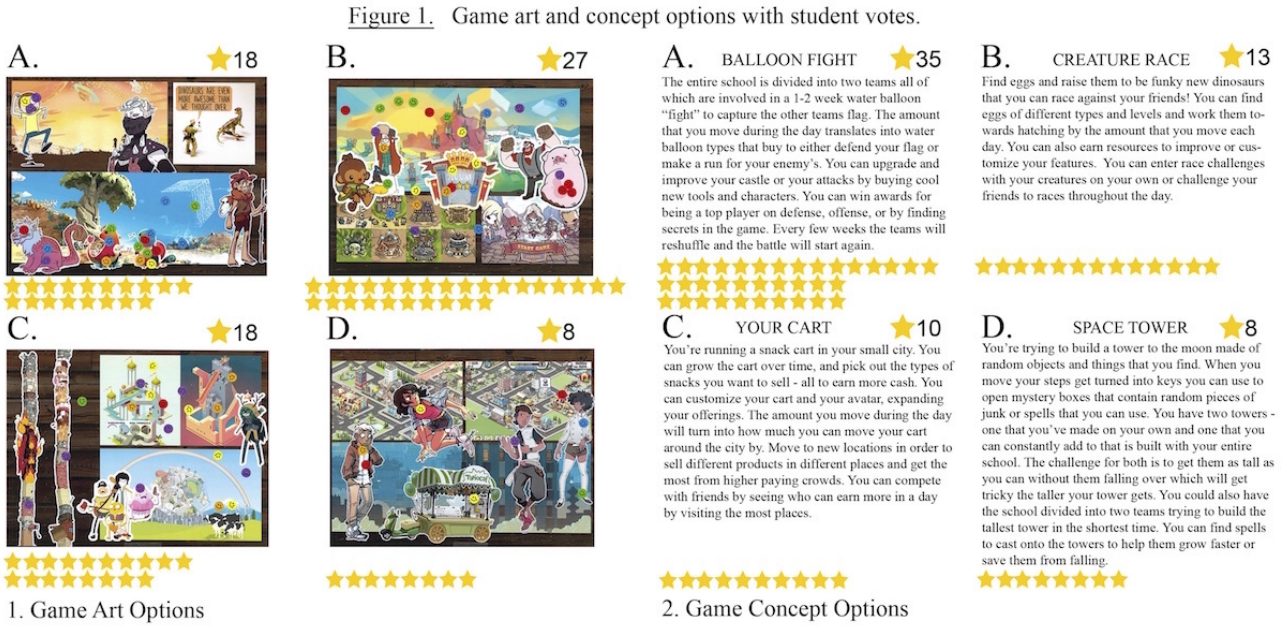
Recruitment and Sample

To design a physical activity gamification program for diverse adolescents; partnering with youth in the design and implementation of the program is necessary to ensure relevancy and program ownership. Youth Participatory Action Research (YPAR) is a methodology in which youth are recruited as research partners to collaboratively assess the needs/assets of their community, inform program design and implementation, as well as assist in analysis and dissemination activities [11]. To this end, YPAR served as an overarching theoretical and methodological framework for our program design and evaluation. The overall objectives of the surveys and focus groups conducted were to: (1) understand the current gaming behaviors and attitudes of the target population, (2) solicit input on barriers and facilitators of their physical activity patterns, and (3) gauge interest and feasibility in integrating physical activity into a Web-based gaming app.

Sophomore and junior students from 5 teacher advisory periods were recruited to participate in focus groups and surveys in a diverse high school in the Northeast (482/1006, 47.91% free/reduced lunch; 362/1006, 35.98% non-White; 181/1006, 17.99% English language learners). A letter was sent home to parents where they could opt out of having their student participate in the study. Students also provided verbal consent to participate in focus groups, and had the option to decline to participate. The University of Vermont’s Committee on Human Research in the Behavioral and Social Sciences approved the study. A total of 5 focus groups were conducted with sophomore and junior students during the fall of 2015 (N=50) to understand gaming behaviors and preferences. School staff and researchers decided on 5 focus groups, as it was thought that 5 groups would allow for ample exposure to students’ varying ideas and perspectives without disrupting the normal school schedule. Focus groups were conducted during the 30-minute teacher advisory period that sophomores and juniors attend daily. The focus groups were led by one of the researchers and the game designer, while a second researcher took notes and recorded responses. Each focus group followed the same script of questions, while also allowing follow-up on new ideas raised by each particular group. The focus groups averaged 10 students each. At the conclusion of the focus groups, all participating students were emailed a 10-question survey that contained similar questions to the focus groups. This survey addressed the types of games and physical activity that students engaged in at school and at home, and was developed specifically for the study. It was thought that students who may have been too shy to speak up during the focus groups could use the survey as a chance to share their thoughts. A subset of 10 students filled out the parallel qualitative survey.

After the initial focus groups, the research and game design team incorporated student input, theory, and best practices from the gaming and behavioral economics fields to design 4 game content schemas and 4 artwork styles for the students to vote on separately using consensus voting techniques. During the spring of 2016, a total of 35 students from the original 5 teacher advisory periods were assembled to vote on artwork and content gaming schema. Each student was given 4 colored stickers, 2 to place on their favorite artwork panels and 2 to place on their favorite content gaming schemas [12]. Both the artwork and content gaming schema had 4 different options for the students to choose from (see Figure 1).

Figure 1. Game art and concept options with student votes.



Game concepts were developed based on previous research on game design and behavioral economics. Each concept was designed to capitalize on principles in basic psychology known to increase motivation, including reinforcement, challenge, cooperation, competition, identity, and surprise [13]. The game concepts were also designed to provide intermittent reinforcement, an incentive strategy known to help guard against behavior extinction [9,14]. Finally, the game concepts were designed to leverage behavioral economic theory by appealing to decision-making heuristics like loss aversion (the tendency to avoid experiencing losses) and the endowment effect (the higher value assessment of things you own vs things you do not own) that humans subconsciously use frequently [10]. Although all 4 game concepts emphasized various gaming engagement techniques, they each incorporated a variety of motivational elements, and differed mostly in the scenario and mechanics proposed. At the conclusion of the voting, researchers discussed the results with the students, and gathered information on elements of each style that students liked.

Analysis

Focus groups were audio recorded and transcribed. Based on theory, previous research, the focus group protocol, and the use of key words in context (eg, the use of in vivo coding) a codebook was developed to identify salient themes. Transcripts were analyzed using this codebook. We followed a coding schema developed by Saldaña [15] in which first cycle codes, including descriptive codes that were used to identify the basic topic of a passage as well as process codes that refer to conceptual action in the data, such as game aesthetics, were applied followed by second cycle codes, or pattern codes, to understand relationships among codes. Transcriptions were coded based on thematic analysis to uncover consistent and major trends as they related to gaming and physical activity attitudes and behavior [16]. Results from the 10 qualitative survey responses were analyzed using the codebook generated from the focus group analysis. Voting was tallied across the artwork and gaming content to determine the majority vote.

Results

Qualitative Results

Focus groups and survey results with youth revealed several themes important for game design: (1) competition, (2) balanced in-game rewards, (3) accessibility, and (4) aesthetic features. Textbox 1 illustrates the major themes from the focus group and survey results aligned with salient quotes to further illustrate youth voice.

Youth described the importance of having an element of competition, particularly a competition with friends or within the school community incorporated into the game design. Additionally, to ensure fairness within the game, youth described various game design features to ensure "balanced in-game rewards" to motivate and reward students who were not as physically active as others. Finally, youth described the importance of accessibility, both technologic and content accessibility, to encourage multiple segments of their demographic to play and continue playing a game. Students also mentioned the importance of "cool graphics" and other aesthetics, specifically cited as critical when trying to encourage someone to play a new game.

Consensus Voting

Consensus voting of students indicated the popularity of a collaborative competitive content design was 53% (35/66) and playful art was 38% (27/71), compared with the other content and art design options presented.

Textbox 1. Results and salient quotes from the focus groups related to gaming content and design.

Focus group themes: game design

1. Competitions:
   - "Like a school wide capture the flag..."
   - "If you can create a game that somehow tracks your steps or activity and then you can compete with that and it can motivate"
2. Balanced in-game rewards:
   - "Progress to be a part of a competition so people don't get discouraged"
   - "There are a lot of games where points can earn you special cash where you can buy special things with and like if you run more you could get special cash"
3. Accessibility:
   - "Any free app that you can get access to that was simple and obnoxious..."
   - "Compatibility with a lot of difference devices"
   - "You have to have something that is really accessible to all, like you could have someone that has been sedentary their entire life and then you could have someone like the captain of the varsity track team and they are in the same school and competing in the same game you are making so you have to have something that...it should probably be targeted to kids who don't do anything...rewarding progress"
4. Aesthetic features:
   - "If it has a nice looking icon and the name sounds like a real app and not some that is fake or suspicious"
   - "I always think it's really cool when you can customize because like there are games that you do what they design and there are games that you can customize to how you like it which is pretty awesome"

Discussion

Importance of Youth Engagement in Intervention Design

To ensure that mHealth and eHealth technologies are effective and resonate with the target population, efforts to incorporate the perspectives, attitudes, and ideas of the target demographic should be solicited and embedded within mobile technologies to improve health. Explicitly incorporating youth voice in the design and construction of eHealth and mHealth interventions is a nascent applied research area [17,18]. Although there are many health and fitness apps available, very few capitalize on effective gaming strategies and behavioral theory [19]. The development process of the current game was designed to use established mixed-methods research, and determine which gaming and health promotion elements were most appealing to an adolescent population, an area of research that has not been widely explored previously. The successful artwork and game concept themes identified are familiar to game designers [19], but not commonly incorporated into physical activity promotion interventions for adolescents, which are themselves sparse and largely consist of curriculum development [20]. Our formative research indicates that gamification is of great interest to adolescents and could be used to add the important element of "fun" to behavior-change interventions. If an intervention is more pleasurable to engage with, it is much more likely that it can be effective in the long-term.

Engaging our target demographic population resulted in important directions for game design, artwork, and integration of physical activity into the game interface. Youth described the importance of aesthetics and accessibility that would encourage them to begin playing a new game, and then provided concrete examples of game design features that would keep them actively engaged and motivated in a physical activity–integrated gamification interface. The 4 themes that emerged from our focus group work helped drive intervention development in ways that were both predictable and more unexpected. For example, the research team was not surprised that competition, rewards, accessibility, and aesthetic features were the predominant themes to emerge from the focus groups. However, the specific manifestations of certain thematic elements suggested by the target audience were often innovative or surprising. The research team believed that an edgy, abstract style was the most attractive artwork style, yet youth overwhelmingly voted for the playful artwork style, illustrating the importance of gathering youth feedback on all aspects of intervention development. Furthermore, youth really wanted to be rewarded for not just goal achievement, but improvement and progress toward goal achievement. Research in business and health has indicated that rewarding progress is a very effective way to keep people on track with their goals, as small rewards for improvement or progress prevent discouragement and motivate continued behavior completion [21,22]. Therefore, it was very interesting that the adolescents could identify progress-based incentives as an important aspect of any game.

In addition to wanting to be rewarded for progress, multiple students also suggested a "capture the flag" style game play design, something that the research team may not have given much thought to without youth involvement. The interest in integrating team-based competition into a game concept may reflect the dominant nature of the social-emotional reward style in adolescents. More so than adults or children, neuroscience research indicates that adolescents' neural reward system

responds to social stimuli [23,24]. The social nature of a team-based competitive game where adolescents collaborate with friends or classmates to work toward a common goal may be particularly appealing. It has also been shown that during adolescence there is a shift from self-oriented behavior to more prosocial behavior, which further illuminates why a competitive team-based gaming concept would be popular [24]. This shift to more prosocial decision-making and ability to take the perspective of another person may partly explain why accessibility and fairness were important gaming concepts for focus group participants. Adolescents wanted peers of all fitness levels to be able to succeed in the game and they also wanted rewards for progress, indicating the importance of allowing even nonathletes to succeed in a physical activity–oriented game. Additionally, adolescents are known to be risk takers with their decision-making [23-25]. This preference for risk may result in wanting to play a game where you're part of a team and there is social pressure to perform a certain behavior so that your team wins.

The mixed-methods structure of the proposed intervention is novel and timely. Mixed-methods research helps provide a more tailored and efficacious intervention, as well as a more complete picture of how intervention outcomes arise. The YPAR method gives the intervention the best chance of success by involving future participants in the development of an engaging, relevant, and appropriate intervention. The YPAR method also allowed us to cultivate intervention champions organically. By involving students throughout the game development year, we began the process of building excitement for the gaming intervention and buy-in to the app.

Strengths and Limitations

This study had several strengths. First, this study included gathering feedback from a diverse group of adolescents at a public high school. Secondly, the multiple modalities used to collect input and feedback also greatly enhanced the study by allowing students to express their opinions in a variety of ways. Finally, the ability of our game design team to talk gaming specifics with the adolescents, and our research team to systematically collect qualitative data both increased the usefulness and rigor of our results. Limitations of the study include the rather small sample size, and the inclusion of only sophomores and juniors. Future work could examine differences in using YPAR methodology in larger samples and age ranges.

Acknowledgments

This study was funded through a USDA-HATCH grant, as well as a University of Vermont Ventures grant.

Conflicts of Interest

None declared.

References

1. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. JAMA 2014;311:806-814. [doi: 10.1001/jama.2014.732] [Medline: 24570244]
2. Guo S, Roche A, Chumlea W, Gardner J, Siervogel R. The predictive value of childhood body mass index values for overweight at age 35 y. Am J Clin Nutr 1994;59:810-819. [Medline: 8147324]
3. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act 2010;7:40 [FREE Full text] [doi: 10.1186/1479-5868-7-40] [Medline: 20459784]
4. Kann L, Kinchen S, Shanklin SL, Flint KH, Kawkins J, Harris WA, Centers for Disease ControlPrevention (CDC). Youth risk behavior surveillance--United States, 2013. MMWR Suppl 2014;63:1-168. [Medline: 24918634]
5. Carlson JA, Schipperijn J, Kerr J, Saelens BE, Natarajan L, Frank LD, et al. Locations of physical activity as assessed by GPS in young adolescents. Pediatrics 2016;137. [doi: 10.1542/peds.2015-2430] [Medline: 26647375]
6. Tergerson J, King K. Do perceived cues, benefits, and barriers to physical activity differ between male and female adolescents? Journal of school health. J Sch Health 2002;72:374-380. [Medline: 12557633]
7. Pope L, Harvey J. The efficacy of incentives to motivate continued fitness-center attendance in college first-year students: a randomized controlled trial. J Am Coll Health 2014;62:81-90. [Medline: 24456510]
8. Garnett B, Becker K, Vierling D, Gleason C, DiCenzo D, Mongeon L. A mixed-methods evaluation of the before-school incentive-based physical activity programme. Health Educ J 2016;76:89-101. [doi: 10.1177/0017896916652432]
9. Jeffery R. Financial incentives and weight control. Prev Med 2013;55(Suppl):S61-S70. [doi: 10.1016/j.ypmed.2011.12.024] [Medline: 2244800]
10. King D, Greaves F, Exeter C, Darzi A. 'Gamification': influencing health behaviours with games. J R Soc Med 2013;106:76-78. [doi: 10.1177/0141076813480996] [Medline: 23481424]
11. Ozer E, Douglas L. The impact of participatory research on urban teens: an experimental evaluation. Am J Community Psychol 2013;5:66-75. [Medline: 22875686]
12. McMurray A. Three decision-making aids: brainstorming, nominal group, and Delphi technique. J Nurs Staff Dev 1994;10:62-65. [Medline: 8176489]
13. Dibble M. TEDx. 2014 Oct 13. Uncovering the Ingredients of FUN URL: https://www.youtube.com/watch?v=IiHQvSG1plM [accessed 2017-08-10] [WebCite Cache ID 6scEFFpGS]

14. Skinner B. Science and Human Behavior. New York, NY: Free Press; 1953.
15. Saldaña J. The Coding Manual for Qualitative Researchers. Thousand Oaks, California: Sage Publications; 2013.
16. Boyatzis R. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, California: Sage Publications; 1998.
17. Hingle M, Nichter M, Medeiros M, Grace S. Texting for health: the use of participatory methods to develop healthy lifestyle messages for teens. J Nutr Educ Behav 2013;45:12-19. [doi: 10.1016/j.jneb.2012.05.001] [Medline: 23103255]
18. Mojica CM, Parra-Medina D, Yin Z, Akopian D, Esparza LA. Assessing media access and use among Latina adolescents to inform development of a physical activity promotion intervention incorporating text messaging. Health Promot Pract 2014;15:548-555 [FREE Full text] [doi: 10.1177/1524839913514441] [Medline: 24357863]
19. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. JMIR Serious Games 2014;2:e9 [FREE Full text] [doi: 10.2196/games.3413] [Medline: 25654660]
20. Dobbins M, DeCorby K, Robeson P, Husson H, Tirilis D. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6-18. Cochrane Database Syst Rev 2009:CD007651. [doi: 10.1002/14651858.CD007651] [Medline: 19160341]
21. Amabile T, Kramer S. The power of small wins. Harvard Business Review 2011;80:70-80.
22. Terry P. The advantage of progress-based incentives. Managed Care 2012;21:32-36. [Medline: 22876520]
23. Galvan A. Adolescent development of the reward system. Front Hum Neurosci 2010;4:6 [FREE Full text] [doi: 10.3389/neuro.09.006.2010] [Medline: 20179786]
24. Crone EA, Dahl RE. Understanding adolescence as a period of social-affective engagement and goal flexibility. Nat Rev Neurosci 2012;13:636-650. [doi: 10.1038/nrn3313] [Medline: 22903221]
25. Smith AR, Chein J, Steinberg L. Impact of socio-emotional context, brain development, and pubertal maturation on adolescent risk-taking. Horm Behav 2013;64:323-332 [FREE Full text] [doi: 10.1016/j.yhbeh.2013.03.006] [Medline: 23998675]

Abbreviations

CDC: Centers for Disease Control and Prevention
MVPA: moderate/vigorous physical activity
YPAR: youth participatory action research

Edited by G Eysenbach; submitted 31.05.17; peer-reviewed by D Thompson, J Bervoets; comments to author 15.07.17; revised version received 21.07.17; accepted 21.07.17; published 29.08.17.

Please cite as:
Pope L, Garnett B, Dibble M
Engaging Adolescents to Inform the Development of a Mobile Gaming App to Incentivize Physical Activity
JMIR Res Protoc 2017;6(8):e161
URL: http://www.researchprotocols.org/2017/8/e161/
doi:10.2196/resprot.8113
PMID:28851676

Protocol

Ascertaining the Value of Noninvasive Measures Obtained Using Color Duplex Ultrasound and Central Aortic Pressure Monitoring During the Management of Cerebral Arteriovenous Malformation Resection: Protocol for a Prospective, Case Control Pilot Study

Kathryn J Busch[1], DMU (Vasc), AMS; Hosen Kiat[1,2,3], MBBS, FRACP, FACP, FACC, FCCP, FCSANZ, FASNC

[1]Faculty of Medicine and Health Sciences, Macquarie University, Sydney, Australia

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

[3]Faculty of Medicine, Western Sydney University, Sydney, Australia

Corresponding Author:
Kathryn J Busch, DMU (Vasc), AMS
Faculty of Medicine and Health Sciences
Macquarie University
Suite 202, 2 Technology Place
Sydney, 2109
Australia
Phone: 61 298123683
Fax: 61 298123983
Email: kathrynjbusch@gmail.com

Abstract

Background: Dramatic hemodynamic changes occur upon removal of an arteriovenous malformation of the brain (bAVM) with a number of potentially serious perioperative complications, such as intracranial hemorrhage and venous occlusive hypertensive syndrome. As these complications largely occur in the postoperative inpatient period, a rapid, repeatable noninvasive investigation to serially monitor relevant intracranial hemodynamics may be of benefit. Though, transcranial Doppler (TCD) and transcranial color duplex (TCCD) are techniques used and available to provide hemodynamic measurements postoperatively, the time course of hemodynamic sequences following bAVM resection remains uncertain.

Objective: This is a prospective, case control pilot study conducted in participants having elective bAVM resection surgery.

Methods: Each participant will undergo a preoperative color duplex ultrasound (CDU) of the bilateral extracranial carotid arteries, a CDU of the circle of Willis including the bAVM vessels, and a central aortic pressure measurement, repeated daily, postoperatively, for a 2-week period.

Results: Patient accrual has commenced with anticipation of first results in 2018.

Conclusions: This protocol aims to strengthen the work of previous authors by providing documentation of the time course of hemodynamic changes following bAVM resection. The protocol is designed to determine whether noninvasive technology, including CDU imaging of the extracranial carotid and intracranial arteries in the form of TCCD along with central aortic pressure measurements, can determine whether there are any hemodynamically significant prognostic markers that may provide insight into the process of vessel remodeling, including insight into venous changes following bAVM resection.

(JMIR Res Protoc 2017;6(8):e173) doi:10.2196/resprot.7991

KEYWORDS

brain; transcranial; arteriovenous malformation; color duplex; hemodynamics; pressure

Introduction

Dramatic hemodynamic changes occur upon removal of an arteriovenous malformation of the brain (bAVM) [1-9]. Before the vasculature returns to normal, the velocity in the former large arteries and veins is markedly reduced concurrent with an elevation of arterial pressure, which is associated with a number of potentially serious postoperative complications, such as arterial-capillary-venous hypertensive (ACVH) syndromes

including hemorrhage, edema and vasospasm [10]. In our institution, by rigorously controlling blood pressure, postoperative hemorrhage has reduced from 4.4% to 1% [11].

