JMIR Formative Research

Contents

Original Papers

Original Paper

Feasibility and Acceptability of an Electronic Health HIV Prevention Toolkit Intervention With Concordant HIV-Negative, Same-Sex Male Couples on Sexual Agreement Outcomes: Pilot Randomized Controlled Trial

Jason William Mitchell[1], MPH, PhD; Ji-Young Lee[2], MPH; Yanyan Wu[1], PhD; Patrick S Sullivan[3], DVM, PhD; Rob Stephenson[4], PhD

[1]Office of Public Health Studies, Myron B Thompson School of Social Work, University of Hawaii, Honolulu, HI, United States

[2]Department of Public Health Sciences, University of Miami, Miami, FL, United States

[3]Rollins School of Public Health, Emory University, Atlanta, GA, United States

[4]The Center for Sexuality and Health Disparities, Department of Systems, Population and Leadership, School of Nursing, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Jason William Mitchell, MPH, PhD
Office of Public Health Studies
Myron B Thompson School of Social Work
University of Hawaii
1960 East-West Rd, Suite T103
Honolulu, HI, 96822
United States
Phone: 1 8082196729
Email: jasonmit@hawaii.edu

Abstract

Background: There is a need to develop innovative and accessible dyadic interventions that provide male couples with the behavioral skills to manage the risk of HIV transmission within their relationship.

Objective: We conducted a pilot randomized controlled trial (RCT) to assess the feasibility and acceptability of the electronic health (eHealth) HIV prevention toolkit intervention to encourage seroconcordant negative male couples in the United States to establish and adhere to a sexual agreement (SA).

Methods: Eligible, consented couples were randomly assigned to the intervention or education control and followed up for 6 months, with assessments occurring every 3 months after baseline. Acceptability items were assessed at both follow-up assessments. Descriptive and comparative statistics summarized cohort characteristics, relationship dynamics, and SA outcomes for the entire cohort and by trial arm. To examine the association between couples' relationship dynamics and their establishment of an SA over time and by trial arm, multilevel logistic regression analyses were performed with a random intercept to account for correlations of repeated measurements of relationship dynamics at months 3 and 6; the odds ratio (OR) of establishment of an SA and the corresponding 95% confidence interval were then reported.

Results: Overall, 7959 individuals initiated screening. Reasons for individual ineligibility varied. An electronic algorithm was used to assess couple-level eligibility, which identified 1080 ineligible and 266 eligible dyads. Eligible couples (n=149) were enrolled in the pilot RCT: 68 received the intervention and 81 received the education control. Retention was 71.5% (213/298 partnered men) over the 6 months. Participants reported high acceptability of the intervention along with some areas for improvement. A significantly higher proportion of couples who received the intervention established an SA at 6 months compared with those who received the education control (32/43, 74% vs 27/50, 54%; $P=.05$). The OR of establishing an SA for couples in the intervention versus those in the control condition was greater than 2 when controlling for a number of different relationship dynamics. In addition, the odds of establishing an SA increased by 88% to 322% for each unit increase in a variety of averaged relationship dynamic scores; the opposite result was found for dynamics of stigma. Differences between trial arms for SA type and adherence were nonsignificant at each assessment. However, changes in these 2 SA aspects were noted over time. The average

number of items couples included in their SA was 18, and about one-fourth to one-third of couples included HIV prevention items.

Conclusions: The findings demonstrate strong evidence for the acceptability and feasibility of the eHealth toolkit as a brief, stand-alone, couples-based HIV prevention intervention. These findings support the need to update the toolkit and evaluate it in a larger clinical trial powered for efficacy.

Trial Registration: ClinicalTrials.gov NCT02494817; http://clinicaltrials.gov/ct2/show/NCT02494817

Introduction

Background

National estimates indicate that between one-third and two-thirds of HIV infections among gay, bisexual, and other men who have sex with men (GBMSM) occur within primary relationships (ie, male couples) [1,2]. In response to these estimates, a growing interest in couples-based approaches to HIV prevention [3-10] has emerged to investigate how relationship dynamics may affect male couples' risk for HIV and other sexually transmitted infections (STIs) and the development of interventions for this population. There is a need to develop innovative dyadic interventions that provide couples with the behavioral skills to manage the risk of HIV transmission within their relationship.

Couples Interdependence Theory

Couples interdependence theory (CIT) is a useful health behavior change theory to understand and examine the process in which relationship dynamics (ie, interaction between primary partners) positively and negatively impact male couples' decisions and behaviors relative to their risk for HIV and other STIs [11,12]. CIT describes one potential process of how relationship partners influence, initiate, and maintain behaviors that impact one another's health (ie, interdependence) [11,12]. Relative to HIV/STI prevention, this theoretical framework takes into consideration predisposing factors of the couple, which includes their relationship functioning (eg, commitment, satisfaction, and trust), communication style, perceptions of HIV/STIs as a health threat, and preferences for outcomes associated with the health threat (eg, condom use and testing).