Accordingly, as these complications largely occur in the postoperative inpatient period, a rapid, repeatable noninvasive investigation to serially monitor relevant intracranial hemodynamics may be of clinical and prognostic benefit.

Transcranial Doppler (TCD) and transcranial color duplex (TCCD) studies of cerebral arteriovenous malformations (bAVM) have established that the feeding arteries exhibit a high-velocity and low-pulsatility flow preoperatively, followed by a postoperative decrease in velocity and increase in pulsatility of the feeding arteries [2,12-21]. A recent systematic review indicated that the time course of hemodynamic sequences remains uncertain; the existing conclusions were of limited clinical value due to variations in pathology, methodology, and timing of measurements during the postoperative period, and that few studies (n=2) have used CDU applied transcranially (TCCD) to evaluate postoperative hemodynamic changes. The review also identified that the venous aspect of bAVMs using transcranial ultrasound remains largely unchartered, possibly due to technical limitations, and, to our knowledge, no study to date has used TCCD to assess both the venous components of a bAVM postoperatively using TCCD.

Proponents of TCCD suggest that in contrast to TCD, TCCD purportedly offers a higher degree of accuracy and reproducibility of velocity measurements enhanced by the ability to visually control sample volume placement [22,23], which can be problematic on TCD given that vessels change waveform characteristics following bAVM resection, making repeated measurements difficult and unreliable. Qualitative advantages of TCCD also include enhanced differentiation between feeding arteries and draining veins in patients with bAVM [23,24].

Further insight into the complex hemodynamics involved in bAVM patients is desirable, especially if it can be derived noninvasively.

Methods

Study Design

This is a prospective, case control pilot study conducted in participants having elective bAVM resection surgery. Each participant will undergo a preoperative CDU of the bilateral extracranial carotid arteries, a TCCD of the circle of Willis, including the bAVM vessels and a concurrent central aortic pressure measurement. This examination will be repeated daily for a 2- week period postoperatively or as long as the patient remains hospitalized between 2 PM and 4 PM to reduce diurnal variations. This protocol sequence will be referred to as the "noninvasive protocol."

The primary outcome measure is to use TCCD and central aortic pressure measurements to establish the time course of hemodynamic changes following bAVM resection, and to determine whether there are any hemodynamically significant prognostic markers that may provide insight into the process of vessel remodeling.

The secondary outcome measure is to determine whether TCCD can provide insight into venous changes following bAVM resection.

Macquarie University Human Research Ethics Committee approved this study (Medical Sciences; reference: 5201400098).

Participants

Eligibility criteria includes adults (>18 years) with willingness and cognitive ability to provide written and informed consent: a control group of healthy participants, a cohort of participants undergoing elective surgical resection with bAVM, and a cohort of participants undergoing elective craniotomies for tumor resection or cerebral aneurysm surgery.

Participants with a history of a psychological illness or other conditions, which may interfere with their ability to understand the study requirements, will be excluded. Participants expressing anxiety or unwillingness following initial consent will be further excluded. For the healthy volunteer cohort, pregnant women and individuals taking medication or hormonal treatment will be excluded.

Control Group

Participants (n=20) will undergo a CDU assessment of the extracranial carotid arteries and TCCD of the circle of Willis in accordance to the protocol. The participants will be comprised of healthy volunteers recruited from research, hospital, and clinic staff at Macquarie University Hospital, and relatives. The examination will be conducted once by a dedicated vascular research sonographer (KB), and then repeated 24 hours later under the same conditions for the purpose of calculating intraobserver measurement reproducibility. The control group will be selected to represent a distribution of ages and sexes.

Arteriovenous Malformation of the Brain Group

The bAVM group includes participants undergoing elective surgical resection of a bAVM. Participants will be recruited from the Department of Neurosurgery, Faculty of Medicine and Health Sciences, Macquarie University. Participants will be enrolled into the study following computed tomography confirmation of an operable bAVM. After the informed consent process has been completed, study participants will receive a study enrollment number and this will be recorded on all study documents.

The participants will undergo the noninvasive protocol by the research sonographer (KB). There will be no change in the participants' routine pre- or postoperative care. Routine digital subtraction angiography will be used to correlate TCCD findings.

Non–Arteriovenous Malformation of the Brain Group

The non-bAVM group includes participants undergoing elective intracranial aneurysm or tumor surgery. Participants will be recruited from the Department of Neurosurgery, Faculty of Medicine and Health Sciences, Macquarie University. Participants will be recruited to establish whether there exists a significant variance in hemodynamic measurements between the bAVM and non-bAVM group. By contrasting the 2 groups, it may assist in identifying hemodynamic changes of vessel

remodeling in bAVM participants as the hemodynamic changes in the non-bAVM group are anticipated to be less dramatic.

The participants will undergo the noninvasive protocol by the vascular research sonographer (KB). Although the noninvasive protocol will be adhered to, data measurements obtained in this group may be for a shorter duration postoperatively, according to varied patient recovery and earlier hospital discharge rates.

Imaging Protocol

The control and study participants will be positioned supine, with the head elevated to approximately 30°, typifying the position of postoperative bAVM patients in the intensive care unit.

The imaging component of the noninvasive protocol includes: (1) full CDU assessment of the extracranial carotid and vertebral arteries, and (2) TCCD assessment of the circle of Willis and cerebral veins, where possible.

Carotid and Vertebral Color Duplex Ultrasound Measurements

The core vascular research sonographer (KB) at Macquarie University Hospital will perform CDU measurements of the carotid and vertebral arteries using a Philips IU22 system and a 9-2 MHz linear transducer. Peak systolic and end diastolic velocity measurements of the distal common carotid, proximal external carotid artery, proximal and distal extracranial internal carotid, and the vertebral artery will be recorded.

In the presence of atherosclerotic disease, maximal peak systolic and end diastolic velocities will be recorded at appropriate locations with stenoses graded accordingly to published diagnostic criteria.

In the absence of carotid disease, peak systolic velocity measurements will be obtained at reproducible locations by convention. Velocity measurements will be obtained by placing the sample volume cursor center stream, parallel to the lumen, and using an angle of 60° for reproducibility of measurements.

Transcranial Color Duplex Measurements

The vascular research sonographer unit at Macquarie University Hospital will perform TCCD measurements using a Philips IU22, 5-1 MHz phased array transducer.

Although the American Society of Neuroimaging Practice Guidelines recommendation advocates assessing the entire circle of Willis, our protocol will assess the middle cerebral artery, anterior cerebral artery, internal carotid artery, and posterior cerebral artery using a transtemporal approach bilaterally. The site of maximal velocity will be recorded. The imaging protocol is based on our experience, whereby following bAVM surgery the evaluation of the entire circle of Willis using windows other than the transtemporal approach is impractical due to lowered patient tolerance levels and imposed restrictions of patient head movement in the intensive care unit.

Attempts to measure cerebral veins will be recorded and repeated using the visual reproducibility capability of TCCD. Normal reference values will be obtained by published criteria obtained from the literature and our control group.

Noninvasive Central Aortic Blood Pressure Measurements

The Uscom BP PLUS device derives central pressure measurements that correlate closely with those obtained invasively from a catheter.

This noninvasive blood pressure device estimates the central pressures from brachial cuff pressure fluctuations, using oscillometry to determine brachial systolic and diastolic pressures during deflation of the cuff. A second inflation, holds cuff pressure approximately 30 mm Hg above the brachial systolic pressure (ie, suprasystolic measurement) for approximately 10 seconds. The intra-arterial pressures in the brachial artery at the cuff measurement site then estimate the aortic pressure using a physics-based model of the left subclavian-to-brachial branch.

This device will be used for the noninvasive protocol at the conclusion of the imaging studies, to obtain a noninvasive pressure measurement otherwise derived from invasive catheterization.

Statistical Analysis

Inter- and Intraobserver Reliability

A Kappa weighted test will be performed on the 20 control participants to determine intraobserver agreement of measurements by comparing CDU of the carotid arteries and TCCD measurements of the circle of Willis 24 hours apart under the same conditions by the core vascular research sonographer. Intraoberserver reliability will be calculated.

Arteriovenous Malformation of the Brain and Non–Arteriovenous Malformation of the Brain Groups

Daily velocity and pulsatility measurements will be entered for each sampled vessel and plotted onto an excel spreadsheet against simultaneous daily pressure measurements. Appropriate statistical analysis will be performed according to perceived trends.

Results

This pilot study is in progress. The data collection has commenced and the findings are planned to be completed in 2018. The results will be submitted to a leading journal for publication.

Discussion

Rationale for Daily Noninvasive Monitoring Following Arteriovenous Malformation Resection

Given the intensive hemodynamic changes that occur following resection or embolization of a bAVM, there is a requirement for stringent blood pressure monitoring and an extended stay in the intensive care unit given the propensity for ACVH complications including edema and hemorrhage [10,11]. Despite progressive advancement in diagnostic armamentarium for hemodynamic measurements, the precise mechanisms for postoperative bAVM hemorrhagic complications remain

controversial [2-9,11]. In 1980, prior to TCD and TCCD, using pulsed echo Doppler flowmeters, Nornes and Grip commented:

> *Blood flow in veins has not received the attention it deserves in relation to disease in man. The dynamics of a bAVM involve venous factors of considerable significance. A bAVM hemorrhage is in principle a venous bleed caused by the increased pressure transmitted through the shunt. The arterial pulses are mediated through the shunt to the venous side. The mean velocity is at the level normally found in cerebral arteries during operations for other diseases, and considerably above what we consider normal intravenous velocities (unpublished data).* [25]

Since 1986, numerous TCD and later (1990) TCCD have studied bAVM hemodynamics post intervention [2,12-17,19-21,26], yet the temporal hemodynamic course following bAVM resection remains uncertain. This is due to variations in pathology, methodology, and timing of measurements during the postoperative period. Furthermore, the venous aspect following bAVM resection using either of these techniques remains unchartered [9].

This is therefore, a field of research we believe is appropriate to pursue.

Limitations of Transcranial Doppler and Transcranial Color Duplex

Despite our aspirations in gaining insight into the venous aspect of bAVMs post resection, there are technical limitations that may prevent successful TCCD imaging of the venous aspect of bAVMs. These include difficulty delineating the boundaries of an artery and vein due to arterialization of the venous signal [24]. Another factor perhaps influencing the omission of venous flow characteristics from previous studies may relate to the lower velocity spectrum, and the close proximity to the arteries that display a more prominent signal [27]. Regardless of such mitigating factors, it remains a worthwhile consideration to attempt to include whatever hemodynamic data can be gleaned from the venous aspect of bAVMs, particularly in the postoperative phase.

Aside from the potential limitations of obtaining meaningful venous data, there may be further likely constraints regarding the feasibility of gaining useful hemodynamic information using TCCD, particularly if there is a poor temporal window or if the bAVMs are small or distally located. Therefore, this study has been implemented initially as a pilot study to assess the viability of obtaining temporal arterial and venous hemodynamic data on a range of bAVM sizes of varied locations. It may be useful, though it also may be technically difficult or even impossible, to perform evaluation of the distal vessels to the AVM and study the changes of autoregulation through the pulsatility index/hyperemic response after AVM resections. Publication of the data obtained from this pilot study will therefore include any such technical limitations.

Conclusion

Both TCCD and CDU are noninvasive modalities that may potentially address some of the gaps in knowledge with regard to the hemodynamic remodeling following the removal of a bAVM from the cerebral vasculature, by providing simultaneous velocity and pulsatility measurements.

This protocol aims to strengthen the work of previous authors by providing documentation of the time course of hemodynamic changes following bAVM resection. The protocol is designed to determine whether noninvasive technology including CDU imaging of the extracranial carotid and intracranial arteries in the form of TCCD, along with central aortic pressure measurements can determine whether there are any hemodynamically significant prognostic markers that may provide insight into the process of vessel remodeling, including insight into venous changes following bAVM resection.

Conflicts of Interest

None declared.

References

1. Morgan M, Winder M. Haemodynamics of arteriovenous malformations of the brain and consequences of resection: a review. J Clin Neurosci 2001;8:216-224. [doi: 10.1054/jocn.2000.0795] [Medline: 11386794]
2. Hassler W, Steinmetz H. Cerebral hemodynamics in angioma patients: an intraoperative study. J Neurosurg 1987;67:822-831. [doi: 10.3171/jns.1987.67.6.0822] [Medline: 3119793]
3. Batjer H, Devous M, Meyer Y, Purdy P, Samson D. Cerebrovascular hemodynamics in arteriovenous malformation complicated by normal perfusion pressure breakthrough. Neurosurgery 1988;22:503-509. [Medline: 3258962]
4. Batjer H, Devous M. The use of acetazolamide-enhanced regional cerebral blood flow measurement to predict risk to arteriovenous malformation patients. Neurosurgery 1992;31:213-217. [Medline: 1513427]
5. Young WL, Kader A, Ornstein E, Baker KZ, Ostapkovich N, Pile-Spellman J, et al. Cerebral hyperemia after arteriovenous malformation resection is related to “breakthrough” complications but not to feeding artery pressure. The Columbia University Arteriovenous Malformation Study Project. Neurosurgery 1996;38:1085-1093; discussion 1093. [Medline: 8727137]
6. al-Rodhan NR, Sundt TM, Piepgras DG, Nichols DA, Rüfenacht D, Stevens LN. Occlusive hyperemia: a theory for the hemodynamic complications following resection of intracerebral arteriovenous malformations. J Neurosurg 1993;78:167-175. [doi: 10.3171/jns.1993.78.2.0167] [Medline: 8421198]

7. Li M. Arteriovenous malformations of the brain. In: Youmans JR, editor. Neurol Surgery,Vol 3. Philadelphia: WB Saunders; 1982:1786-1806.
8. Rangel-Castilla L, Spetzler R, Nakaji P. Normal perfusion pressure breakthrough theory: a reappraisal after 35 years. Neurosurg Rev 2015;38:339-404. [Medline: 25483235]
9. D'Aliberti G, Talamonti G, Piparo M, Debernardi A, Zella S, Boccardi E, et al. Venous flow rearrangement after treatment of cerebral arteriovenous malformations: a novel approach to evaluate the risks of treatment. World Neurosurg 2014;82:160-169. [Medline: 23396073]
10. Morgan MK, Sekhon LH, Finfer S, Grinnell V. Delayed neurological deterioration following resection of arteriovenous malformations of the brain. J Neurosurg 1999;90:695-701. [doi: 10.3171/jns.1999.90.4.0695] [Medline: 10193615]
11. Morgan MK, Winder M, Little NS, Finfer S, Ritson E. Delayed hemorrhage following resection of an arteriovenous malformation in the brain. J Neurosurg 2003;99:967-971. [doi: 10.3171/jns.2003.99.6.0967] [Medline: 14705722]
12. Hassler W. Hemodynamic aspects of cerebral angiomas. Acta Neurochir Suppl 1986;37:1-136. [Medline: 3548227]
13. Lindegaard K, Grolimund P, Aaslid R, Nornes H. Evaluation of cerebral AVM's using transcranial Doppler ultrasound. J Neurosurg 1986;65:335-344. [Medline: 3525779]
14. Guo H, Pan Y, Zhou L, Shi Y. Research on hemodynamics of cerebral arteriovenous malformation by Doppler ultrasound. Chin Med J (Engl) 1993;106:351-356. [Medline: 8404277]
15. Harders A, Bien S, Eggert H, Laborde G, Merland J, Rufenacht D. Haemodynamic changes in arteriovenous malformations induced by superselective embolization: transcranial Doppler evaluation. Neurol Res 1988;10:239-245. [Medline: 2907113]
16. Kiliç T, Pamir M, Budd S, Ozek M, Erzen C. Grading and hemodynamic follow-up study of arteriovenous malformations with transcranial Doppler ultrasonography. J Ultrasound Med 1998;17:729-739. [Medline: 9849944]
17. Chioffi F, Pasqualin A, Beltramello A, Da Pian R. Hemodynamic effects of preoperative embolization in cerebral arteriovenous malformations: evaluation with transcranial Doppler sonography. Neurosurgery 1992;31:877-884; discussion 884. [Medline: 1436411]
18. Lindegaard K, Bakke S, Grolimund P, Aaslid R, Huber P, Nornes H. Assessment of intracranial hemodynamics in carotid artery disease by transcranial Doppler ultrasound. J Neurosurg 1985;63:890-898. [Medline: 3903072]
19. Petty G, Massaro AR, Tatemichi T, Mohr J, Hilal S, Stein B, et al. Transcranial Doppler ultrasonographic changes after treatment for arteriovenous malformations. Stroke 1990;21:260-266. [Medline: 2406994]
20. Tyagi S, Mahapatra AK, Mishra NK. Transcranial Doppler evaluation of blood flow velocity changes in basal cerebral arteries in cerebral AVMs following embolisation and surgery. Neurol India 2000;48:112-115 [FREE Full text] [Medline: 10878772]
21. Park S, Hwang S. Transcranial Doppler study of cerebral arteriovenous malformations after gamma knife radiosurgery. J Clin Neurosci 2009;16:378-384. [Medline: 19138850]
22. Klotzsch C, Henkes H, Nahser HC, Kuhne D, Berlit P. Transcranial color-coded duplex sonography in cerebral arteriovenous malformations. Stroke 1995;26:2298-2301. [doi: 10.1161/01.STR.26.12.2298]
23. Schöning M, Buchholz R, Walter J. Comparative study of transcranial color duplex sonography and transcranial Doppler sonography in adults. J Neurosurg 1993;78:776-784. [doi: 10.3171/jns.1993.78.5.0776] [Medline: 8468608]
24. Bartels E. Evaluation of arteriovenous malformations (AVMs) with transcranial color-coded duplex sonography: does the location of an AVM influence its sonographic detection? J Ultrasound Med 2005;24:1511-1517. [Medline: 16239654]
25. Nornes H, Grip A. Hemodynamic aspects of cerebral arteriovenous malformations. J Neurosurg 1980;53:456-464. [doi: 10.3171/jns.1980.53.4.0456] [Medline: 7420165]
26. Lindegaard KF, Sorteberg W, Nornes H. Transcranial Doppler in neurosurgery. Adv Tech Stand Neurosurg 1993;20:39-80. [Medline: 8104401]
27. Martin PJ, Gaunt ME, Naylor AR, Terence Hope D, Orpe V, Evans DH. Intracranial aneurysms and arteriovenous malformations: transcranial colour-coded sonography as a diagnostic aid. Ultrasound in Medicine & Biology 1994 Jan;20(8):689-698. [doi: 10.1016/0301-5629(94)90026-4] [Medline: 7863558]

Abbreviations

ACVH: arterial-capillary-venous hypertensive
bAVM: arteriovenous malformation of the brain
CDU: color duplex ultrasound
TCD: transcranial Doppler
TCCD: transcranial color duplex

Edited by G Eysenbach; submitted 07.05.17; peer-reviewed by R Kirollos; comments to author 15.07.17; revised version received 26.07.17; accepted 27.07.17; published 31.08.17.

Please cite as:
Busch KJ, Kiat H
Ascertaining the Value of Noninvasive Measures Obtained Using Color Duplex Ultrasound and Central Aortic Pressure Monitoring During the Management of Cerebral Arteriovenous Malformation Resection: Protocol for a Prospective, Case Control Pilot Study
JMIR Res Protoc 2017;6(8):e173
URL: http://www.researchprotocols.org/2017/8/e173/
doi:10.2196/resprot.7991
PMID:28860105

Original Paper

An Electronic Health Record–Based Strategy to Systematically Assess Medication Use Among Primary Care Patients With Multidrug Regimens: Feasibility Study

Stacy Cooper Bailey[1], PhD, MPH; Christine U Oramasionwu[1], PhD, PharmD; Alexandra C Infanzon[1], MPH; Emily R Pfaff[2], MS; Izabela E Annis[1], MS; Daniel S Reuland[3], MD, MPH

[1]Division of Pharmaceutical Outcomes and Policy, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]North Carolina Translational and Clinical Sciences Institute, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[3]Division of General Medicine and Clinical Epidemiology, Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Stacy Cooper Bailey, PhD, MPH
Division of Pharmaceutical Outcomes and Policy
Eshelman School of Pharmacy
University of North Carolina at Chapel Hill
2204 Kerr Hall, Campus Box 7573
Chapel Hill, NC,
United States
Phone: 1 919 843 9834
Email: scbailey@unc.edu

Abstract

Background: Medication nonadherence and misuse are public health and patient safety concerns. With the increased adoption of electronic health records (EHRs), greater opportunities exist to communicate directly with, and collect data from, patients through secure portals linked to EHRs.

Objective: The study objectives were to develop and pilot test a method of monitoring patient medication use in outpatient settings and determine the feasibility and acceptability of this approach.

Methods: Adult primary care patients on multidrug regimens were recruited from an academic internal medicine clinic by a trained research assistant. After completing a baseline, in-person interview, patients were sent a link to a questionnaire about medication use via the patient portal. One week later, the RA contacted patients to complete a follow-up telephone interview assessing patient satisfaction and experience with the questionnaire. Patient EHRs were also reviewed to determine the questionnaire completion rate.

Results: Of 100 patients enrolled, 89 completed the follow-up interview and 82 completed the portal questionnaire. The mean age of the sample was 61.8 (range 31-88) years. Approximately half (54/100, 54%) of the sample was male, two-thirds were white (67/100, 67%) and 26% (26/100) African-American. A total of 44% reported an annual household income of <$50,000 per year, and 17% (17/100) reported a high school or less level of education. No significant differences were found in questionnaire completion rates by sociodemographic characteristics or prior portal use. Most (68/73, 93%) found the questionnaire easy to access, easy to complete (72/73, 99%), and valuable (73/89, 82%). Time constraints and log-in difficulties were the main reasons for noncompletion.

Conclusions: The portal questionnaire was well received by a socioeconomically diverse group of patients with high completion rates achieved. Routine use of a portal-based questionnaire could provide a valuable signal to providers and care teams about patient medication use and identify patients needing additional support.

(JMIR Res Protoc 2017;6(8):e157) doi:10.2196/resprot.7986

KEYWORDS

patient portal; medication adherence; health literacy

Introduction

While medication nonadherence has long been recognized as a public health and patient safety concern, it is likely to become increasingly important with the growing prevalence of chronic disease and the aging of the US population [1-3]. Estimates indicate the number of people taking prescription drugs is increasing over time, with the percentage of the US population who take 5 or more medications nearly doubling from 8% in 1999-2000 to 15% in 2011-2012 [4,5]. Such complex, multidrug regimens usually require greater patient self-management skills and clinical oversight. However, providers often lack the time and resources needed to identify and assist patients who exhibit poor medication adherence or misunderstand instructions for use.

With the increased implementation and use of electronic health records (EHRs), greater opportunities exist to engage patients in care and collect patient data on health behaviors outside the clinic setting [6]. Yet while the potential value of using EHRs for patient engagement and monitoring has been well recognized, the actual acceptability and practicality of EHR-based strategies have not been as well explored [7,8]. In this feasibility study, we sought to leverage a commonly used EHR platform to develop and pilot test a strategy to monitor patient medication use in outpatient settings and determine the feasibility and acceptability of this approach. Findings from this study can inform future implementation and use of similar EHR-based tools.

Methods

Envisioning the Electronic Health Record–Based Strategy

The intent of this study was to develop and pilot test a low-cost, sustainable method of systematically collecting data from patients on their prescription medication use at home. The belief was that such data could help identify patients who were struggling with medication adherence or had other medication concerns that could reasonably impact their health care and health outcomes. Given limited clinic resources, a portal-based questionnaire, which could be automatically delivered to any patient or group of patients on a routine basis and for virtually no cost, was selected as the most appropriate means of collecting patient data on medication use outside the clinic visit. The strategy was envisioned to include 3 steps: (1) patients receive and complete an online questionnaire via the patient portal, (2) questionnaire results are recorded in the EHR for review by the clinical care team, and (3) based upon results, the clinical care team provides additional counseling or resources to patients on appropriate medication use according to a clinic protocol. The study reported herein is focused on the development and initial pilot-testing of the portal questionnaire itself (steps 1 and 2); additional studies are planned to evaluate the strategy as a whole.

Building the Electronic Health Record Strategy

To create the portal questionnaire, a survey consisting of 12 items from the Measure of Drug Self-Management (MeDS) was built using Epic EHR (Epic Systems Corp) [9]. The MeDS, which deconstructs the tasks associated with taking medications focusing on the patient knowledge, skills, and behaviors needed to correctly take multiple drugs, was chosen because it was originally designed for use in an EHR environment. Items and response options were programmed individually and then grouped together into a survey. The questionnaire was designed to be delivered to a patient via My Chart, the health care system's patient portal powered by the Epic EHR system, along with an accompanying new message notification sent to the patient's on-file email address. A link was included in the message to direct patients to the questionnaire.

After completion, individual item responses were tallied and an overall score was calculated and filed in the patient's chart for viewing by the clinical care team. The intention was for questionnaire results to help inform patient-provider discussions on medication use during future clinical appointments and/or to trigger immediate outreach to the patient should a serious medication concern be identified. The exact response would be determined by the nature of the questionnaire results and clinical care protocols. The provider's view within the EHR was also custom built to allow questionnaire results to be graphed longitudinally over time along with other relevant patient data. This presentation was chosen as it would enable providers to monitor trends in medication use and determine the effects, if any, of clinical interventions to improve adherence.

Pilot-Testing the Electronic Health Record–Based Strategy

Overview

Following its development, this EHR-based strategy was pilot-tested among patients receiving care at an academic general internal medicine clinic. Convenience sampling was used to recruit patients. Specifically, a research assistant reviewed the list of daily appointments at the study clinic and approached patients who were potentially eligible (by age, medications prescribed, and patient portal access) to introduce the study and determine patient interest and eligibility. Patients were considered eligible if they were age 18 years or older; were English-speaking; had primary responsibility for managing their own medications; were prescribed at least 3 medications including 2 specific medications for diabetes and high cholesterol (ie, metformin, atorvastatin); had access to high-speed Internet at home or had a smartphone; were a clinic patient with an activated patient portal account; and had no severe cognitive, visual, or hearing impairment that would preclude informed consent or study participation.

After verifying eligibility, the RA engaged the patient in the informed consent process and conducted a structured, in-person interview. The RA then manually assigned the questionnaire to the participant within the portal. A follow-up telephone interview was conducted 1 week later by the same RA. Participants were compensated $20. An institutional review board approved study procedures.

Outcomes

Process outcomes for this study included whether the questionnaire was delivered to participants via the portal, whether participants completed the questionnaire, and whether

the questionnaire results were recorded and displayed correctly within the EHR. The RA recorded outcomes by abstracting data from the EHR. Additionally, participants were asked to self-report receipt and completion of the questionnaire during the follow-up interview. The interview also assessed patient experiences with and perceptions of the electronic tool and strategy.

Sociodemographic Variables and Covariates

The baseline battery also included questions measuring key patient health and sociodemographic characteristics as well as prior portal and Internet use. Health literacy was measured using the Rapid Estimate of Adult Literacy in Medicine (REALM), a word pronunciation test that is commonly used to assess patient literacy skills [10].

Statistical Analyses

Descriptive statistics were calculated for patient sociodemographic variables and study outcomes. To assess whether there were any systematic, statistically significant differences between patients who completed the follow-up interview and those who did not, we used Pearson chi-square test or Fisher exact test for categorical variables and Student *t* test for age. The same tests were used to compare the differences between the patients who completed the portal questionnaire and those who did not. Specifically, we examined if completion varied by age, sex, race/ethnicity, education, income, literacy skills, and average use of the patient portal. We used α=.05 to determine statistical significance. All analyses were performed using SAS version 9.3 (SAS Institute Inc).

Results

Participant Characteristics

Recruitment took place from March 2016 to August 2016. A total of 171 patients were approached; 39 patients declined, 31 patients were ineligible, and 1 patient consented but did not initiate the interview. A total of 100 patients completed the baseline interview.

Table 1 describes the characteristics of the study sample (N=100). The mean age of this study sample was 61.8 (range 31-88) years. Approximately half (54/100, 54%) of the sample was male, two-thirds were white (67/100, 67%), and 17% (17/100) reported a high school or less level of education. Most (92/100, 92%) participants had adequate literacy skills according to the REALM. Few patients (5/100, 5%) reported never having used the patient portal; 39% (39/100) stated that they used the portal once per month or less. Most patients (62/100, 62%) reported having used a computer to access the site, 19% (19/100) reported using a smartphone, 13% (13/100) a tablet, and 21% (21/100) said they used all 3 types of devices (categories not mutually exclusive). We observed no sociodemographic differences between patients who participated in the follow-up interview (n=89) versus those who did not (n=11).

Process Outcomes

A review of the EHR revealed that the portal questionnaire was successfully generated and delivered to all participants (N=100). Of those, 82 of 100 participants completed and submitted the questionnaire via the portal; all questionnaire scores and item responses were recorded and displayed correctly in patient charts.

A total of 73 participants who completed the questionnaire also completed the follow-up interview. Of these, the majority (65/73, 89%) reported completing the portal questionnaire on a computer, laptop or tablet while the remaining 11% (8/73) patients reported completing the measure via a smartphone. There were no sociodemographic differences between participants who completed the portal questionnaire versus those who did not (Table 2). Similarly, patients who reported never having used the patient portal or using it less than once per month were as likely to complete the portal questionnaire as more frequent patient portal users (*P*=.61).

Satisfaction with the Questionnaire

Most patients (68/73, 93%) who completed both the questionnaire and the follow-up interview reported that the tool was very easy to access on the portal, and 99% (72/73) of patients reported the tool was very easy to complete. Most patients reported that the questionnaire was an acceptable length, with only 2 patients stating that it was too long. Overall, satisfaction with the questionnaire was high, with the tool receiving an average satisfaction score of 9.3 (range of 5 to 10), with a score of 10 indicating highly satisfied.

Reasons for Noncompletion

A total of 16 people completed the follow-up interview but did not complete the portal questionnaire. The majority of these participants (10/16) cited time constraints as the reason for not completing the tool. A total of 3 participants reported that they had difficulty logging into the portal, and 2 stated that they did not receive a notification email. The remaining patient did not provide a specific reason for noncompletion.

Participants' Perceived Usefulness of the Electronic Health Record Strategy

Of the patients participating in the follow-up interview, 82% (73/89) believed that a medication adherence questionnaire would be valuable to complete prior to clinic visits to help measure and track patient medication self-management skills over time. While 16% (14/89) of patients stated that they did not know if the questionnaire would be valuable, 2% (2/89) of respondents stated that they did not believe such a tool would be useful.

Table 1. Characteristics of study sample.

Variable	Participants N=100
Age, years, mean (range)	61.8 (31-88)
Sex, n	
Male	54
Female	46
Race/ethnicity, n	
African American	26
White	67
Other	7
Educational attainment, n	
High school diploma, general education diploma, or less	17
Some college	26
College degree or more	57
Income, n	
<$19,999	16
$20,000 to $49,999	28
≥$50,000	45
Don't know/refused	11
Literacy skills, n	
Inadequate	8
Adequate	92
Methods of prior portal access, n[a]	
Never accessed	5
Computer or laptop	62
Tablet	13
Smartphone	19
All above devices	21

[a]Responses are not mutually exclusive.

Table 2. Patient characteristics, stratified by completion of portal questionnaire.

Characteristic	Completed N=82	Not completed N=18	*P* value[a]
Age, years, mean (SD)	61.6 (11.8)	62.6 (12.0)	.74
Sex, n (%)			.50
Male	43 (52)	11 (61)	
Female	39 (48)	7 (39)	
Race/ethnicity, n (%)			.09
White	58 (71)	9 (50)	
Other	24 (29)	9 (50)	
Educational attainment, n (%)			.45
High school diploma, general education diploma, or less	12 (15)	5 (28)	
Some college	22 (27)	4 (22)	
College degree or more	48 (59)	9 (50)	
Income, n (%)			.19
<$50,000	34 (42)	10 (56)	
≥$50,000	40 (49)	5 (28)	
Don't know/refused	8 (10)	3 (17)	
Average patient portal use, n (%)			.61
Once per month or less	31 (38)	8 (44)	
2-3 times per month	33 (40)	5 (28)	
At least once weekly	18 (22)	5 (28)	

[a]Differences between groups were tested using Student *t* test for continuous variables and Pearson chi-square test or Fisher exact test for categorical variables.

Discussion

Principal Findings

Results from this feasibility study indicate that a patient portal-based questionnaire could be used among primary care patients to routinely and systematically assess medication self-management skills and identify medication-related concerns. Most patients found a portal-based questionnaire to be easy to use, and high completion rates were achieved among a sociodemographically diverse set of patients. Of note, only 18 of the 100 patients who were sent the questionnaire did not complete it. Analyses suggest that there were no systematic differences between those who completed the portal questionnaire and those who did not; the most commonly cited reason for not completing the online assessment was lack of time, not difficulty with the tool itself. The majority of patients reported believing that the EHR-based strategy could be beneficial in terms of keeping their providers informed of their outpatient medication use.

While these findings are promising, it is necessary to note that portal-based tools are only likely to be beneficial to those who can use them, namely, those who have access to the Internet and have the cognitive and computer skills necessary to complete an online assessment. Individuals lacking Internet access have historically been more likely to be older, low-income, members of racial/ethnic minority groups, and to have limited literacy skills, raising concerns that implementing EHR-based interventions may further exacerbate disparities [11,12]. However, recent national data suggests that Internet access is on the rise among many of these groups, particularly when access via smartphones is taken into account [13,14]. An estimated 77% of US adults now report that they own a smartphone, with similar rates of ownership among white, African American, and Hispanic adults [15]. Use of mobile devices may therefore help reduce racial disparities in portal access and increase overall use. This is reflected in this study, where 11% of participants completed the portal-based assessment via their smartphone and 40% reported having used a smartphone (either as the sole method or one of multiple methods) to access the portal in the past.

Beyond increasing accessibility via mobile devices, other advances could also promote patient use of EHR-based tools. Recently deployed technology at this study clinic will now allow for electronic questionnaires to be completed securely by patients outside the portal environment, with the results still populating in patient charts. This is likely to remove some of the noted barriers to completing portal-based assessments, such as difficulties logging into the portal or remembering portal passwords.

Limitations

There are limitations to the study that should be noted. First, it was a small feasibility study conducted among 100 primary

care patients at one university-based clinic. All patients enrolled in the study had Internet access and a portal account, although the account may not have been in active use. Results may not be generalizable to other populations. We were also limited in terms of the analyses that could be conducted given the small sample size and lack of variability in certain outcomes. In particular, we were unable to determine the acceptability of this approach specifically among patients with limited health literacy. Additional studies will need to be conducted to determine if this strategy is suitable for this patient population.

As this was a preliminary feasibility study, we did not assess provider access and use of questionnaire results nor did we examine its effectiveness or long-term impact in actual use. Informal feedback from health care providers at the study clinic indicated that many believed the questionnaire could be beneficial, although they acknowledged that physicians often experience information overload and may have difficulty reviewing scores and responding to identified concerns during time-limited clinic visits. It is also unknown whether the utility of the measure would diminish over time after multiple administrations or if it would be improved if tailored to patients based on their own unique, medication-related challenges. Additional research will be needed to examine the effectiveness of the EHR-based approach and its utility to patients, providers, and care teams.

Conclusions

Routine use of a portal-based questionnaire could provide a valuable signal to providers and their care teams about patient use of medications and help identify those patients in need of additional counseling or resources. Such an approach could help improve quality of care and promote medication adherence and safety as questionnaire results could alert providers to patient confusion that could lead to nonadherence, potential medication errors or even a preventable adverse drug event. This strategy could also help elucidate if a patient's inability to achieve therapeutic goals is related to biologic failure of the drug regimen or poor self-care. While more research is needed, preliminary results from this study indicate portal questionnaires can be successfully generated, delivered to, and completed by patients with results stored in charts for subsequent review and monitoring by members of the primary health care team. Additional research will be necessary to fully evaluate the strategy in actual use.

Acknowledgments

This study was supported in part by a research grant from the Investigator-Initiated Studies Program of Merck Sharp & Dohme Corp. The opinions expressed in this paper are those of the authors and do not necessarily represent those of Merck Sharp & Dohme Corp. Merck Sharp & Dohme Corp was not involved in the design or conduct of the study or in the reporting of study results. The project described was also supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through grant award number UL1TR001111. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Dr Bailey has served as a consultant to Merck Sharp & Dohme Corp for work unrelated to this manuscript. Drs Bailey, Oramasionwu, and Reuland have received grant funding via their institution from Merck Sharp & Dohme Corp and Eli Lilly and Company.

Conflicts of Interest

None declared.

References

1. Aspden P, Wolcott JA, Bootman JL, Croenwett LR. Preventing Medication Errors. Washington: National Academies Press; 2007.
2. World Health Organization. Adherence to Long-Term Therapies: Evidence for Action. Geneva: World Health Organization; 2003.
3. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007 Dec;22 Suppl 3:391-395 [FREE Full text] [doi: 10.1007/s11606-007-0322-1] [Medline: 18026807]
4. Gu Q, Dillon C, Burt V. Prescription drug use continues to increase: U.S. prescription drug data for 2007-2008. NCHS Data Brief 2010 Sep(42):1-8 [FREE Full text] [Medline: 20854747]
5. Kantor ED, Rehm CD, Haas JS, Chan AT, Giovannucci EL. Trends in prescription drug use among adults in the United States from 1999-2012. JAMA 2015 Nov 03;314(17):1818-1831 [FREE Full text] [doi: 10.1001/jama.2015.13766] [Medline: 26529160]
6. Office of the National Coordinator for Health Information Technology. Health IT Dashboard: quick stats. 2017. URL: https://dashboard.healthit.gov/quickstats/quickstats.php [accessed 2017-07-21] [WebCite Cache ID 6s7wx1L0Q]
7. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
8. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res 2015;17(2):e40 [FREE Full text] [doi: 10.2196/jmir.3703] [Medline: 25707035]

9. Bailey SC, Annis IE, Reuland DS, Locklear AD, Sleath BL, Wolf MS. Development and evaluation of the Measure of Drug Self-Management. Patient Prefer Adherence 2015;9:1101-1108. [doi: 10.2147/PPA.S85411] [Medline: 26257515]
10. Davis T, Crouch M, Long S, Jackson R, Bates P, George R, et al. Rapid assessment of literacy levels of adult primary care patients. Fam Med 1991 Aug;23(6):433-435. [Medline: 1936717]
11. Bailey SC, O'Conor R, Bojarski EA, Mullen R, Patzer RE, Vicencio D, et al. Literacy disparities in patient access and health-related use of Internet and mobile technologies. Health Expect 2014 Nov 2. [doi: 10.1111/hex.12294] [Medline: 25363660]
12. Gordon NP, Hornbrook MC. Differences in access to and preferences for using patient portals and other eHealth technologies based on race, ethnicity, and age: a database and survey study of seniors in a large health plan. J Med Internet Res 2016 Mar 04;18(3):e50 [FREE Full text] [doi: 10.2196/jmir.5105] [Medline: 26944212]
13. Perrin A, Duggan M. Americans' Internet access 2000-2015. Washington: Pew Internet and American Life Project URL: http://www.pewinternet.org/files/2015/06/2015-06-26_internet-usage-across-demographics-discover_FINAL.pdf [accessed 2017-07-21] [WebCite Cache ID 6s7xMzcYy]
14. Smith A. US smartphone use in 2015. Washington: Pew Internet and American Life Project; 2015. URL: http://www.pewinternet.org/files/2015/03/PI_Smartphones_0401151.pdf [accessed 2017-07-21] [WebCite Cache ID 6s7xTvRdU]
15. Mobile fact sheet. Washington: Pew Internet and American Life Project URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2017-07-21] [WebCite Cache ID 6s7xc5EiX]

Abbreviations

EHR: electronic health record
MeDS: Measure of Drug Self-Management
REALM: Rapid Estimate of Adult Literacy in Medicine

Edited by G Eysenbach; submitted 09.05.17; peer-reviewed by C Kruse, JJ Mira; comments to author 01.06.17; revised version received 08.06.17; accepted 09.06.17; published 10.08.17.

Please cite as:
Bailey SC, Oramasionwu CU, Infanzon AC, Pfaff ER, Annis IE, Reuland DS
An Electronic Health Record–Based Strategy to Systematically Assess Medication Use Among Primary Care Patients With Multidrug Regimens: Feasibility Study
JMIR Res Protoc 2017;6(8):e157
URL: http://www.researchprotocols.org/2017/8/e157/
doi:10.2196/resprot.7986
PMID:28798013

Protocol

Feasibility of a Facebook Intervention for Exercise Motivation and Cardiac Rehabilitation Adherence: Study Protocol

Lee Anne Siegmund[1], RN, PhD; Haitham M Ahmed[2], MD, MPH; Michael Todd Crawford[2], MS; James Frank Bena[3], MS

[1]Nursing Institute, Office of Nursing Research and Innovation, The Cleveland Clinic, Cleveland, OH, United States

[2]Heart and Vascular Institute, Department of Cardiovascular Medicine, The Cleveland Clinic, Cleveland, OH, United States

[3]Lerner Research Institute, Quantitative Health Sciences, The Cleveland Clinic, Cleveland, OH, United States

Corresponding Author:
Lee Anne Siegmund, RN, PhD
Nursing Institute
Office of Nursing Research and Innovation
The Cleveland Clinic
T4
9500 Euclid Ave
Cleveland, OH, 44195
United States
Phone: 1 216 445 3457
Fax: 1 216 445 1776
Email: siegmul@ccf.org

Abstract

Background: While cardiac rehabilitation has been shown to be effective at improving coronary heart disease (CHD), participation is generally poor. Attempts to increase uptake and adherence often fail. Use of a Facebook intervention for this population may be a unique opportunity to support self-determined motivation and affect adherence.

Objective: To evaluate the impact of a Facebook intervention on motivation for exercise and adherence to cardiac rehabilitation in patients with CHD during a 12-week, Phase II cardiac rehabilitation program.

Methods: A prospective, randomized controlled pilot study, grounded in Self-Determination Theory, will be conducted. Participants will be recruited from inpatient, or the intake visit to outpatient, cardiac rehabilitation, and then randomly assigned to the intervention or comparison group. Participants in the intervention group will take part in a private Facebook group. Weekly posts will be designed to support self-determined motivation, measured at baseline and postcardiac rehabilitation by the Behavioral Regulation in Exercise Questionnaire-3 (BREQ-3). The Psychological Need Satisfaction for Exercise (PNSE) scale will measure fulfillment of needs that affect motivation. Participants in the comparison group will be given the same materials, but these will be supplied via handouts and email. The number of sessions attended will be tallied and analyzed using *t* tests. Overall motivation will be evaluated using analysis of covariance (ANCOVA) models. Multivariate analysis of variance models will be used to evaluate differences in the change across motivation subtypes. If significant, ANCOVA models for each subtype will be fit. ANCOVA models will be used to compare changes in needs satisfaction, overall and separately among the three subscales, between groups. Engagement in the Facebook group will be measured by number of “likes” and self-report of weekly visits to the group.