In CIT, these predisposing factors are posited to affect couples' transformation of motivation and communal coping, 2 other key features of CIT. Transformation of motivation refers to the couple's cognitive interpretation and emotional response to the health threat as being meaningful (ie, important) to their relationship. In other words, partners move from a primarily individual-focused motivation to one that is more prorelationship and health enhancing (ie, how both partners as a couple benefit instead of only one) [11,12]. Transformation of motivation also lends itself to the couple, creating joint goals for long-term relationship functioning, and each partner's willingness to accommodate for the relationship is a function of the dynamics present in the relationship [13,14].

Another key component to CIT is communal coping, which refers to partners having a shared assessment of HIV/STIs as a health threat, a vision of shared action about managing and reducing their risk for HIV/STIs (via behaviors) and engaging in related HIV/STI prevention behaviors that are beneficial to them as a couple [11,12]. Couples' coping strategies for HIV/STI prevention are largely determined by the degree that both partners appraise HIV/STIs from an individual standpoint to one as a collective team (ie, transformation of motivation), such that their shared emotional and cognitive responses lead to a greater likelihood of them making a joint effort, partaking in planning and decision making and communicating about how best to reduce their risk for HIV and other STIs [15,16]. CIT provides a useful theoretical framework to examine how couples' dynamics in general and changes in their dynamics (eg, predisposing factors, transformation of motivation, and communal coping) may lead relationship partners to working together to engage in and achieve their joint health goals as it applies to HIV/STI prevention.

Sexual Agreements

Sexual agreements (SAs) are one dynamic of male couples' relationships that have implications for HIV/STI prevention, as supported by a number of investigations identified in a recent scoping view [17]. An SA is formed when partners have explicit conversations with decision making that leads them to having a mutual understanding about which sexual and other relational behaviors, they want to occur with each other (ie, in their relationship) and if applicable, with anyone else (eg, casual sex partners) [18,19]. To date, much research has been conducted about male couples' SAs, including circumstances and reasons for forming an agreement [18,20-23], investment in one [24-27], and adherence rates and disclosure, and reasons when an agreement is broken [17,28]. SAs are common among male couples [18,20,21,25,29], vary by type [18,19,29,30] and are dynamic, such that changes in composition or type may occur over time [31]. Types of SAs and the composition of these types come in many forms, yet they generally fall into 3 broad categories of closed, open with guidelines, and open without guidelines. A closed agreement represents that sex only occurs between the primary relationship partners, whereas an open agreement with or without guidelines permits certain (or any) sexual and relational behaviors to occur with casual sex partners.

Implicitly, SAs have direct implications for HIV/STI prevention as this dynamic pertains to couples' sexual behaviors, which may or may not affect their risk for HIV and other STIs. A

recent scoping review summarized the associations reported from previous studies on male couples' establishment and adherence to the agreement and their engagement in condomless anal sex (CAS) within and/or outside of the relationship [17]. In general, negative associations were found between engagement in CAS outside of the relationship and couples who concurred about having an agreement (including type) and to adhering to it [17]. Other work has found that partnered GBMSM's likelihood of having had CAS within and outside of the relationship significantly decreased as their scores of being invested in the SA increased [32].

The associations between couples' SAs and their attitudes toward couple's HIV testing and counseling (CHTC) and pre-exposure prophylaxis (PrEP) [33] has also been explored [34-38]. Findings from these studies point for the need to tailor content and messaging that account for couples' perceived concerns and benefits about using these prevention strategies relative to how it may affect their relationship and agreement.

By definition and in consideration of CIT, the process of establishing an SA could be advantageous for helping couples to reduce their risk for HIV and other STIs in several ways. First, it may provide couples with opportunities to learn and practice communication and negotiation skills, including the facilitation of discussions about their previous and current behaviors pertinent to prevention (eg, sex and substance use) and ways forward. Second, creation of an SA could help couples foster having a joint responsibility and identify ways for partners to support one another and for them to make associated decisions for how best to prevent HIV and other STIs in their relationship. For instance, creating an SA could enable couples with opportunities to decide if and when to use various evidence-based HIV/STI prevention strategies in their relationship according to their HIV serostatus and relational and sexual needs. Such strategies could include condom use, individual HIV/STI testing, CHTC, PrEP, and/or treatment as prevention (TasP) [39-42] with antiretroviral treatment (ART) to obtain and maintain an undetectable viral load to decrease the risk of onward HIV transmission among those living with HIV. As noted in prior work with male couples [43-49], the strategies which couples could include in their agreement may depend on the support and needs of each partner in the relationship, their attitudes toward these strategies, and their value and engagement in behaviors (eg, CAS and substance use) that may increase their risk for HIV yet be at odds with certain dynamics of their relationship (eg, trust and intimacy).