Results: This project was funded in July 2017 and recruitment is currently underway. The recruitment goal is 60 cardiac rehabilitation patients. Data collection is anticipated to be complete by July 2018.

Conclusions: This pilot study will be the first to examine the effect of a Facebook intervention on patient adherence and motivation for exercise in a cardiac rehabilitation setting. Engagement in the Facebook group and participation in the study will help to determine the feasibility of using Facebook to affect adherence and motivation in cardiac rehabilitation patients, potentially improving outcomes through the use of a unique intervention.

Trial Registration: ClinicalTrials.gov NCT02971813; https://clinicaltrials.gov/ct2/show/NCT02971813 (Archived by WebCite at http://www.webcitation.org/6sRsz8Zpa)

(JMIR Res Protoc 2017;6(8):e162) doi:10.2196/resprot.7554

KEYWORDS

cardiac rehabilitation; social media; adherence; motivation; Facebook; Self-Determination Theory; Behavioral Regulation in Exercise Questionnaire-3; BREQ-3; Psychological Need Satisfaction in Exercise scale

Introduction

Background

Coronary heart disease (CHD) is the leading killer of men and women and currently accounts for 15.5 million cases in the United States [1]. Phase II cardiac rehabilitation, a Class 1 recommendation by the American College of Cardiology Foundation and the American Heart Association, is a secondary prevention program that has been shown to be safe and effective in treating patients diagnosed with existing CHD [2-10]. However, despite the reported effectiveness of cardiac rehabilitation, many at high risk for CHD are less likely to adhere to the program [11]. Utilization of cardiac rehabilitation is low overall, particularly for women, minorities, and those with comorbidities [4], and attempts to increase uptake and adherence often fail [12].

In recent years, Web-based interventions have been used to examine exercise adherence and theory-supported apps have enabled feedback on exercise intensity and adherence in remotely delivered cardiac rehabilitation [13]. Interventions utilizing the Web improved daily step counts [14] and physical activity intensity [15]. The use of such apps has been shown to be feasible and acceptable for use in special populations, including patients with cystic fibrosis [16] and cancer survivors [17]. A recent randomized controlled trial utilizing online social media to test its effect on physical activity found that the social support provided by the program resulted in an increase in group cohesion [18]. The perception of group cohesion may be important to patients in cardiac rehabilitation since social support was found to be an important component in exercise adherence [19]. Due to the vital role that social support has played in helping people to become more self-motivated [20], it is appropriate to examine unique ways to foster a sense of belonging or connectedness.

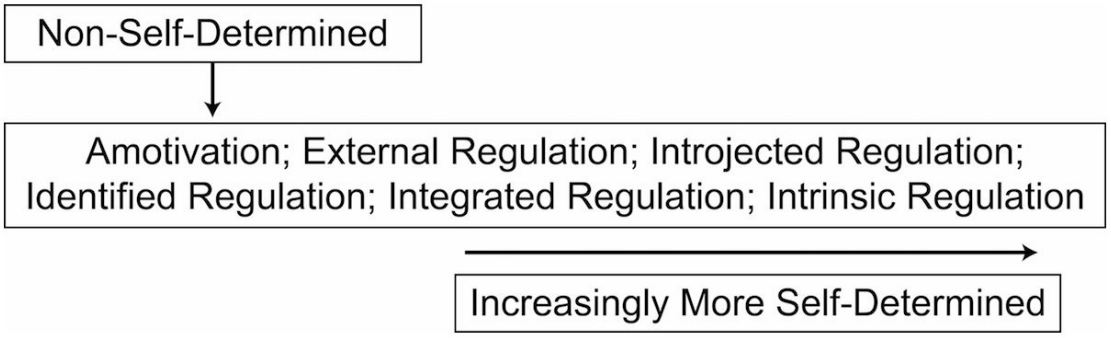
Social media is growing in popularity, making it an interesting venue for delivery of an intervention designed to affect cardiac rehabilitation adherence. Facebook in particular has the most engaged users of all social media sites, with more than 75% logging in daily [21]. Use of social network platforms on the Web, such as Twitter or Facebook, has helped patients manage personal health and has increased adherence to medical treatment [22], possibly through a sense of involvement and social support. Joseph et al [15] showed pilot data that supports Facebook as a tool for promoting physical activity by utilizing education and group discussions. Facebook, relative to other social media or Web-based interventions, has been reported to have high retention rates when used to affect health behaviors [23]. While Facebook has been studied as a means to improve physical activity in a number of populations [17,24-27], there is a knowledge gap regarding the effectiveness of social networking interventions used to promote health [28] and its use in cardiac rehabilitation as a tool to improve motivation.

Theoretical Framework

This study is grounded in Self-Determination Theory [20,29,30], which defines motivation in terms of intrinsic and extrinsic sources (see Figure 1). Self-Determination Theory focuses on social and cognitive factors and how those factors influence an individual's motivation. The theory describes motivation as being on a continuum with behavioral regulators ranging from amotivation, in which a person lacks intention to do an activity, to intrinsic regulation, in which an individual may do the activity simply for the joy of it [20]. Self-Determination Theory specifically examines conditions that lead to self-determined (ie, internalized) motivation and states that three psychological needs are necessary for it to exist: competence, autonomy, and relatedness [20]. A motivationally supportive environment supports these three needs in several ways. Competence, in essence self-efficacy, can be supported through provision of structure, offering participants positive feedback and helping them to set realistic goals [31-33]. Competence, according to Cognitive Evaluation Theory, a subtheory of Self-Determination Theory, will not lead to intrinsic motivation in the absence of autonomy [34]. Autonomy may be supported by helping the individual make decisions for personal reasons and helping them to make choices with minimal pressure [31-33]. Relatedness can be promoted by providing a sense of connectedness to others. An environment that helps a person feel socially included and supported by others may help facilitate intrinsic motivation [20,32,35].

Motivation for exercise is an important concept in the examination of cardiac rehabilitation adherence. Self-Determination Theory was previously used as a theoretical framework for motivational research in a cardiac rehabilitation setting [36]. Thorup and colleagues [36] showed qualitative evidence that a pedometer-based cardiac rehabilitation intervention supported autonomy, competence, and relatedness. It is possible that increasingly more self-determined (ie, internalized) motivation may be enough to help patients overcome the many obstacles associated with nonadherence to exercise and cardiac rehabilitation.

Figure 1. Self-Determination Theory.



Study Objectives

The purpose of the current randomized pilot trial is to determine the feasibility of using a Facebook intervention that provides education, peer support, and provider support to affect change in motivation and self-determination for exercise and adherence to cardiac rehabilitation in patients with CHD during a 12-week, Phase II cardiac rehabilitation program. It is hypothesized that the following will occur:

1. Scores for motivation for exercise overall will increase for patients exposed to a Facebook intervention and across individual motivational subtypes (ie, regulations) relative to a comparison group who receive educational handouts and emails.

2. The percentage of cardiac rehabilitation sessions attended will be higher relative to a comparison group who receive educational handouts and emails.

3. Engagement in the private Facebook group (ie, number of visits to the group and "likes") will predict the number of cardiac rehabilitation sessions attended and the change in motivation. The feasibility of a larger trial will be based on sample size and participants' engagement in the Facebook group.

Methods

Design

This is a prospective, randomized controlled pilot trial to evaluate the feasibility of using a social media intervention to affect change in motivation for exercise and adherence to cardiac rehabilitation sessions.

Setting and Sample

The setting for this study will be in the outpatient cardiac rehabilitation at the main campus of a large tertiary care center in Northeast Ohio, USA, several satellite facilities in the region, and in patients' homes or other locations where home computers might be accessed. This cardiac rehabilitation program provides electrocardiogram-monitored and supervised exercise, dietary guidance, and smoking cessation, behavioral counseling and stress reduction. All patients receive an individualized exercise prescription based on functional capacity at intake. Most patients, depending on insurance coverage, will be able to attend up to three sessions per week for a total of 36 sessions. In addition, patients are given guidance for unsupervised exercise at home.

All patients who are current and regular Facebook users, have qualified for cardiac rehabilitation (ie, diagnosed with CHD), and are entering cardiac rehabilitation at the main campus of this tertiary care center will qualify to participate in the study prior to beginning Phase II cardiac rehabilitation. Current Facebook users were chosen, as it is important that participants are skilled at using the Internet and familiar with social media. Regular use will be defined as logging on to Facebook at least two times in the last month. Inclusion criteria will include both men and women 18 years of age or older who speak English and live within 100 miles of the main campus of this tertiary care center. Participants must be able to read and understand English in order to complete the consent form, the Psychological Need Satisfaction in Exercise (PNSE) scale [37], and the Behavioral Regulation in Exercise Questionnaire-3 (BREQ-3). There will be no exclusion based on secondary diagnosis; however, participants must be able to exercise well enough to qualify to take part in cardiac rehabilitation.

Measures

The primary hypothesis, change in motivation for exercise, will be measured at baseline and postintervention using the BREQ-3. The BREQ-3 is a 24-question validated instrument that measures forms of intrinsic and extrinsic regulation of exercise behavior [34] and is based on Self-Determination Theory. Psychometrics were first completed for the Behavioral Regulation in Exercise Questionnaire-2 (BREQ-2) by Markland and Tobin [38]. Cronbach alpha reliabilities were as follows: amotivation, .83; external regulation, .79; introjected regulation, .80; identified regulation, .73; and intrinsic regulation, .86. The BREQ-3 includes five questions in addition to those on the BREQ-2 and has a new subscale for integrated regulation [33]. The subscales (ie, regulations) of the BREQ-3 are used to calculate a relative autonomy index (RAI) [39]. Each question is answered on a 5-point Likert scale (range 0-4) and represents one of the regulations. The regulations are weighted then summed to give a single score. The resulting score or index gives an indication of the individual respondent's self-determination for exercise. The RAI will place individual motivational subtypes or behavioral regulations on the self-determination continuum from amotivated (ie, lacking intention to exercise, score of -3) to intrinsically motivated (ie, self-determined or autonomously motivated, score of +3).

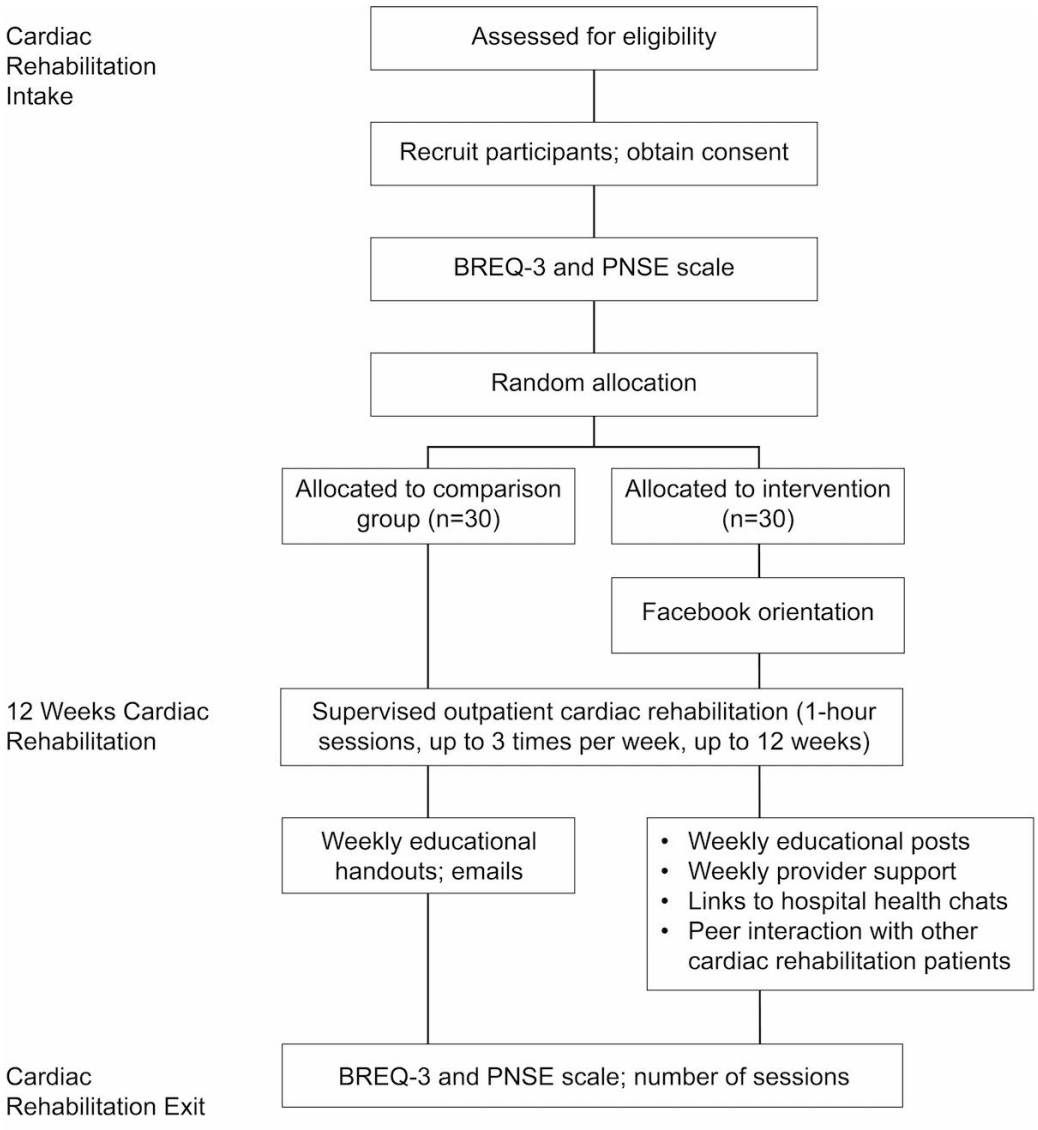
The PNSE scale will be used to assess need satisfaction with exercise. This scale was designed to assess the perception of

psychological need satisfaction associated with self-determined motivation for exercise and consists of 18 items on a 6-point Likert scale, with three subscales measuring perceived competence, autonomy, and relatedness. The scale has shown high internal consistency (Cronbach alpha>.90) [37].

The secondary hypothesis, the percentage of cardiac rehabilitation sessions attended, will be measured at the time of cardiac rehabilitation completion or dropout. It will be calculated by dividing the number of sessions attended in a 3-month period of time by the total number of sessions allowed by insurance and multiplying by 100.

The tertiary hypothesis, Facebook engagement, will be assessed by measuring the number of "likes" by individuals on the private Facebook group. "Likes" (ie, the number of times a participant clicks "like" on any of the Facebook posts) will be counted and, along with visits to the group, will be used to examine the association between engagement in the social media intervention (ie, Facebook) and cardiac rehabilitation adherence and change in motivation. A postintervention questionnaire will be given to determine number of visits to the group. The participants will be asked to circle the number of times they accessed the private Facebook group per week: 0, 1-5, 6-10, 11-15, or >15 times. The questionnaire will also be used to collect qualitative data on participants' perceptions of the intervention, including whether they felt supported in their care and more in touch with providers, whether or not they chatted with other Facebook members, and if the Facebook group affected their exercise behaviors. The questionnaire will use a 5-point Likert scale from 1 (*not at all*) to 5 (*quite a bit*) for all questions in addition to a section for comments. Participants may also grant permission for the evaluation of comments made on the private Facebook group, allowing the researchers to explore themes for qualitative analysis. Examination of comments will allow for a better understanding of the effectiveness of individual posts and the satisfaction of needs that may lead to self-determined motivation.

Figure 2. Study flowchart. BREQ-3: Behavioral Regulation in Exercise Questionnaire-3; PNSE: Psychological Need Satisfaction in Exercise.



Patient characteristics will be collected and will include key demographic variables (ie, age, gender, race, employment, distance to cardiac rehabilitation, and socioeconomic status), engagement (ie, number of visits to the group and "likes"), and key clinical variables (ie, cardiac rehabilitation indication, hypertension, diabetes, hyperlipidemia, and waist circumference), which will be obtained from the electronic medical record.

Data Collection Procedures

Overview

Volunteers will be recruited from the main campus of this tertiary care center during their inpatient stay or the intake visit for cardiac rehabilitation at the main campus and satellite facilities in the region. Volunteers will be screened for Facebook use and interest in the study, the protocol will be explained, and volunteers will then be asked to sign consent forms. The consent form will address the fact that Facebook is a public forum and names and comments are seen by other participants and the research team. The Facebook group will be private in the sense that those not in the group will not be able to see the content. Participants will complete a baseline BREQ-3 questionnaire and PNSE scale at the time of consent. Participants will then be randomized to Facebook versus comparison groups using blocked randomization (see Figure 2).

Intervention

The Facebook intervention will include peer support, education, provider support, and text message prompts when new posts are added. These interventions are designed to minimize pressure, offer choices, and allow for peer interaction, positive feedback, guidance, and direction in order to provide support for competence, autonomy, and relatedness. Competence will primarily be supported with the use of educational posts in the Facebook group. Autonomy support will come from the provider posts. Finally, relatedness will be supported by peer interaction and engagement in the Facebook group.

Educational posts will cover 12 topics that will encourage participants to practice preventive heart care while offering a variety of suggestions and encouragement for making personal health care choices. The educational portion of the intervention is designed to offer clear information and structure, thus supporting competence, which may help to enhance intrinsic motivation. These 12 educational topics will be standardized such that they will be posted on the Facebook group, one each week, and then the same ones will be reposted again every 12 weeks. The posts may be in the form of text, video, and/or pictures; they will include materials from the hospital's health library and other fact sheets and videos produced by the hospital, the American Heart Association, and the US Centers for Disease Control.

Provider posts will include topics such as motivational quotes, encouragement, reminders to exercise independently, and reminders to contact providers with questions. These postings are designed to promote a sense of choice and help participants feel that providers see them as having a unique frame of reference, thus being supportive of autonomy. Providers will be nurses on the research team, exercise physiologists, nurse practitioners, and physicians who may or may not choose to reveal their personal identities. All Facebook participants will see the same content. Provider support will also include links to provider health chats, in which patients can chat online with providers at set dates and times.

Peer interaction on Facebook will be as frequent as the participant freely chooses and will be monitored daily by the research team for appropriateness of content. Engagement in Facebook is designed to offer an opportunity for social inclusion and a sense of involvement, allowing for relatedness.

The comparison group will receive the same educational and provider support materials as the Facebook group, but will receive it in the form of a handout or via email in the event the patient cannot be contacted or misses cardiac rehabilitation on a particular week. Both groups will have the opportunity for weekly education classes and typical peer interactions, which will involve up to 3 hours of group cardiac rehabilitation per week.

Upon cardiac rehabilitation completion or dropout, post data will be collected. It is anticipated that this pilot will take up to 1 year and will be completed when 30 participants for each group have been obtained (see Table 1).

Table 1. Study calendar.

	Study timeline				
	Month 1	Months 2-7	Months 4-11	Months 4-12	Month 12
Study event	Begin recruitment; Complete intake for first 8 subjects (BREQ-3[a], PNSE[b] scale); No data used for first 8 subjects	Collection of in-take data	Collection of exit data	Data cleaning	Statistical analysis; Begin manuscript writing and preparation for longer trial

[a]BREQ-3: Behavioral Regulation in Exercise Questionnaire-3.

[b]PNSE: Psychological Need Satisfaction in Exercise.

Data Analysis

Statistical Methods

This is a feasibility study and the sample size obtained will determine if the study is appropriately powered to detect the desired effect size. Patient characteristics will be summarized by group using frequencies and percentages for categorical factors, and using means and standard deviations for continuous measures. In order to examine the primary outcome—differences in change in motivation between groups—overall motivation using the RAI from the BREQ-3 will be evaluated using analysis of covariance (ANCOVA) models. Mean differences with 95% CIs for group differences will be presented. Multivariate analysis of variance models will be used to evaluate differences in the change across individual motivation subtypes (ie, regulations), using the BREQ-3, between groups overall. If significant, separate ANCOVA models for each subtype will be fit. Similar ANCOVA models will be used to compare changes in needs satisfaction scores, overall and separately among the three subscales, between groups. Two-sample *t* tests will be used to compare number of sessions completed. As a secondary analysis, the relationships between patient characteristics, numbers of visits to the group and "likes," and the outcome variables—RAI change, number of sessions, and needs satisfaction change—will be examined using *t* tests and Pearson correlations. The correlation between changes in RAI and needs satisfaction will also be evaluated. Analyses will be performed using SAS version 9.4 software (SAS Institute Inc). An overall significance level of .05 will be assumed for all tests.

Sample Size

The investigators plan to enroll 30 patients in each group. In the first 9 months of 2016, cardiac rehabilitation at the main campus of this tertiary care center had approximately 170 patient intakes. It is assumed that there will be a similar number of patient intakes for a 9-month period in 2017. Based on Facebook participation rates for those over 50 years of age [21] and the high participation rates in previous research projects in this facility's cardiac rehabilitation center, it is estimated that 40% may meet eligibility requirements and agree to participate. Allowing for use of the first 8 participants to establish the Facebook group, the estimated sample size would then be 60 total participants for randomization to study groups who can then be included in analysis. With this sample size, there will be 86% power to detect large effect sizes (Cohen d=0.8) for our study outcomes [40]. The primary aims of this sample size determination is to evaluate whether the proposed intervention is feasible and to estimate the differences that might exist so that a larger trial that would have adequate power to detect smaller differences could be designed based on what was learned in this pilot study. The sample size of 30 per group was chosen primarily to facilitate a large intervention group, since the value of the intervention is predicated upon interaction among the participants.

Human Subjects Protection

This feasibility study has been approved by the Institutional Review Board of this tertiary care center (Study No. 16-1456) and is registered at ClinicalTrials.gov (NCT02971813).

Participants will be assured that participation in the study at all times is voluntary and will not affect their care in any way. Protection of human subjects for this study will be further ensured through the process of informed consent. Participants in the intervention and comparison groups will be informed that privacy of medical information will be ensured. However, due to the nature of social media, information or comments posted by the patients in the Facebook group will be visible to others in the group as well as the study team. For this reason, the informed consent will address the fact that comments may be seen by others.

All responses from participants on the Facebook group will be assigned a number and all other identifying information will be removed for data analysis. Any data on paper will be kept in the principal investigator's (LAS) locked office in a locked filing cabinet. All electronic data will be stored on the principal investigator's computer, which requires password entry, in a folder accessible only to the principal investigator and the research team, and on an encrypted thumb drive. Dissemination of findings will be deidentified and reported numerically in narrative form or in aggregate, with no personal identifiers.

Results

This project was funded in July 2017 and recruitment is currently underway. The recruitment goal is 60 cardiac rehabilitation patients. Data collection is anticipated to be complete by July 2018.

Discussion

Overview

The main objective of this project will be to examine the feasibility of a novel Facebook intervention to address patient adherence to cardiac rehabilitation. Improving uptake and adherence to cardiac rehabilitation is of paramount importance in the secondary prevention of CHD. This study will use the validated BREQ-3 and the PNSE scale and examine the effect of a Facebook intervention on number of cardiac rehabilitation sessions attended. Applying the Self-Determination Theory, the research team will provide educational and provider support postings on a private Facebook group. The participants will have the opportunity to learn and interact with other participants in this social media platform. This study has the potential to affect a change in patient motivation for exercise and cardiac rehabilitation adherence, thus reducing complications and hospital readmissions among patients eligible for cardiac rehabilitation.

Limitations and Unanticipated Problems

Limitations for this study include the variable number of sessions paid for by non-Medicare and non-Medicaid insurance. This could potentially affect motivation or participation in the Facebook group and cardiac rehabilitation itself if the patient has few sessions that are covered by insurance. Feasibility concerns for the pilot include obtaining a large enough cohort of patients in order to have peer support, especially for those who enroll in the early stages of the study. Data will not be included for the first 8 participants in order to ensure that there

is a large enough group of Facebook users to enable social networking among participants. Additionally, patient visits to the Facebook group rely on self-report and are therefore subject to reporting bias.

There are limitations to this feasibility study that can be addressed in a larger trial. Patients who are not current Facebook users were excluded from the pilot trial. If Facebook is demonstrated to be a feasible venue for presenting and testing motivation for exercise, those who are not currently on Facebook should be included in a larger trial. The fact that patients may see each other in cardiac rehabilitation sessions presents a potential for diffusion bias, demoralization, or rivalry. This has been minimized to the extent that few participants are likely to communicate about the study in cardiac rehabilitation sessions due to the number of classes and facilities; however, it will need to be a consideration for this and larger studies.

Conclusions

The findings of this study will help to determine the feasibility of using a Facebook intervention to affect adherence and motivation. It has the possibility of opening doors to other technological interventions and unique approaches to improving health outcomes in this population. The results of this study will determine if a larger-scale intervention is feasible. Further, this pilot study will be the first to examine the effect of a Facebook intervention on patient adherence and motivation for exercise in a cardiac rehabilitation setting. The established private, cardiac rehabilitation Facebook group will enable a larger-scale intervention to be implemented and will allow for the examination of additional outcome variables. This intervention has the potential to add innovative approaches to the body of evidence seeking ways to improve patient outcomes in cardiac rehabilitation.

Conflicts of Interest

None declared.

References

1. Sanchis-Gomar F, Perez-Quilis C, Leischik R, Lucia A. Epidemiology of coronary heart disease and acute coronary syndrome. Ann Transl Med 2016 Jul;4(13):256 [FREE Full text] [doi: 10.21037/atm.2016.06.33] [Medline: 27500157]
2. Servey JT, Stephens M. Cardiac rehabilitation: Improving function and reducing risk. Am Fam Physician 2016 Jul 01;94(1):37-43. [Medline: 27386722]
3. Zeng W, Stason WB, Fournier S, Razavi M, Ritter G, Strickler GK, et al. Benefits and costs of intensive lifestyle modification programs for symptomatic coronary disease in Medicare beneficiaries. Am Heart J 2013 May;165(5):785-792. [doi: 10.1016/j.ahj.2013.01.018] [Medline: 23622916]
4. Savage PD, Sanderson BK, Brown TM, Berra K, Ades PA. Clinical research in cardiac rehabilitation and secondary prevention: Looking back and moving forward. J Cardiopulm Rehabil Prev 2011;31(6):333-341 [FREE Full text] [doi: 10.1097/HCR.0b013e31822f0f79] [Medline: 21946418]
5. Ribeiro F, Oliveira NL, Silva G, Campos L, Miranda F, Teixeira M, et al. Exercise-based cardiac rehabilitation increases daily physical activity of patients following myocardial infarction: Subanalysis of two randomised controlled trials. Physiotherapy 2017 Mar;103(1):59-65. [doi: 10.1016/j.physio.2015.12.002] [Medline: 27012822]
6. Coyan GN, Reeder KM, Vacek JL, Coyan GN, Reeder KM, Vacek JL. Diet and exercise interventions following coronary artery bypass graft surgery: A review and call to action. Phys Sportsmed 2014 May;42(2):119-129 [FREE Full text] [doi: 10.3810/psm.2014.05.2064] [Medline: 24875979]
7. Ghashghaei FE, Sadeghi M, Marandi SM, Ghashghaei SE. Exercise-based cardiac rehabilitation improves hemodynamic responses after coronary artery bypass graft surgery. ARYA Atheroscler 2012;7(4):151-156 [FREE Full text] [Medline: 23205048]
8. Williams MA, Ades PA, Hamm LF, Keteyian SJ, LaFontaine TP, Roitman JL, et al. Clinical evidence for a health benefit from cardiac rehabilitation: An update. Am Heart J 2006 Nov;152(5):835-841. [doi: 10.1016/j.ahj.2006.05.015] [Medline: 17070142]
9. Martin B, Hauer T, Arena R, Austford LD, Galbraith PD, Lewin AM, et al. Cardiac rehabilitation attendance and outcomes in coronary artery disease patients. Circulation 2012 Aug 07;126(6):677-687 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.066738] [Medline: 22777176]
10. American Association of Cardiovascular and Pulmonary Rehabilitation, American College of Cardiology Foundation, American Heart Association Task Force on Performance Measures (Writing Committee to Develop Clinical Performance Measures for Cardiac Rehabilitation), Thomas RJ, King M, Lui K, et al. AACVPR/ACCF/AHA 2010 update: Performance measures on cardiac rehabilitation for referral to cardiac rehabilitation/secondary prevention services endorsed by the American College of Chest Physicians, the American College of Sports Medicine, the American Physical Therapy Association, the Canadian Association of Cardiac Rehabilitation, the Clinical Exercise Physiology Association, the European Association for Cardiovascular Prevention and Rehabilitation, the Inter-American Heart Foundation, the National Association of Clinical Nurse Specialists, the Preventive Cardiovascular Nurses Association, and the Society of Thoracic Surgeons. J Am Coll Cardiol 2010 Sep 28;56(14):1159-1167 [FREE Full text] [doi: 10.1016/j.jacc.2010.06.006] [Medline: 20863958]

11. Siegmund LA, Naylor J, Bena J, McClelland M. The relationship between metabolic syndrome and adherence to cardiac rehabilitation. Physiol Behav 2017 Feb 01;169:41-45. [doi: 10.1016/j.physbeh.2016.11.005] [Medline: 27840095]
12. Karmali KN, Davies P, Taylor F, Beswick A, Martin N, Ebrahim S. Promoting patient uptake and adherence in cardiac rehabilitation. Cochrane Database Syst Rev 2014(6):CD007131. [doi: 10.1002/14651858.CD007131.pub3] [Medline: 24963623]
13. Rawstorn JC, Gant N, Meads A, Warren I, Maddison R. Remotely delivered exercise-based cardiac rehabilitation: Design and content development of a novel mHealth platform. JMIR Mhealth Uhealth 2016;4(2):e57 [FREE Full text] [doi: 10.2196/mhealth.5501] [Medline: 27342791]
14. Devi R, Powell J, Singh S. A Web-based program improves physical activity outcomes in a primary care angina population: Randomized controlled trial. J Med Internet Res 2014;16(9):e186 [FREE Full text] [doi: 10.2196/jmir.3340] [Medline: 25217464]
15. Joseph RP, Keller C, Adams MA, Ainsworth BE. Print versus a culturally relevant Facebook and text message-delivered intervention to promote physical activity in African American women: A randomized pilot trial. BMC Womens Health 2015;15:30 [FREE Full text] [doi: 10.1186/s12905-015-0186-1] [Medline: 25886945]
16. Cox NS, Alison JA, Button BM, Wilson JW, Holland AE. Feasibility and acceptability of an Internet-based program to promote physical activity in adults with cystic fibrosis. Respir Care 2015 Mar;60(3):422-429 [FREE Full text] [doi: 10.4187/respcare.03165] [Medline: 25425703]
17. Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]
18. Irwin B, Kurz D, Chalin P, Thompson N. Testing the efficacy of OurSpace, a brief, group dynamics-based physical activity intervention: A randomized controlled trial. J Med Internet Res 2016 May 06;18(4):e87 [FREE Full text] [doi: 10.2196/jmir.5342] [Medline: 27154301]
19. Cadmus-Bertram L, Irwin M, Alfano C, Campbell K, Duggan C, Foster-Schubert K, et al. Predicting adherence of adults to a 12-month exercise intervention. J Phys Act Health 2014 Sep;11(7):1304-1312. [doi: 10.1123/jpah.2012-0258] [Medline: 24176780]
20. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
21. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2017 Jan 12. Social media fact sheet URL: http://www.pewinternet.org/fact-sheet/social-media/ [accessed 2017-08-04] [WebCite Cache ID 6sTCpc5Hl]
22. Magnezi R, Bergman YS, Grosberg D. Online activity and participation in treatment affects the perceived efficacy of social health networks among patients with chronic illness. J Med Internet Res 2014;16(1):e12 [FREE Full text] [doi: 10.2196/jmir.2630] [Medline: 24413148]
23. Maher C, Ferguson M, Vandelanotte C, Plotnikoff R, De Bourdeaudhuij I, Thomas S, et al. A Web-based, social networking physical activity intervention for insufficiently active adults delivered via Facebook app: Randomized controlled trial. J Med Internet Res 2015;17(7):e174 [FREE Full text] [doi: 10.2196/jmir.4086] [Medline: 26169067]
24. Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31. [doi: 10.1002/oby.20232] [Medline: 23505165]
25. Struik LL, Baskerville NB. The role of Facebook in Crush the Crave, a mobile- and social media-based smoking cessation intervention: Qualitative framework analysis of posts. J Med Internet Res 2014;16(7):e170 [FREE Full text] [doi: 10.2196/jmir.3189] [Medline: 25016998]
26. Cavallo DN, Tate DF, Ries AV, Brown JD, DeVellis RF, Ammerman AS. A social media-based physical activity intervention: A randomized controlled trial. Am J Prev Med 2012 Nov;43(5):527-532 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.019] [Medline: 23079176]
27. Cavallo DN, Tate DF, Ward DS, DeVellis RF, Thayer LM, Ammerman AS. Social support for physical activity-role of Facebook with and without structured intervention. Transl Behav Med 2014 Dec;4(4):346-354 [FREE Full text] [doi: 10.1007/s13142-014-0269-9] [Medline: 25584083]
28. Balatsoukas P, Kennedy CM, Buchan I, Powell J, Ainsworth J. The role of social network technologies in online health promotion: A narrative review of theoretical and empirical factors influencing intervention effectiveness. J Med Internet Res 2015;17(6):e141 [FREE Full text] [doi: 10.2196/jmir.3662] [Medline: 26068087]
29. Deci EL, Ryan RM. The support of autonomy and the control of behavior. J Pers Soc Psychol 1987 Dec;53(6):1024-1037. [Medline: 3320334]
30. Ryan RM, Deci EL. Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemp Educ Psychol 2000 Jan;25(1):54-67. [doi: 10.1006/ceps.1999.1020] [Medline: 10620381]
31. Deci EL, Ryan RM. In: Aronson E, editor. Intrinsic Motivation and Self-Determination in Human Behavior. New York, NY: Plenum Press; 1985.
32. Haakma I, Janssen M, Minnaert A. Understanding the relationship between teacher behavior and motivation in students with acquired deafblindness. Am Ann Deaf 2016;161(3):314-326. [doi: 10.1353/aad.2016.0024] [Medline: 27477039]

33. Wilson PM, Rodgers WM, Loitz CC, Scime G. “It's who I am...really!” The importance of integrated regulation in exercise contexts. J Appl Biobehav Res 2006 Apr;11(2):79-104. [doi: 10.1111/j.1751-9861.2006.tb00021.x]
34. DeCharms R. Personal Causation: The Internal Affective Determinants of Behavior. New York, NY: Academic Press; 1968.
35. Lindwall M, Ivarsson A, Weman-Josefsson K, Jonsson L, Ntoumanis N, Patrick H, et al. Stirring the motivational soup: Within-person latent profiles of motivation in exercise. Int J Behav Nutr Phys Act 2017 Jan 14;14(1):4 [FREE Full text] [doi: 10.1186/s12966-017-0464-4] [Medline: 28088208]
36. Thorup CB, Grønkjær M, Spindler H, Andreasen JJ, Hansen J, Dinesen BI, et al. Pedometer use and self-determined motivation for walking in a cardiac telerehabilitation program: A qualitative study. BMC Sports Sci Med Rehabil 2016;8:24 [FREE Full text] [doi: 10.1186/s13102-016-0048-7] [Medline: 27547404]
37. Wilson P, Rogers WT, Rodgers WM, Wild T. The Psychological Need Satisfaction in Exercise scale. J Sport Exerc Psychol 2006 Sep;28(3):231-251. [doi: 10.1123/jsep.28.3.231]
38. Markland D, Tobin V. A modification to the Behavioural Regulation in Exercise Questionnaire to include an assessment of amotivation. J Sport Exerc Psychol 2004 Jun;26(2):191-196. [doi: 10.1123/jsep.26.2.191]
39. Markland D. Bangor University. 2014 Oct 24. Exercise motivation measurement URL: http://pages.bangor.ac.uk/~pes004/exercise_motivation/scales.htm [accessed 2017-08-04] [WebCite Cache ID 6sTETAp7P]
40. 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]

Abbreviations

ANCOVA: analysis of covariance
BREQ-2: Behavioral Regulation in Exercise Questionnaire-2
BREQ-3: Behavioral Regulation in Exercise Questionnaire-3
CHD: coronary heart disease
PNSE: Psychological Need Satisfaction in Exercise
RAI: relative autonomy index

Edited by G Eysenbach; submitted 13.04.17; peer-reviewed by S Sweet, R Rahman, N ter Hoeve, A Venkatakrishnan; comments to author 01.06.17; revised version received 16.06.17; accepted 12.07.17; published 18.08.17.

Please cite as:
Siegmund LA, Ahmed HM, Crawford MT, Bena JF
Feasibility of a Facebook Intervention for Exercise Motivation and Cardiac Rehabilitation Adherence: Study Protocol
JMIR Res Protoc 2017;6(8):e162
URL: http://www.researchprotocols.org/2017/8/e162/
doi:10.2196/resprot.7554
PMID:28821473

Original Paper

Conventional Cognitive Behavioral Therapy Facilitated by an Internet-Based Support System: Feasibility Study at a Psychiatric Outpatient Clinic

Kristoffer NT Månsson[1,2,3], PhD; Hugo Klintmalm[4], MSc; Ragnar Nordqvist[4], MSc; Gerhard Andersson[3,5], PhD

[1]Department of Psychology, Stockholm University, Stockholm, Sweden

[2]Department of Adult Psychiatry, PRIMA Barn- och Vuxenpsykiatri, Stockholm, Sweden

[3]Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

[4]Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

[5]Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

Corresponding Author:
Kristoffer NT Månsson, PhD
Department of Psychology
Stockholm University
Frescati Hagväg 8
Stockholm, 106 91
Sweden
Phone: 46 (0)705803267
Email: kristoffer.mansson@psychology.su.se

Abstract

Background: Cognitive behavioral therapies have been shown to be effective for a variety of psychiatric and somatic disorders, but some obstacles can be noted in regular psychiatric care; for example, low adherence to treatment protocols may undermine effects. Treatments delivered via the Internet have shown promising results, and it is an open question if the blend of Internet-delivered and conventional face-to-face cognitive behavioral therapies may help to overcome some of the barriers of evidence-based treatments in psychiatric care.

Objective: We evaluated the feasibility of an Internet-based support system at an outpatient psychiatric clinic in Sweden. For instance, the support system made it possible to send messages and share information between the therapist and the patient before and after therapy sessions at the clinic.

Methods: Nine clinical psychologists participated and 33 patients were enrolled in the current study. We evaluated the usability and technology acceptance after 12 weeks of access. Moreover, clinical data on common psychiatric symptoms were assessed before and after the presentation of the support system.

Results: In line with our previous study in a university setting, the Internet-based support system has the potential to be feasible also when delivered in a regular psychiatric setting. Notably, some components in the system were less frequently used. We also found that patients improved on common outcome measures for depressive and anxious symptoms (effect sizes, as determined by Cohen *d*, ranged from 0.20-0.69).

Conclusions: This study adds to the literature suggesting that modern information technology could be aligned with conventional face-to-face services.

(JMIR Res Protoc 2017;6(8):e158) doi:10.2196/resprot.6035

KEYWORDS

cognitive behavioral therapy; Internet-treatment; psychiatry; blended therapy

Introduction

During the last decade, there has been a growing interest in alternative ways of delivering psychological treatments. The development of Internet-delivered interventions targeting common psychiatric and somatic disorders is one promising method [1,2]. Therapist-guided Internet-delivered treatments based on cognitive behavioral therapy (ICBT) have commonly

shown promising effects in studies of both research studies (efficacy) [3], and in more clinically representative settings (effectiveness) [4]. A growing body of evidence suggests similar outcomes of ICBT and conventional face-to-face cognitive behavioral therapies (CBT) [2], with therapist-guided ICBT being less time-consuming for the clinician. Using the Internet to deliver health care may open new avenues to treatment, especially in societies where the distance to care is far away. Thus, ICBT has the potential to increase access to evidence-based psychological treatment [3].

In primary or psychiatric care, there may be some obstacles of providing conventional CBT delivered face-to-face. For instance, therapists may be prone to drift away from implementing effective interventions (ie, therapist drift) [5], and they may also fail to adhere to evidence-based treatment manuals [6]. One way to overcome such obstacles could be to provide computer-assisted support in therapeutic work [7]. In a previous study, we developed an Internet-based support system to facilitate the delivery of conventional CBT [8]. The basic idea of the system is to support the delivery of CBT in a clinical setting where the therapist meets their patients face to face. By providing support, our objective was to improve the delivery of regular treatment components present in CBT, for example, homework assignments. A potential strength of the approach is that it conceptually shifts the focus of research away from specific digital interventions towards the system level (ie, capable of delivering many interventions). The approach also highlights the potential impact of introducing digital communication channels in face-to-face psychotherapy. The initial study showed some promising findings in terms of user experiences (eg, the ease of providing written information as a complement to the therapy sessions), and we observed reliable reductions of depressive and anxious symptoms. The study was conducted in a university setting, and there is a need to test the support system in clinical psychiatric care (eg, with a more severe clinical population and across different disorders).

This feasibility study aimed at evaluating the experiences and effects of an Internet-based support system used as an adjunct to conventional CBT delivered face to face. The system was designed to support the delivery of face-to-face CBT and not replace in-session treatment activities. The system was used for communication between therapy sessions, sharing media, and clarifying homework assignments [8]. Clinicians and patients were recruited from a psychiatric clinic in Sweden, and the users were given access to the support system during 12 weeks. At follow-up, we evaluated support system usability and technology acceptance. Moreover, self-report questionnaires targeting clinical symptoms at baseline and 12-week follow-up were also administered.

Methods

Procedure

Nine clinical psychologists participated as therapists in the study. The clinicians were asked to recruit patients from the clinic in accordance with the standard procedures at the clinic. In line with the ethics committee agreement (ID: 2013/452-31), all patients were informed about the objectives of the study via a document printed on paper and asked to provide written informed consent before inclusion. All patients answered questionnaires regarding clinical and demographic characteristics via the Internet. After inclusion, the clinician registered the patient in the support system and distributed an online follow-up survey after 12 weeks access of the support system. Mean time between assessments was 91 days (range 61-116).

The Support System

The Internet-based support system used in this study was previously developed and tested in a pilot study conducted in a university clinic setting [8]. Also, the support system has been used in audiologic practice in supporting first-time hearing aid clients [9]. In brief, the support system was accessible via personal computers through an encrypted secured socket layer connection to the Internet. Users were assigned personal login identifications via email. Also, to increase security, an additional temporary password was sent via mobile phone text messages at each attempt to log on.

The support system facilitated a variety of functions and the therapists decided themselves on how to use the content, tailored to the patients' needs, and components included communication between sessions with the ability to send mobile phone text messages. Via the support system, the therapist also had the opportunity to send mobile phone text messages to the patients. The support system included a library that mainly provided text documents, but also other media such as audio and movies were made accessible. These resources were compiled primarily from prior studies on Internet-delivered CBT for anxiety and depression [10], and they were not presented as separate treatments but rather as part of the face-to-face treatment (eg, as online handouts). Topics covered in the online handouts contained supplemental information on CBT, such as behavioral activation, activity scheduling, exposure therapy, common cognitive biases, and maintenance of avoidance via safety behaviors. We also provided some audio files, such as relaxation instructions. In addition, the support system included common questionnaires and forms used in homework assignments, such as guides to create a fear hierarchy, daily thought records, and sleep diaries. For an overview of all the functions, see Table 1.

Table 1. List of included components in the Internet-based support system.

	Editable by[a]	
	Patient	Clinician
Formulate homework assignments	No	Yes
Library (sharing content)	No	Yes
Registration forms	Yes	Yes
Sending internal messages	Yes	Yes
Sending mobile phone text messages	No	Yes
Setting an agenda	No	Yes
Setting goals	Yes	Yes
Uploading files to the patient's personal library	Yes	Yes
Uploading new files to the library	No	Yes
Writing memos	Yes	Yes

[a]The term editable means that the user was able to edit, add, or delete content within that specific component/function in the support system.

Technical Issues

During the study period, we had one main technical problem with the support system. As a way of warranting the security of the support system, it was designed to automatically log out inactive users (as determined by no clicks with the pointer). First, the support system automatically disconnected users after 10 minutes of inactivity. A number of users gave us feedback that text had been lost due to this function (eg, while writing a long message exceeding 10 minutes, the user was incorrectly disconnected). Consequently, we increased this time frame to 40 minutes during the study period.

Before study initiation, we invited a group of clinicians to a 2-hour workshop offering a brief overview of the support system. Also, the clinicians logged in to the system and were instructed to complete five tasks in order to acquire some knowledge on basic functions in the support system, for example, log in to the system, create a new user (patient), send the patient a message, share a file from the library with the patient, as well as a registration form for behavioral experiments.

Participants and Recruitment

The included clinicians' professional status and demographic characteristics are presented in Table 2. The clinicians volunteered and did not receive any compensation for their participation.

During the study period, 52 patients were registered in the support system. However, data from 4 patients were missing at the baseline assessment, 12 patients were missing at follow-up, and for 7 patients assessment data were completely missing (ie, both at baseline and follow-up). In total, 29 patients contributed with complete data from the pre- and posttreatment assessments. The patients' demographic characteristics and computer experience at baseline are presented in Table 3. Participants self-rated their level of experience of using computers on a 5-point Likert scale, ranging from 0 (very limited) to 4 (very much). We did not include any clinical interview in order to determine diagnostic criterion and comorbidity. The patients received treatment but no compensation for participating in the study.

Table 2. Demographic and professional characteristics of the clinicians (n=9).

Characteristics	
Age in years, mean (median, SD, range)	37.78 (35, 9.2, 28-54)
Year as licensed psychologist, mean (median, SD, range)	2.44 (2, 2.5, 0-7)
Sex, female, n (%)	3 (33.3)
Professional status, n (%)	
Pre-licensed under supervision	4 (44.4)
Licensed clinical psychologist	5 (55,6)
Clinical work (% per month), n (%)	
0-25%	3 (33.3)
26-50%	1 (11.1)
76-100%	5 (55.6)

Table 3. Demographic and clinical characteristics of the patients (n=45).

Characteristics	
Age in years, mean (median, SD, range)	30.58 (28,10.6, 18-60)
Work time (% per month), mean (median, SD, range)	36.11 (40, 30.7, 0-100)
Sex, female, n (%)	36 (80.0)
Having children, n (%)	29 (64.4)
Computer experience, n (%)	
Less	11 (24.4)
More	34 (75.6)
Educational status, n (%)	
<High school	22 (48.9)
>High school	23 (51.1)

The patients were either recruited from an existing wait-list at the clinic or were currently undergoing a conventional CBT at the clinic. In order to receive treatment at the psychiatric clinic, the patients had to be over 18 years of age. Eligible patients in this study were required to have some computer experience (ie, being able to handle their bank account via the Internet) and have access to a computer and mobile phone during the study period. Patients not considered eligible, or denied participation in the study, were offered conventional face-to-face CBT in line with routines at the clinic.

All procedures contributing to this work comply with the standards of the national ethical committee and with the Helsinki Declaration of 2008.

Cognitive Behavioral Therapy

This study did not follow a manualized CBT protocol, nor did all the clinicians receive clinical supervision as part of the study. The clinicians tailored the CBT according to their patient's needs (eg, based on cognitive case formulation or behavior analysis) and each clinician-patient pair individually decided how to use the Internet-based support system during the treatment.

We evaluated the use of the support system during a period of 12 weeks. Therefore, our assessments at baseline and 12-week follow-up were not fixed at pre- and posttreatment (ie, at baseline, some patients had already started CBT, and for some patients the CBT was not terminated at the 12-week follow-up).

Support System Usability

For all the users (ie, clinicians and patients), we monitored the number of logins, the total time spent logged in, as well as the number of messages sent within the support system. After 12 weeks of accessing the support system, we evaluated the users' experiences. We also asked questions targeting specific functions within the support system, for example, how often the participant read and downloaded text documents, listened to audio files from the library, set goals for the treatment, asked questions, and requested guidance via internal messages. The questions were rated on a 6-point scale ranging from never to very often. In addition, the clinicians were also asked to rate for how many of their patients the features in the support system had been, or would have been, relevant for their patients in their regular clinical practice, ranging from no one, less than 50%, more than 50%, or for most patients.

Technology Acceptance, Perceived Usefulness, and Ease of Use

We used 19 questions targeting usability of the Internet-based support system. The questions were adopted from questionnaires of technology acceptance [11], perceived usefulness, and perceived ease of use [12] and were translated into Swedish. We used only a sample of questions and customized them to fit the current study. All questions were rated on a 7-point Likert scale ranging from "Strongly disagree" to "Strongly agree." All participants were asked to answer these questions (ie, both clinicians and patients).

Clinical Outcome and Quality of Life

The Beck Anxiety Inventory (BAI) [13] and the Generalized Anxiety Disorder Screener-7 (GAD-7) [14] were used both at baseline and as outcome measures of anxiety symptoms. Both questionnaires have been shown to have excellent internal consistency (Cronbach alpha >.90) [13,14]. The Montgomery Åsberg Depression Rating Scale self-rating version (MADRS-S) [15] and the Patient Health Questionnaire-9 (PHQ-9) [16] were used to measure symptoms of depression and suicidality. MADRS-S and PHQ-9 also have excellent internal consistency (alpha >.89) [15,16]. In the MADRS-S, suicidality was defined as a score of at least three points on item 9. Similarly, patients scoring one point (or above) on item 9 on the PHQ-9 were also considered suicidal in this study.

In addition to change in symptoms of anxiety and depression, the Quality of Life Inventory (QOLI) [17] was administered both at baseline and at 12-week follow-up. QOLI has shown good to excellent internal consistency (alpha >.77) in a clinical population with both anxious and depressive disorders [18]. In agreement with our previous studies [3], all self-report questionnaires were administered via a secured Internet-based platform.

Data Analysis

The STATA v13.1 statistical software for Mac OS X (StataCorp) was used to analyze the data. We evaluated user experiences

across patients with high versus low activity in the support system and dichotomized high versus low frequent users by performing a median split on number of times the patients accessed the support system (ie, ≥12 defined high users). Differences between users (ie, low versus high activity) groups (ie, clinicians versus patients) were analyzed using logistic regression.

We also performed analyses on clinical outcome of anxious and depressive symptoms. Similarly, quality of life was measured at baseline and 12-week follow-up. In order to account for dependency in the data (ie, longitudinal clinical outcomes), we used generalized estimating equations (GEE) with an exchangeable correlation structure, assuming that all missing data were completely at random [19]. Outcomes are presented as coefficients or odds ratios (OR). Within-group effect sizes were calculated based on the pooled standard deviation and correlation between time points, expressed as Cohen *d* with 95% confidence intervals. Furthermore, we also investigated if the number of times accessing the support system was associated with change in the patient's symptoms of anxiety and depression.

As a way to control for multiple comparisons, we performed Bonferroni corrections within each sector of the analyses (ie, one sector corresponds to support system usability, and another was clinical outcome).

Furthermore, we explored what time of the day the patients accessed the support system. Specifically, we were interested in the proportion of logins made after the clinic was closed (ie, before 8 a.m. and after 5 p.m.).

Results

Support System Usability

Clinicians

The mean number of times the clinicians accessed the support system during the 12-week period was 94 (SD 54, median 89), and across all the clinicians the average time logged in to the support system was 1008 minutes (16.8 h, SD 784 min, median 770 min). On average, 64 messages were sent per clinician (SD 25, median 62, range 17-100). Moreover, the mean number of sent mobile phone text messages was 32 (SD 14, median 35, range 9-51).

As shown in Table 4, the clinicians' ratings of usability demonstrate how often specific components were assigned to the patient, as well as the proportion of patients for whom this component was considered relevant in the therapeutic work. For example, sharing forms and studying information in the library for own professional development were on average used 2.8 times (ie, less used than "sometimes"). Yet, most of the clinicians rated these functions to be relevant for more than 50% of their patients.

Table 4. Clinicians' (n=9) evaluation of support system usability on a 6-point Likert scale (0=never and 5=very often), sorted by mean values.

Questions			The component was relevant for n patients, n (%)	
	Mean	SD	Less than 50%	More than 50%
Sending reminders via mobile phone text messages	4.11	1.0	3 (33.3)	6 (66.7)
Shared documents, images, and audio files via the library	4.00	1.0	3 (33.3)	6 (66.7)
Answered questions	3.89	1.0	4 (44.4)	5 (55.6)
Providing support and encouragement	3.89	0.9	4 (44.4)	5 (55.6)
Formulated homework assignments	3.78	1.1	5 (55.6)	4 (44.4)
Provided psychoeducation from the library	3.44	1.9	4 (44.4)	5 (55.6)
Reading the patients reports on homework assignments	3.44	0.7	5 (55.6)	4 (44.4)
Asked for feedback on information in the library	3.22	1.2	7 (78.0)	2 (22.2)
Examined the patient work with homework	3.11	0.9	5 (55.6)	4 (44.4)
Studied information from the library for own professional development	2.88	1.3	4 (44.4)	5 (55.6)
Distributed registration forms	2.78	1.5	4 (44.4)	5 (55.6)
Reviewed homework assignments reported by the patient	2.78	1.3	7 (78.0)	2 (22.2)
Worked with assignments from the library	2.44	1.2	8 (89.0)	1 (11.1)
Corrected and revised homework assignments	2.44	1.4	6 (66.7)	3 (33.3)
Formulated goals for therapy	1.88	1.2	7 (78.0)	2 (22.2)
Setting an agenda	1.44	1.3	7 (78.0)	2 (22.2)
Play audio during the therapy session	0.11	0.3	9 (100)	0 (0)

Table 5. The patients' (n=33) evaluation of support system usability on a 6-point Likert scale (0=never and 5=very often), sorted by mean values.

Questions	Mean	SD	Median
Proportion of completed homework assignments	3.27	1.3	4
Accessed your psychologists formulations of homework assignments	3.18	1.6	4
Provided information about the progress of your homework assignments[a]	2.33	1.7	3
Asked for guidance via internal messages[a]	2.18	1.8	3
Answered forms	2.15	1.7	3
Downloaded and saved information on your computer	1.91	1.7	2
Reading information from the library	1.88	1.6	2
Reading and reviewed treatment goals during the therapy	1.76	1.7	1
Printed documents	1.42	1.7	1
Reading the agenda	1.21	1.6	0
Saved your own therapy-related information (text and/or images) in your personal library	1.03	1.4	0
Listened to audio files	0.90	1.4	0
Wrote notes regarding questions to discuss with your psychologist	0.33	0.8	0
Wrote memos	0.30	0.8	0

[a]Indicating differences between the high versus low frequent users.

Patients

Across 12 weeks of access, the patients' average number of logins was 14 (SD 15.3, median 11, range 1-95), and they (n=49) spent on average 92 minutes (SD 157, median 42) on the support system. One patient was an outlier and spent more than 1000 minutes logged into the support system. After excluding this outlier, the average number of minutes was reduced to 72 (SD 72, median 40), which corresponds to an average of 6 minutes of access per week (72/12) and patient. In addition, the patients sent on average 6 messages to their therapist (SD 10, median 3, range 0-58), although there is a large variation across users.

The patient's usability ratings of specific components in the support system are presented in Table 5. High and low frequent users ratings differed significantly on two items: (1) providing information about the progress homework assignments (high users mean 3.1, SD 1; low users mean 1.6, SD 2; β=0.64, Z=2.40, P=.02), and (2) asked for guidance via internal messages (high users mean 3.1, SD 1; low users mean 1.2, SD 2; β=0.70, Z=2.71, P=.002). However, after controlling for multiple comparisons (ie, Bonferroni correction) the differences were not significant.

Technology Acceptance, Perceived Usefulness, and Ease of Use

The clinician and the patient ratings of technology acceptance, perceived usefulness, and ease of use are shown in Table 6. The clinicians and the patients rated two items significantly differently. First, the clinicians were more motivated to use the support system after the study termination (β=0.68, Z=2.10, P=.036). Second, the patients, relative to the clinicians, highlighted that the support system reminded them about tasks to complete in the support system (β=–0.50, Z=2.37, P=.018). However, by controlling for multiple comparisons, these differences were not significant.

Clinical Outcome and Quality of Life

Total scores on the BAI, MADRS-S, and PHQ-9 decreased from baseline to 12-week follow-up, yet the GAD-7 only showed a trend towards statistical significance. Moreover, quality of life, as measured by QOLI, increased over time (see Table 7).

Suicidal ideations, as measured by MADRS-S item 9, decreased by 14% from baseline to follow-up (OR 0.86, Z= 2.05, P=.040). However, the scored item on suicidal ideation in PHQ-9 did not change over time (OR 0.89, Z=1.43, P=.152). With the exception of change on MADRS-S suicidality and QOLI, the other results on clinical symptoms remained statistically significant following Bonferroni correction (P<.05).

We did not find that the number of times accessing the support system was associated with any change in clinical symptoms of anxiety or depression. We found that 30.52% (420/1376) of the patients' logins were made after working hours at the clinic.

Table 6. Questionnaire targeting technology acceptance and ease of use of the support system. Ratings provided on a 7-point Likert scale (1=strongly disagree to 7=strongly agree).

Questions	Patients (n=33)		Clinicians (n=9)	
	Mean	SD	Mean	SD
1. Using the platform improves the quality of the work I do	4.09	1.9	5.22	1.0
2. The platform enables me to accomplish tasks more quickly	3.61	1.9	3.56	1.6
3. The platform increases my productivity in my work with the therapy	3.91	1.9	4.00	1.4
4. The platform improves my work and the things I need to do in therapy, such as homework assignments and practical exercises	4.58	2.0	4.56	1.0
5. Using the platform can increase my effectiveness working with the therapy	4.09	1.8	4.44	0.9
6. Using the platform makes it easier to do my tasks	4.70	2.0	4.11	1.6
7. Learning to operate the platform is easy for me	5.67	1.9	5.00	1.4
8. Performing an operation in the platform always leads to a predicted result	4.12	1.8	3.22	1.5
9. The platform has a clear interface that helps me do what I want	4.52	1.9	3.22	1.1
10. The platform is flexible and easy to interact with	4.33	1.8	3.11	1.8
11. It is easy for me to become skillful at using the platform	4.42	1.9	4.89	1.0
12. Overall, I find the platform easy to use	5.30	1.9	4.67	1.6
13. I feel confident in finding information in the platform	5.12	1.8	4.22	1.3
14. I feel confident in receiving and sending messages	5.24	1.8	5.33	1.6
15. I feel confident in downloading files	5.12	2.1	4.78	1.8
16. The platform was visually appealing	3.58	1.9	3.78	1.8
17. The platform reminds me about tasks to complete	4.79	2.1	2.67	1.9
18. The organization of information in the platform is clear	4.45	1.9	3.44	2.0
19. I would like to use the platform on a regular basis in the future	4.36	2.3	6.44	0.9

Table 7. Generalized estimating equations (GEE) regarding clinical symptoms and quality of life at baseline, and 12-week follow-up (total N=33 patients in the GEE; 29 contributed with complete data from baseline to follow-up).

Measure	Pre, mean (SE)	Post, mean (SE)	Coefficient	*Z*	*P*	Cohen *d*[a] (95% CI)
BAI	20.54 (1.7)	15.82 (1.8)	-4.72	-2.88	.004	0.45 (0.1 to 0.8)
GAD-7	10.35 (0.8)	9.15 (0.9)	-1.20	-1.77	.077	0.20 (-0.1 to 0.5)
MADRS-S	20.96 (1.3)	18.18 (1.4)	-2.78	-2.61	.009	0.32 (0.0 to 0.6)
PHQ-9	13.32 (0.9)	9.19 (1.0)	-4.13	-3.68	<.001	0.69 (0.2 to 1.1)
QOLI	-0.37 (0.3)	0.03 (0.3)	0.40	2.12	.034	-0.18 (-0.4 to 0.1)

[a]Effect sizes (Cohen *d*) were calculated on observed data.

Discussion

Principal Findings

The aim of this feasibility study was to test an Internet-based support system in a clinical psychiatric setting with a focus on both clinician and patient experiences, and also including patient outcomes. Overall, we found that clinicians, as compared to the patients, rated some functions of the support system as more useful and that ratings by patients tended to be fairly low for some functions. We also asked clinicians to rate the proportion of patients for which the components of the support system would be useful. Less than half of the clinicians rated that the components would be useful for more than half of their patients. As there were few clinicians in the study, these estimates should be interpreted with caution but at least they signal that some functions, like sending reminders and sharing documents, may be appealing to clinicians in their work. At the same time, using the support system to formulate therapy goals, agenda setting, and playing audio files were barely used by the clinicians. Overall, usefulness ratings, ease of use, and technology acceptance varied but were fairly high for some items. Moreover, 30% of the times the patients accessed the support system were after working hours at the clinic. This indicated that this support system also has the potential to increase the availability of psychiatric care. In line with a large body of literature on the effects of CBT and ICBT, symptom ratings decreased over the study period.

This feasibility study raises many questions. First of all, the support system tested in our first study [8] appears to work when delivered in a more regular psychiatric setting with regular clinicians and patients [20]. Yet, it is important to keep in mind that a few specific functions within the system were rarely used by the clinicians. We hesitate to refer to this study as an effectiveness study as use of the support system per se was not part of regular practice, and we introduced and tested the system simultaneously. In the first study [8], we had a smaller sample and used interviews to gather information on experiences of clinicians and patients. In this study, we investigated differences between high versus low activity users and differences between the clinicians and the patients. No difference turned out to be statistically significant after controlling for multiple comparisons but indicated that the high frequent users more often sent messages to their therapist.

A second aspect to discuss relates to attitudes towards technology use and preferences (eg, [21,22]). There is a growing literature on these topics relating to ICBT, but far less work on the use of technology within face-to-face CBT, sometimes referred to as blended treatments [23], has been conducted. In addition, a recent stakeholder survey indicated that blended treatments are rated as more acceptable than ICBT with less therapist contact [24]. We expect more studies to appear in the field of blended interventions [23]. One recent example was a study on depression in which a mobile phone app was used [25]. Moreover, in this study we did not really focus on the technical aspects of the system, and there has been increased interest in the use of novel technologies and how they can be best incorporated and correctly described in digital health interventions [26].

Third, what can we expect to achieve with the support system? There is emerging literature on knowledge acquisition in CBT [27], and we believe that the support system can serve as a facilitator for patients when they learn more about themselves and the treatment presented. This might not necessarily lead to better outcomes in the short run but is also unlikely to lead to worse outcomes. In the long run, it is possible that the enhanced learning and support provided by the system could help to prevent relapse.

Fourth, this study raises questions regarding training of therapists and adherence to treatment manuals. It is possible that clinicians with less training can benefit more from blending information technology with face-to-face services. There are examples of using computerized support [7] with good outcomes, but to the best of our knowledge, there is a lack of controlled trials testing if clinicians with less training can perform as well as more skilled clinicians if they work with a support system. More experienced and well-trained clinicians may also be more effective if tasks can be delegated to the computer (eg, handling outcome measures).

Limitations

It is important to keep in mind that this study is limited by a number of factors. First, the within-group design limits any causal inferences, and we cannot answer whether or not the support system made any substantial, positive or negative, contributions beyond the effect of conventional CBT. However, in this feasibility study (without a predefined feasibility criteria) we focused on usability and technology acceptance at an outpatient psychiatric clinic. Second, the clinicians decided whether or not to ask a specific patient about participating in the study (ie, possible self-selection bias). Consequently, it is possible that the outcome of the study is affected by confounding by indication. Third, in terms of CBT interventions, we cannot demonstrate the specific interventions the clinicians delivered. Moreover, we did not measure the therapist's competence in delivering CBT. Nevertheless, by the use of the current support system we were able to monitor the use of some fundamental CBT components (eg, that the clinicians provided homework assignments). Fourth, the number of patients lost to follow-up may be an important sign of dissatisfaction. Nevertheless, it is plausible that this was related to issues regarding procedure of the study (eg, the clinicians were primarily responsible for initiating the follow-up assessments). In our previous study in a university setting, we had no missing data.

Conclusions

In spite of the limitations, this study adds to the literature showing that modern information technology can be aligned with conventional face-to-face services. Future studies should investigate the added value of using a support system in psychiatric care. Another option is to evaluate the usability of the support system when training new therapists.

Acknowledgments

We thank Cecilia Olsson Lynch, Hugo Klintmalm, Mikael Salomonsson, Nahla Elkholy, Pelle Alexandersson, Ragnar Nordqvist, Rikard Sunnhed, Susanne Westh, and Thomas Hesslow who volunteered as clinical psychologists. Anders Berntsson and Dr Anna Åberg Wistedt made significant contributions initiating the current research project. Also, many thanks to Dr Sarah-Marie Vigerland for contributions with proofreading. We gratefully thank all the patients participating in the study.

PRIMA Psychiatry Research Foundation funded this research (KM). GA is funded by a professor grant at Linköping University.

Conflicts of Interest

None declared.

References

1. Andersson G. Using the Internet to provide cognitive behaviour therapy. Behav Res Ther 2009 Mar;47(3):175-180. [doi: 10.1016/j.brat.2009.01.010] [Medline: 19230862]
2. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
3. Andersson G. The Internet and CBT: a clinical guide. Boca Raton: CRC Press; 2014:1-144.
4. Andersson G, Hedman E. Effectiveness of Guided Internet-Based Cognitive Behavior Therapy in Regular Clinical Settings. Verhaltenstherapie 2013;23(3):140-148. [doi: 10.1159/000354779]
5. Waller G. Evidence-based treatment and therapist drift. Behav Res Ther 2009 Feb;47(2):119-127. [doi: 10.1016/j.brat.2008.10.018] [Medline: 19036354]
6. Tobin DL, Banker JD, Weisberg L, Bowers W. I know what you did last summer (and it was not CBT): a factor analytic model of international psychotherapeutic practice in the eating disorders. Int J Eat Disord 2007 Dec;40(8):754-757. [doi: 10.1002/eat.20426] [Medline: 17610255]
7. Roy-Byrne P, Craske MG, Sullivan G, Rose RD, Edlund MJ, Lang AJ, et al. Delivery of evidence-based treatment for multiple anxiety disorders in primary care: a randomized controlled trial. JAMA 2010 May 19;303(19):1921-1928 [FREE Full text] [doi: 10.1001/jama.2010.608] [Medline: 20483968]
8. Månsson KNT, Skagius RE, Gervind E, Dahlin M, Andersson G. Development and initial evaluation of an Internet-based support system for face-to-face cognitive behavior therapy: a proof of concept study. J Med Internet Res 2013 Dec 10;15(12):e280 [FREE Full text] [doi: 10.2196/jmir.3031] [Medline: 24326065]
9. Brännström KJ, Öberg M, Ingo E, Månsson KNT, Andersson G, Lunner T, et al. The Process of Developing an Internet-Based Support System for Audiologists and First-Time Hearing Aid Clients. Am J Audiol 2015 Sep;24(3):320-324. [doi: 10.1044/2015_AJA-14-0094] [Medline: 26649539]
10. Andersson G, Carlbring P, Ljótsson B, Hedman E. Guided Internet-Based CBT for Common Mental Disorders. J Contemp Psychother 2013 May 21;43(4):223-233. [doi: 10.1007/s10879-013-9237-9]
11. Roca JC, Chiu C, Martínez FJ. Understanding e-learning continuance intention: An extension of the Technology Acceptance Model. International Journal of Human-Computer Studies 2006 Aug;64(8):683-696. [doi: 10.1016/j.ijhcs.2006.01.003]
12. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
13. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [Medline: 3204199]
14. 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]
15. Svanborg P, Asberg M. A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. Acta Psychiatr Scand 1994 Jan;89(1):21-28. [Medline: 8140903]
16. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
17. Frisch MB, Cornell J, Villanueva M, Retzlaff PJ. Clinical validation of the Quality of Life Inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. Psychological Assessment 1992;4(1):92-101. [doi: 10.1037/1040-3590.4.1.92]
18. Lindner P, Andersson G, Ost L, Carlbring P. Validation of the internet-administered Quality of Life Inventory (QOLI) in different psychiatric conditions. Cogn Behav Ther 2013;42(4):315-327. [doi: 10.1080/16506073.2013.806584] [Medline: 23837710]
19. Little R, Rubin D. Statistical analysis with missing data. Hoboken, NJ: Wiley; 2002.
20. Shadish WR, Matt GE, Navarro AM, Phillips G. The effects of psychological therapies under clinically representative conditions: a meta-analysis. Psychol Bull 2000 Jul;126(4):512-529. [Medline: 10900994]
21. Gun SY, Titov N, Andrews G. Acceptability of Internet treatment of anxiety and depression. Australas Psychiatry 2011 Jun;19(3):259-264. [doi: 10.3109/10398562.2011.562295] [Medline: 21682626]
22. Mohr DC, Siddique J, Ho J, Duffecy J, Jin L, Fokuo JK. Interest in behavioral and psychological treatments delivered face-to-face, by telephone, and by internet. Ann Behav Med 2010 Aug;40(1):89-98 [FREE Full text] [doi: 10.1007/s12160-010-9203-7] [Medline: 20652466]
23. van DVR, Witting M, Riper H, Kooistra L, Bohlmeijer ET, van GLJ. Blending online therapy into regular face-to-face therapy for depression: content, ratio and preconditions according to patients and therapists using a Delphi study. BMC Psychiatry 2014 Dec 14;14:355 [FREE Full text] [doi: 10.1186/s12888-014-0355-z] [Medline: 25496393]
24. Topooco N, Riper H, Araya R, Berking M, Brunn M, Chevreul K, et al. Attitudes towards digital treatment for depression: A European stakeholder survey. Internet Interventions 2017 Jun;8:1-9. [doi: 10.1016/j.invent.2017.01.001]
25. Ly KH, Topooco N, Cederlund H, Wallin A, Bergström J, Molander O, et al. Smartphone-Supported versus Full Behavioural Activation for Depression: A Randomised Controlled Trial. PLoS One 2015;10(5):e0126559 [FREE Full text] [doi: 10.1371/journal.pone.0126559] [Medline: 26010890]

26. Murray E, Hekler E, Andersson G, Collins L, Doherty A, Hollis C, et al. Evaluating Digital Health Interventions. American Journal of Preventive Medicine 2016 Nov;51(5):843-851. [doi: 10.1016/j.amepre.2016.06.008]
27. Harvey AG, Lee J, Williams J, Hollon SD, Walker MP, Thompson MA, et al. Improving Outcome of Psychosocial Treatments by Enhancing Memory and Learning. Perspect Psychol Sci 2014 Mar;9(2):161-179. [doi: 10.1177/1745691614521781] [Medline: 25544856]

Abbreviations

BAI: Beck Anxiety Inventory
CBT: cognitive behavioral therapy
GAD-7: Generalized Anxiety Disorder Screener 7 items
GEE: generalized estimation equations
ICBT: Internet-delivered cognitive behavioral therapy
MADRS-S: Montgomery-Åsberg Depression Rating Scale self-rating version
OR: odds ratio
PHQ-9: Patient Health Questionnaire 9 items
QOLI: Quality of Life Inventory

Edited by G Eysenbach; submitted 27.05.16; peer-reviewed by J Apolinário-Hagen, J Ruwaard; comments to author 22.09.16; revised version received 12.06.17; accepted 12.07.17; published 24.08.17.

Please cite as:
Månsson KNT, Klintmalm H, Nordqvist R, Andersson G
Conventional Cognitive Behavioral Therapy Facilitated by an Internet-Based Support System: Feasibility Study at a Psychiatric Outpatient Clinic
JMIR Res Protoc 2017;6(8):e158
URL: http://www.researchprotocols.org/2017/8/e158/
doi:10.2196/resprot.6035
PMID:28838884

Corrigenda and Addenda

Correction of: Development of a Web-Based Intervention for Addressing Distress in Caregivers of Patients Receiving Stem Cell Transplants: Formative Evaluation With Qualitative Interviews and Focus Groups

Nicole Amoyal Pensak[1], PhD; Tanisha Joshi[2], PhD; Teresa Simoneau[3], PhD; Kristin Kilbourn[4], PhD; Alaina Carr[4], B A; Jean Kutner[2], PhD, MSPH; Mark L Laudenslager[5], PhD

[1]Hackensack University Medical Center, Department of Research, Cancer Prevention and Control, John Theurer Cancer Center, Hackensack, NJ, United States

[2]University of Colorado Anschutz Medical Campus, Aurora, CO, United States

[3]VA Eastern Colorado Healthcare System, Golden, CO, United States

[4]University of Colorado-Denver, Denver, CO, United States

[5]University of Colorado Anschutz Medical Campus, Department of Psychiatry, Aurora, CO, United States

Corresponding Author:
Nicole Amoyal Pensak, PhD
Hackensack University Medical Center, Department of Research, Cancer Prevention and Control, John Theurer Cancer Center
40 Prospect Avenue
Hackensack, NJ, 07601
United States
Phone: 1 5519964267
Email: Nicole.Pensak@HackensackMeridian.org

Related Article:

Correction of: http://www.researchprotocols.org/2017/6/e120/

(JMIR Res Protoc 2017;6(8):e130) doi:10.2196/resprot.8322

An acknowledgment in the paper by Pensak et al, entitled “Development of a Web-Based Intervention for Addressing Distress in Caregivers of Patients Receiving Stem Cell Transplants: Formative Evaluation With Stakeholder Interviews and Focus Groups” [JMIR Res Protoc 2017;6(6):e120], was inadvertently omitted during typesetting. The following information concerning funding should have appeared: "Acknowledgements: Funded in part by grants from NIA T32AG044296 (JK to NAP) and NCI CA126071 (MLL) and a contract from PCORI CE1304-6208 (MLL)."

In addition, the article's original title has been changed to: “Development of a Web-Based Intervention for Addressing Distress in Caregivers of Patients Receiving Stem Cell Transplants: Formative Evaluation With Qualitative Interviews and Focus Groups.”

The acknowledgment and new title will appear in the online version of the paper on the JMIR website on August 15, 2017, together with the publication of this correction notice. Because this was made after submission to PubMed, the correction notice has been submitted to PubMed. The corrected metadata have also been resubmitted to CrossRef.

Edited by G Eysenbach; submitted 29.06.17; this is a non–peer-reviewed article;accepted 30.06.17; published 15.08.17.

Please cite as:
Pensak NA, Joshi T, Simoneau T, Kilbourn K, Carr A, Kutner J, Laudenslager ML
Correction of: Development of a Web-Based Intervention for Addressing Distress in Caregivers of Patients Receiving Stem Cell Transplants: Formative Evaluation With Qualitative Interviews and Focus Groups
JMIR Res Protoc 2017;6(8):e130
URL: http://www.researchprotocols.org/2017/8/e130/
doi:10.2196/resprot.8322
PMID:30578231

Original Paper

Correction of: Effectiveness of Adaptive E-Learning Environments on Knowledge, Competence, and Behavior in Health Professionals and Students: Protocol for a Systematic Review and Meta-Analysis

Guillaume Fontaine[1], RN, MSc; Sylvie Cossette[1,2], RN, PhD; Marc-André Maheu-Cadotte[1,2], RN, BSc; Tanya Mailhot[1,2,3], RN, PhD; Marie-France Deschênes[2,4], RN, MSc; Gabrielle Mathieu-Dupuis[5], BSc

[1]Montreal Heart Institute Research Center, Montreal, QC, Canada
[2]Faculty of Nursing, Université de Montréal, Montreal, QC, Canada
[3]Faculty of Medicine, Université de Montréal, Montreal, QC, Canada
[4]Center for Innovation in Nursing Education, Université de Montréal, Montreal, QC, Canada
[5]School of Librarianship and Information Science, Université de Montréal, Montreal, QC, Canada

Corresponding Author:
Guillaume Fontaine, RN, MSc
Montreal Heart Institute Research Center
5000 Belanger Street
Room S-2490
Montreal, QC, H1T 1C8
Canada
Phone: 1 514 376 3330 ext 3069
Fax: 1 514 593 7441
Email: guillaume.fontaine@umontreal.ca

Related Article:

Correction of: http://www.researchprotocols.org/2017/7/e128/

(JMIR Res Protoc 2017;6(8):e131) doi:10.2196/resprot.8377

In the paper by Fontaine et al, entitled "Effectiveness of Adaptive E-Learning Environments on Knowledge, Competence, and Behavior in Health Professionals and Students: Protocol for a Systematic Review and Meta-Analysis" [JMIR Res Protoc 2017;6(7):e128], a passage was inadvertently omitted during copyediting. Under the subheading, *Secondary Outcome Measures*, the following should have been included:

"Primary studies reporting an objective measure of users' competence (eg, behavior change counseling competence scores) or a subjective measure of users' competence (eg, self-reported skills) will be considered for inclusion."

The paragraph given above should have preceded the one below:

"Primary studies reporting an objective measure of users' behavior (eg, clinical interventions reported in patients' medical file, number of tests ordered) and a subjective measure of users' clinical behavior (eg, self-reported performance of clinical interventions) will be considered for inclusion."

In addition, the trial registration number and associated url links were left out of the abstract. This information should have appeared as follows: *Trial Registration:* PROSPERO International Prospective Register of Systematic Reviews: CRD42017065585; https://www.crd.york.ac.uk/PROSPERO/display_record.asp? ID=CRD42017065585 (Archived by WebCite® athttp://www.webcitation.org/6rXGdDwf4).

This omission has been corrected in the online version of the paper on the JMIR website on August 15, 2017, together with the publication of this correction notice. Because this was made after submission to PubMed, the correction notice has been submitted to PubMed. The corrected metadata have also been resubmitted to CrossRef.

Edited by G Eysenbach; submitted 07.07.17; this is a non–peer-reviewed article;accepted 07.07.17; published 15.08.17.

Please cite as:
Fontaine G, Cossette S, Maheu-Cadotte MA, Mailhot T, Deschênes MF, Mathieu-Dupuis G
Correction of: Effectiveness of Adaptive E-Learning Environments on Knowledge, Competence, and Behavior in Health Professionals and Students: Protocol for a Systematic Review and Meta-Analysis
JMIR Res Protoc 2017;6(8):e131
URL: http://www.researchprotocols.org/2017/8/e131/
doi:10.2196/resprot.8377
PMID:30578209

Protocol

Beliefs, Knowledge, Implementation, and Integration of Evidence-Based Practice Among Primary Health Care Providers: Protocol for a Scoping Review

Filipa Pereira[1], BSN, MNSc; Mireille Salvi[1]; Henk Verloo[1], RN, PhD

School of Health Sciences, HES-SO Valais/Wallis, Department of Nursing Sciences, University of Applied Sciences and Arts of Western Switzerland, Sion, Switzerland

Corresponding Author:
Filipa Pereira, BSN, MNSc
School of Health Sciences, HES-SO Valais/Wallis
Department of Nursing Sciences
University of Applied Sciences and Arts of Western Switzerland
5 Chemin de l'Agasse
Sion, CH-1950
Switzerland
Phone: 41 27 606 84 00
Fax: 41 27 606 84 01
Email: filipa.pereira@hevs.ch

Abstract

Background: The adoption of evidence-based practice (EBP) is promoted because it is widely recognized for improving the quality and safety of health care for patients, and reducing avoidable costs. Providers of primary care face numerous challenges to ensuring the effectiveness of their daily practices. Primary health care is defined as: the entry level into a health care services system, providing a first point of contact for all new needs and problems; patient-focused (not disease-oriented) care over time; care for all but the most uncommon or unusual conditions; and coordination or integration of care, regardless of where or by whom that care is delivered. Primary health care is the principal means by which to approach the main goal of any health care services system: optimization of health status.

Objective: This review aims to scope publications examining beliefs, knowledge, implementation, and integration of EBPs among primary health care providers (HCPs).

Methods: We will conduct a systematic scoping review of published articles in the following electronic databases, from their start dates until March 31, 2017: Medical Literature Analysis and Retrieval System Online (MEDLINE) via PubMed (from 1946), Embase (from 1947), Cumulative Index to Nursing and Allied Health Literature (CINAHL; from 1937), the Cochrane Central Register of Controlled Trials (CENTRAL; from 1992), PsycINFO (from 1806), Web of Science (from 1900), Joanna Briggs Institute (JBI) database (from 1998), Database of Abstracts of Reviews of Effects (DARE; from 1996), Trip medical database (from 1997), and relevant professional scientific journals (from their start dates). We will use the predefined search terms of, "evidence-based practice" and, "primary health care" combined with other terms, such as, "beliefs", "knowledge", "implementation", and "integration". We will also conduct a hand search of the bibliographies of all relevant articles and a search for unpublished studies using Google Scholar, ProQuest, Mednar, and WorldCat. We will consider publications in English, French, Spanish, and Portuguese.

Results: The electronic database searches were completed in April 2017. Retrieved articles are currently being screened, and the entire study is expected to be completed by November 2017.

Conclusions: This systematic scoping review will provide a greater understanding of the beliefs, knowledge, implementation, and integration of EBPs among primary HCPs. The findings will inform clinical practice and help to draw a global picture of the EBP research topics that are relevant to primary care providers.

(JMIR Res Protoc 2017;6(8):e148) doi:10.2196/resprot.7727

KEYWORDS

evidence-based practice; primary healthcare; primary healthcare providers; beliefs; knowledge; implementation

Introduction

Evidence-based practice (EBP) is an emerging, breakthrough approach among health care providers (HCPs) [1,2]. EBP has its origins in evidence-based medicine and is defined as, "the conscientious and judicious use of current best evidence in making decisions about the care of individual patients" [3]. HCPs are expected to use EBPs as a standard approach to daily practice [4,5], thereby integrating research, patient preferences, clinical expertise, and innovative technologies [6-8]. However, the implementation of EBPs remains a controversial process [9,10], and not all HCPs are convinced that it improves the quality of care [11,12]. Implementing EBPs is challenging, especially in primary health care settings [13,14].

Primary health care is defined as: the entry level into a health care services system, providing a first point of contact for all new needs and problems; patient-focused (not disease-oriented) care over time; care for all but the most uncommon or unusual conditions; and coordination or integration of care, regardless of where or by whom that care is delivered. Primary health care is the primary means by which to approach the main goal of any health care services system: optimization of health status [15]. Health care provided by primary HCPs includes health promotion, prevention and diagnosis, detection, intervention, treatment, and case and care management [16,17]. Furthermore, primary HCPs provide first-line health care services to home-dwelling adult patients and long-term nursing home patients. Nevertheless, in some acute health situations, home-dwelling individuals will need to be referred to medical specialists or acute hospital services for additional health care advice [15,17]. Primary HCPs play a crucial decision-making role, which strengthens communication and collaboration between community HCPs and specialized HCPs to provide the best available overall health care to community-dwelling individuals [14]. Primary HCPs include general practitioners, community health care nurses and nurse practitioners, midwives and allied health care professionals (occupational therapists, physical therapists, speech and language therapists, podiatrists, dieticians, psychologists, social workers, and radiological and medical imaging technologists), pharmacists, and dentists [18].

Although every HCP is generally considered accountable for providing the best available evidence-based health care [19,20], recent research has concluded that only a small percentage of HCPs consistently do so [21-25]. The implementation rate of EBPs among HCPs in hospital institutions has been largely documented [26,27] and multiple barriers have been reported [19,22,28,29]. These barriers include time constraints, a lack of personal motivation and negative attitudes, professional resistance to research, and inadequate knowledge of (and skills needed for) EBPs among clinicians [30-32]. Additionally, several authors have documented administrative and organizational problems in the workplace, a lack of mentors for EBPs, inadequate resources at the point of care, a gap between theory and practice, a lack of any meaningful transition between training courses on EBP and the clinical reality, and an absence or lack of basic education on the subject [21,25,33]. Finally, different authors have highlighted that HCPs' beliefs about EBPs are associated with their capacity to implement such practices [22,34,35].

Over the last two decades, the use of EBPs in health care has been documented in exploratory and observational studies in different settings. However, there have been no wide-ranging overviews or comparisons examining the beliefs, knowledge, implementation, and integration of EBPs in primary health care settings, or among primary HCPs [36,37].

This scoping review aims to explore the different studies examining the beliefs, knowledge, implementation, and integration of EBPs in primary health care settings and among primary care HCPs. The following research questions will be explored:

1. What is the extent of the research exploring beliefs, knowledge, implementation, and integration of EBPs among primary HCPs?
2. What is the nature of the research exploring beliefs, knowledge, implementation, and integration of EBPs among primary HCPs?
3. How do the extent and nature of the research exploring beliefs, knowledge, implementation, and integration of EBPs vary across primary HCPs?

Methods

We will use the scoping review methodological framework conceived by Arksey and O'Malley [38], with the refinements described by Levac et al [39] and Colquhoun et al [40], and consider the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement [41].

This multistage model involves: (1) identifying the research questions (listed above); (2) identifying relevant studies (search methods used); (3) selecting studies; (4) charting the data; (5) collating, summarizing, and reporting the results; and, if necessary, (6) consulting with key stakeholders. The refinements to the original framework include: establishing clear research questions, purposes, and expected outcomes of the scoping review; assembling a team with content and methodological expertise; searching the literature using an iterative process involving inclusion and exclusion criteria; using at least two reviewers to independently review abstracts and full-text papers, with a consensus procedure in cases of disagreement; developing a data extraction form onto which two researchers can upload the data independently; conducting a quality assessment of included papers; and performing an analysis that includes a descriptive quantitative summary of papers, as well as a qualitative thematic analysis.

Eligibility Criteria

Searches in scoping reviews are recommended to be as comprehensive as possible, in order to identify every possible study that is relevant to the field [39,42]. Original studies eligible for inclusion encompass descriptive studies, cohort studies, interventional studies, qualitative studies, mixed-design studies, and reports studying EBPs among primary HCPs. Studies should provide a measure or level of: (1) belief, (2)

knowledge, (3) implementation, or (4) integration of EBPs, as defined by the authors.

Studies should be conducted among primary HCPs in primary health care settings, and among HCPs carrying out parallel activities in primary, secondary, and tertiary health care settings. Studies should also include regional and national surveys among general practitioners, community health care nurses and nurse practitioners, midwives, occupational therapists, physical therapists, speech and language therapists, podiatrists, dieticians, psychologists, social workers, radiological and medical imaging technologists, dentists, and pharmacists. Grey literature containing information on beliefs, knowledge, implementation, and integration of EBPs is eligible. Studies including expert opinions and editorials will be excluded from the original set of studies collected.

Information Sources and Search Strategy

We will conduct a systematic scoping search of published articles in the following electronic databases, from their start dates until March 31, 2017: Medical Literature Analysis and Retrieval System Online (MEDLINE) via PubMed (from 1946), Embase (from 1947), Cumulative Index to Nursing and Allied Health Literature (CINAHL; from 1937), the Cochrane Central Register of Controlled Trials (CENTRAL; from 1992), PsycINFO (from 1806), Web of Science (from 1900), Joanna Briggs Institute (JBI) database (from 1998), Database of Abstracts of Reviews of Effects (DARE; from 1996), Trip medical database (from 1997) and relevant professional scientific journals for the primary HCPs mentioned (from their start dates). We will use predefined search terms and Medical Subject Headings (MeSH) terms such as, “evidence-based practice”, “evidence-based nursing”, “evidence-based approach”, “best practices”, “beliefs”, “knowledge”, “implementation”, “integration”, and “health personnel”. The “health personnel” MeSH term’s tree will be expanded using the MeSH terms and keywords “health care” and “allied health care professionals”. In addition to searching electronic databases, we will conduct a hand search of the bibliographies of all relevant articles and a search of unpublished studies using Google Scholar, ProQuest, Mednar, the WorldCat and Health Technology Assessment (Canadian Search Interface) databases, and OpenGrey. We will consider publications in English, French, Spanish, and Portuguese. If studies are identified in languages other than those mastered by the research team, we will contact their authors to complete the data extraction and quality assessment form. A draft MEDLINE search strategy is included in Multimedia Appendix 1. When the MEDLINE strategy has been finalized, subject headings and syntax will be adapted for the other databases. The search strategy’s results will be reported following the PRISMA statement guidelines (Multimedia Appendix 2) [43].

Bibliographic Management

Review Manager software version 5.3 and Endnote 8.0 will be used for collecting and analyzing the bibliographic references retrieved using the search strategy [44].

Article Selection Process

We will include studies by using a two-step process with two independent reviewers for each step. First, each abstract initially selected will be evaluated. Then, each potentially relevant full article will be retrieved for consideration of inclusion. Second, data will be selected and extracted. In cases of a disagreement between the two reviewers, the other team members will contribute to the decision. Disagreements will be resolved through discussion among the team members or, if needed, a consensus will be reached following a further discussion with the authors. Article selection will be based on the methodological framework for scoping studies recommended by Arksey and O'Malley and Levac et al [38,39] and the systematic review by Leung et al [45].

Data Collection Process

We will use standardized data collection forms (Microsoft Excel sheets for the data on the studies and interventions, as well as a quality assessment of the studies included) developed by the research team. Data extraction will be conducted independently by two separate reviewers using a specially designed, standardized data extraction form based on a relevant, previously published extraction form [46]. Discrepancies will be resolved through discussion and consultation with the co-authors.

Data Items

The following information will be extracted from each included study and put into an appropriate usable form: (1) study author, year of publication, and country in which the study was conducted; (2) study characteristics (including research questions, study setting and design, instruments used, duration of follow-up, and sample size); (3) participants’ characteristics (including sex, age, professional activity and experience, level of education, and setting); and (4) types of outcome measures [47].

Outcomes and Prioritization

The primary outcomes will be the nature, number, and comparison of studies examining beliefs, knowledge, implementation, and/or integration of EBPs among primary HCPs. These outcomes will allow us to generate an overview of the existing implementation and integration of EBPs in primary health care.

We hypothesize that significant differences will exist in the designs and instruments of EBPs, both in different primary health care settings and among different primary HCPs. Secondary outcomes will focus on: the beliefs, knowledge, implementation, and integration measures/levels of EBPs; the theoretical frameworks and instruments used for assessment; and the level of evidence provided in the included studies [48]. Overall, the outcomes of this scoping review will provide useful suggestions and recommendations for identifying possible gaps in the research [38].

Data Synthesis

We will summarize the results using a descriptive (noninterpretive) narrative synthesis [38,39,42] and content analysis [49]. All data on the beliefs, knowledge, implementation, and integration measures/levels for EBPs will

be summarized in a table. By summarizing and critically appraising all studies, we will be able to identify gaps in the current evidence and avenues for future research.

Confidence in Cumulative Evidence and Risk of Bias

As recommended by Daudt et al [42], we will assess the quality of the eligible observational and mixed-method studies by using the appropriate recommended tools: Grading of Recommendations Assessment, Development, and Evaluation; and the Mixed-Methods Appraisal Tool [47,50]. Qualitative studies will be assessed using the Letts Qualitative Review Form [51]. The risk of bias in the retrieved studies will be assessed using the Cochrane Collaboration tool [44]. We will not, however, exclude studies based on a quality assessment, because we wish to provide a comprehensive overview of the available evidence and its extent. The quality of evidence will be assessed using the following items: risk of bias, inconsistency, imprecision, indirectness, publication bias, and the confidence effect [44]. Using these tools, each of the following domains will be rated as: (1) very low quality, (2) low quality, (3) moderate quality, or (4) high quality [52]. Any disagreements regarding the quality assessment will be resolved through discussion among the team members.

Ethics and Dissemination

All data in this project will be gathered through searches of literature databases, and recommendations and guidelines available online. No information on individuals will be collected within the framework of this project, thus approval from a research ethics committee is not required.

Results

The electronic database searches were completed in July 2017. Retrieved articles are currently being screened and the entire study is expected to be completed by November 2017.

Discussion

Providing the best available, safe, high-quality health care is the gold standard objective in all health care settings. To the best of our knowledge, there exists no global overview of the beliefs, knowledge, implementation, and integration of EBPs among primary HCPs and institutions. This documentary research project will provide a picture of the state of the art of research in this domain, and reveal to what extent EBPs are implemented and integrated. This review will, therefore, provide valuable information to practitioners, policy makers, and other stakeholders.

Strengths and Limitations

This systematic scoping research protocol's strengths are: (1) obtaining a broad overview of the studies dealing with beliefs, knowledge, implementation, and integration of EBPs among primary HCPs; (2) the use of an appropriate search strategy designed in collaboration with two specialist librarians who are experienced in conducting such reviews; and (3) the inclusion criteria, which impose no restrictions on time period or geographic location.

Nevertheless, this protocol does include some limitations which may introduce bias: (1) the exclusion of articles written in languages other than English, French, Spanish, or Portuguese; and (2) the personal judgements of the reviewers' study assessments.

Acknowledgments

The authors wish to thank Pauline Melly, librarian, for the revision of the PubMed equation. This scoping review is funded by the Department of Nursing Sciences at the University of Applied Sciences and Arts Western Switzerland (funding reference number: Sagex 71796). This funding will support data collection and analyses. The funder is not involved in protocol design, the plan of analyses, or the interpretation or publication of the scoping review's results.

Authors' Contributions

FP is the guarantor. FP, MS, and HV drafted the protocol manuscript. All authors contributed to the development of the selection criteria, data extraction and analyses, and the search strategy. FP, MS, and HV provided expertise on EBPs. All authors approved the final protocol manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Draft MEDLINE - PubMed search.

[PDF File (Adobe PDF File), 21KB - resprot_v6i8e148_app1.pdf]

Multimedia Appendix 2

Flow diagram based on the PRISMA-p guidelines.

[PDF File (Adobe PDF File), 38KB - resprot_v6i8e148_app2.pdf]

References

1. Heiwe S, Kajermo KN, Tyni-Lenné R, Guidetti S, Samuelsson M, Andersson I, et al. Evidence-based practice: attitudes, knowledge and behaviour among allied health care professionals. Int J Qual Health Care 2011 Apr;23(2):198-209. [doi: 10.1093/intqhc/mzq083] [Medline: 21242158]
2. Melnyk BM, Gallagher-Ford L. Implementing the new essential evidence-based practice competencies in real-world clinical and academic settings: moving from evidence to action in improving healthcare quality and patient outcomes. Worldviews Evid Based Nurs 2015 Apr;12(2):67-69. [doi: 10.1111/wvn.12089] [Medline: 25810000]
3. Sackett DL, Rosenberg WM, Gray JA, 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] [Medline: 8555924]
4. Institute of Medicine. The Future of Nursing: Leading Change, Advancing Health. Washington, DC: National Academies Press; 2011.
5. Hughes R. Patient safety and quality: an evidence-based handbook for nurses. AHRQ Publication No. 08-0043 2008 Mar:1403 Rockville, MD. [Medline: 21328752]
6. Goulet C, Lampron A, Morin D, Héon M. [Evidence-based practice. Part 1: origins, definitions, critiques, obstacles, advantages and impact]. Rech Soins Infirm 2004 Mar(76):12-18. [Medline: 15085564]
7. Melnyk B. Evidence-Based Practice in Nursing & Healthcare: A Guide to Best Practice. Philadelphia, PA: Wolters Kluwer Health; 2014.
8. Goulet C, Lampron A, Morin D, Héon M. [Evidence-based practice: Part 2: steps in the process]. Rech Soins Infirm 2004 Mar(76):19-29. [Medline: 15085565]
9. Levin RF, Feldman HR. The EBP controversy: misconception, misunderstanding, or myth. Res Theory Nurs Pract 2006;20(3):183-186. [Medline: 16986352]
10. Zeitz K, McCutcheon H. Evidence-based practice: to be or not to be, this is the question!. Int J Nurs Pract 2003 Oct;9(5):272-279. [Medline: 14531848]
11. McEvoy MP, Williams MT, Olds TS, Lewis LK, Petkov J. Evidence-based practice profiles of physiotherapists transitioning into the workforce: a study of two cohorts. BMC Med Educ 2011 Nov 29;11:100 [FREE Full text] [doi: 10.1186/1472-6920-11-100] [Medline: 22126299]
12. Wilkinson SA, Hough J, Hinchliffe F. An Evidence-based approach to influencing evidence-based practice in allied health. J Allied Health 2016;45(1):41-48. [Medline: 26937881]
13. Taylor AC, Bond GR. Employment specialist competencies as predictors of employment outcomes. Community Ment Health J 2014 Jan;50(1):31-40. [doi: 10.1007/s10597-012-9554-6] [Medline: 23054156]
14. Brownstein JN, Hirsch GR, Rosenthal EL, Rush CH. Community health workers “101” for primary care providers and other stakeholders in health care systems. J Ambul Care Manage 2011;34(3):210-220. [doi: 10.1097/JAC.0b013e31821c645d] [Medline: 21673520]
15. Van Lerberghe W. The world health report 2008: primary health care: now more than ever.: World Health Organization; 2008. URL: http://www.who.int/whr/2008/whr08_en.pdf [accessed 2017-06-18] [WebCite Cache ID 6rJRKBbNx]
16. Witmer A, Seifer SD, Finocchio L, Leslie J, O'Neil EH. Community health workers: integral members of the health care work force. Am J Public Health 1995 Aug;85(8 Pt 1):1055-1058. [Medline: 7625495]
17. Adashi EY, Geiger HJ, Fine MD. Health care reform and primary care--the growing importance of the community health center. N Engl J Med 2010 Jun 03;362(22):2047-2050. [doi: 10.1056/NEJMp1003729] [Medline: 20427777]
18. Swider SM. Outcome effectiveness of community health workers: an integrative literature review. Public Health Nurs 2002;19(1):11-20. [Medline: 11841678]
19. Swinkels A, Albarran JW, Means RI, Mitchell T, Stewart MC. Evidence-based practice in health and social care: where are we now? J Interprof Care 2002 Nov;16(4):335-347. [doi: 10.1080/1356182021000008265] [Medline: 12487842]
20. Ubbink DT, Vermeulen H, Knops AM, Legemate DA, Oude RK, Heineman MJ, et al. Implementation of evidence-based practice: outside the box, throughout the hospital. Neth J Med 2011 Feb;69(2):87-94 [FREE Full text] [Medline: 21411849]
21. Scurlock-Evans L, Upton P, Upton D. Evidence-based practice in physiotherapy: a systematic review of barriers, enablers and interventions. Physiotherapy 2014 Sep;100(3):208-219. [doi: 10.1016/j.physio.2014.03.001] [Medline: 24780633]
22. Estrada N. Exploring perceptions of a learning organization by RNs and relationship to EBP beliefs and implementation in the acute care setting. Worldviews Evid Based Nurs 2009;6(4):200-209. [doi: 10.1111/j.1741-6787.2009.00161.x] [Medline: 19686224]
23. Glegg SM, Holsti L. Measures of knowledge and skills for evidence-based practice: a systematic review. Can J Occup Ther 2010 Oct;77(4):219-232. [doi: 10.2182/cjot.2010.77.4.4] [Medline: 21090063]
24. Friesen-Storms JH, Moser A, van der Loo S, Beurskens AJ, Bours GJ. Systematic implementation of evidence-based practice in a clinical nursing setting: a participatory action research project. J Clin Nurs 2015 Jan;24(1-2):57-68. [doi: 10.1111/jocn.12697] [Medline: 25258116]

25. Verloo H, Desmedt M, Morin D. Beliefs and implementation of evidence-based practice among nurses and allied healthcare providers in the Valais hospital, Switzerland. J Eval Clin Pract 2017 Feb;23(1):139-148. [doi: 10.1111/jep.12653] [Medline: 27687154]
26. Melnyk BM, Fineout-Overholt E, Feinstein NF, Sadler LS, Green-Hernandez C. Nurse practitioner educators' perceived knowledge, beliefs, and teaching strategies regarding evidence-based practice: implications for accelerating the integration of evidence-based practice into graduate programs. J Prof Nurs 2008;24(1):7-13. [doi: 10.1016/j.profnurs.2007.06.023] [Medline: 18206837]
27. Wallen GR, Mitchell SA, Melnyk B, Fineout-Overholt E, Miller-Davis C, Yates J, et al. Implementing evidence-based practice: effectiveness of a structured multifaceted mentorship programme. J Adv Nurs 2010 Dec;66(12):2761-2771 [FREE Full text] [doi: 10.1111/j.1365-2648.2010.05442.x] [Medline: 20825512]
28. Earle-Foley V. Evidence-based practice: issues, paradigms, and future pathways. Nurs Forum 2011;46(1):38-44. [doi: 10.1111/j.1744-6198.2010.00205.x] [Medline: 21306394]
29. Fernández MJ. Evidence for practice--growing slowly but surely. Australas Emerg Nurs J 2014 Feb;17(1):1. [doi: 10.1016/j.aenj.2013.12.001] [Medline: 24507175]
30. Wittmann-Price R, Celia L, Dunn R. Successful implementation of evidence-based nursing practice: the indispensable role of staff development. J Nurses Prof Dev 2013;29(4):202-204. [doi: 10.1097/NND.0b013e31829b2212] [Medline: 23877292]
31. Breimaier HE, Halfens RJG, Lohrmann C. Nurses' wishes, knowledge, attitudes and perceived barriers on implementing research findings into practice among graduate nurses in Austria. J Clin Nurs 2011 Jun;20(11-12):1744-1756. [doi: 10.1111/j.1365-2702.2010.03491.x] [Medline: 21362075]
32. Majid S, Foo S, Luyt B, Zhang X, Theng Y, Chang Y, et al. Adopting evidence-based practice in clinical decision making: nurses' perceptions, knowledge, and barriers. J Med Libr Assoc 2011 Jul;99(3):229-236 [FREE Full text] [doi: 10.3163/1536-5050.99.3.010] [Medline: 21753915]
33. Upton D, Stephens D, Williams B, Scurlock-Evans L. Occupational therapists' attitudes, knowledge, and implementation of evidence-based practice: a systematic review of published research. Br J of Occup Ther 2014 Jan 15;77(1):24-38. [doi: 10.4276/030802214x13887685335544]
34. Melnyk BM, Fineout-Overholt E, Giggleman M, Cruz R. Correlates among cognitive beliefs, EBP implementation, organizational culture, cohesion and job satisfaction in evidence-based practice mentors from a community hospital system. Nurs Outlook 2010;58(6):301-308. [doi: 10.1016/j.outlook.2010.06.002] [Medline: 21074647]
35. Thorsteinsson HS, Sveinsdóttir H. Readiness for and predictors of evidence-based practice of acute-care nurses: a cross-sectional postal survey. Scand J Caring Sci 2014 Sep;28(3):572-581. [doi: 10.1111/scs.12083] [Medline: 24111971]
36. González-Torrente S, Pericas-Beltrán J, Bennasar-Veny M, Adrover-Barceló R, Morales-Asencio JM, De Pedro-Gomez J. Perception of evidence-based practice and the professional environment of primary health care nurses in the Spanish context: a cross-sectional study. BMC Health Serv Res 2012 Jul 31;12:227 [FREE Full text] [doi: 10.1186/1472-6963-12-227] [Medline: 22849698]
37. Al Omari M, Khader Y, Jadallah K, Dauod AS, Al-Shdifat AAK. Awareness, attitude and practice of evidence-based medicine among primary health care doctors in Jordan. J Eval Clin Pract 2009 Dec;15(6):1131-1136. [doi: 10.1111/j.1365-2753.2009.01223.x] [Medline: 20367716]
38. 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]
39. 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]
40. Colquhoun HL, Letts LJ, Law MC, MacDermid JC, Missiuna CA. A scoping review of the use of theory in studies of knowledge translation. Can J Occup Ther 2010 Dec;77(5):270-279. [doi: 10.2182/cjot.2010.77.5.3] [Medline: 21268509]
41. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015 Jan 02;349:g7647 [FREE Full text] [Medline: 25555855]
42. Daudt HML, van Mossel C, Scott SJ. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Med Res Methodol 2013 Mar 23;13:48 [FREE Full text] [doi: 10.1186/1471-2288-13-48] [Medline: 23522333]
43. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, 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):e1-34 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.06.006] [Medline: 19631507]
44. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions. Chichester, England: Wiley-Blackwell; 2008.
45. Leung K, Trevena L, Waters D. Systematic review of instruments for measuring nurses' knowledge, skills and attitudes for evidence-based practice. J Adv Nurs 2014 Oct;70(10):2181-2195. [doi: 10.1111/jan.12454] [Medline: 24866084]
46. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. J Health Serv Res Policy 2005 Jan;10(1):45-53. [Medline: 15667704]

47. Souto RQ, Khanassov V, Hong QN, Bush PL, Vedel I, Pluye P. Systematic mixed studies reviews: updating results on the reliability and efficiency of the Mixed Methods Appraisal Tool. Int J Nurs Stud 2015 Jan;52(1):500-501. [doi: 10.1016/j.ijnurstu.2014.08.010] [Medline: 25241931]
48. Aguayo-Albasini JL, Flores-Pastor B, Soria-Aledo V. [GRADE system: classification of quality of evidence and strength of recommendation]. Cir Esp 2014 Feb;92(2):82-88 [FREE Full text] [doi: 10.1016/j.ciresp.2013.08.002] [Medline: 24361098]
49. 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]
50. Pace R, Pluye P, Bartlett G, Macaulay AC, Salsberg J, Jagosh J, et al. Testing the reliability and efficiency of the pilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review. Int J Nurs Stud 2012 Jan;49(1):47-53. [doi: 10.1016/j.ijnurstu.2011.07.002] [Medline: 21835406]
51. Letts L, Wilkins S, Law M, Stewart D, Bosch J, Westmorland M. Critical Review Form - Qualitative Studies (Version 2.0). 2007. URL: https://www.scienceopen.com/document?vid=3d7a0f0e-1fe0-472d-b901-30636dd74f0a [accessed 2017-06-18] [WebCite Cache ID 6rJV9nVWG]
52. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol 2011 Dec;64(12):1311-1316. [doi: 10.1016/j.jclinepi.2011.06.004] [Medline: 21802902]

Abbreviations

CENTRAL: Cochrane Central Register of Controlled Trials
CINAHL: Cumulative Index to Nursing and Allied Health Literature
DARE: Database of Abstracts of Reviews of Effects
EBP: evidence-based practice
HCP: health care provider
JBI: Joanna Briggs Institute
MEDLINE: Medical Literature Analysis and Retrieval System Online
MeSH: Medical Subject Headings
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols

Edited by G Eysenbach; submitted 21.03.17; peer-reviewed by G Malarvizhi, D Tao; comments to author 11.05.17; revised version received 02.06.17; accepted 20.06.17; published 01.08.17.

Please cite as:
Pereira F, Salvi M, Verloo H
Beliefs, Knowledge, Implementation, and Integration of Evidence-Based Practice Among Primary Health Care Providers: Protocol for a Scoping Review
JMIR Res Protoc 2017;6(8):e148
URL: http://www.researchprotocols.org/2017/8/e148/
doi:10.2196/resprot.7727
PMID:28765102

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org