Couples-Based HIV Prevention Interventions for Male Couples

One meta-analysis has concluded that couples-based interventions are more effective in promoting sexual risk reduction behaviors and testing for HIV and other STIs when compared with interventions delivered to individual partners [8]. Although the evidence to support this conclusion is tempered by the limited number of efficacious HIV prevention interventions available for male couples [7], several theoretically informed, couples-based HIV prevention interventions have been developed for male couples [50-59], with several pending dissemination of outcome findings [50,51,54,58]. Many of these

current and upcoming interventions use a tailored approach to accommodate couples' specific needs, incorporate communication and other dynamics in relationship skills–building activities (eg, problem solving), provide sexual health education and HIV/STI prevention-related resources, and encourage the formation of an SA or risk-reduction plan.

Several of the interventions include CHTC as one of the core components, either delivered in person [52,53,55,56] or remotely (ie, video Web-based platform) [51,54]. The number of sessions in the interventions vary, from 2 [50,51], 3 [55] and 4 [52,58] up to 7 sessions [59]. CHTC is a single-session intervention [56] that has also been pilot tested with an added component to address substance use [53]. The delivery time for these interventions also varies, ranging from 45 min for a single session (eg, CHTC) up to 10 or more hours for all sessions.

With respect to specific populations of male couples, 2 of the interventions were designed for young GBMSM in relationships [51,52]: one for methamphetamine-using black male couples [59] and another for predominantly Spanish-speaking Latino GBMSM and their same-sex partners [57,58]. Two of the interventions were developed to attend to the HIV care and adherence needs of male couples with one or both partners living with HIV (ie, serodiscordant and seroconcordant positive) [50,55], whereas some focus on the HIV prevention needs of seroconcordant negative and serodiscordant male couples [51,53,56]. Other interventions address the HIV prevention needs of all 3 groups of couples: seroconcordant negative, seroconcordant positive, and serodiscordant [52,58,59].

To date, 2 of the in-progress interventions are being delivered on the Web [51,54], whereas the rest are being or have been provided in person. In-person interventions for male couples may have limited impact and reachability because of structural barriers (eg, stigma of same-sex behaviors and lack of lesbian, gay, bisexual, transgender, and questioning/queer [LGBTQ]-affirming environments) and the number of resources (eg, appropriately trained personnel and cost) required for successful dissemination and implementation [60-64]. Interventions delivered by a digital health platform (ie, mobile health [mHealth] and electronic health [eHealth]) may help negate some of these limitations and required resources. Couples-based HIV prevention interventions that are delivered by a digital health platform would offer male couples the convenience of accessing the intervention from anywhere with an internet connection and being able to use it in a private setting, thereby providing further privacy, security, safety, and confidentiality. Pending the structure of the intervention, digital health platforms could also help increase reachability as more male couples would be able to use the intervention at any given time compared with those offered in person.

Specific Aims of the Pilot Randomized Controlled Trial of the Electronic Health HIV Prevention Toolkit Intervention

To help increase the number of accessible HIV prevention interventions for male couples in the United States, we leveraged the digital platform of eHealth. The present eHealth, couples-based HIV prevention toolkit intervention, was

developed for seroconcordant HIV-negative male couples, theoretically guided by CIT for couples' health behavior change [11,12], and was based on preliminary work conducted with the target population [29,31,57] and the extant literature [7,28]. The toolkit intervention is an interactive, directed, experiential website aimed to help prepare each couple with the knowledge and skills needed to create a tailored SA that meets the needs of their relationship and for HIV/STI prevention. The specific aims of the pilot randomized controlled trial (RCT) were to (1) assess the feasibility to recruit, enroll, and retain an eligible and consented sample of couples for 6 months; (2) assess the overall acceptability of the toolkit intervention; (3) examine the preliminary impact that using the toolkit intervention will result in a greater proportion of couples to establish and adhere to an SA compared with couples in the control condition; (4) describe the composition of couples' SAs relative to HIV/STI prevention; and (5) examine which relationship dynamics were associated

with couples' establishment and/or adherence to an SA. The trial was not adequately powered to find meaningful differences between trial arms.

Methods

Study Design

All procedures for the pilot RCT occurred on the Web, with couples randomly assigned to 1 of 2 conditions after completion of the baseline assessment. An electronic algorithm was employed to screen and verify couples for study eligibility, followed by manually checking the validity of their data before inviting them to enroll into the pilot RCT. Figure 1 illustrates the Consolidated Standards of Reporting Trials diagram of the RCT. The University of Miami's institutional review board approved all the study procedures. The pilot RCT was registered on ClinicalTrials.gov (NCT02494817).

Figure 1. Results of eligibility screening.
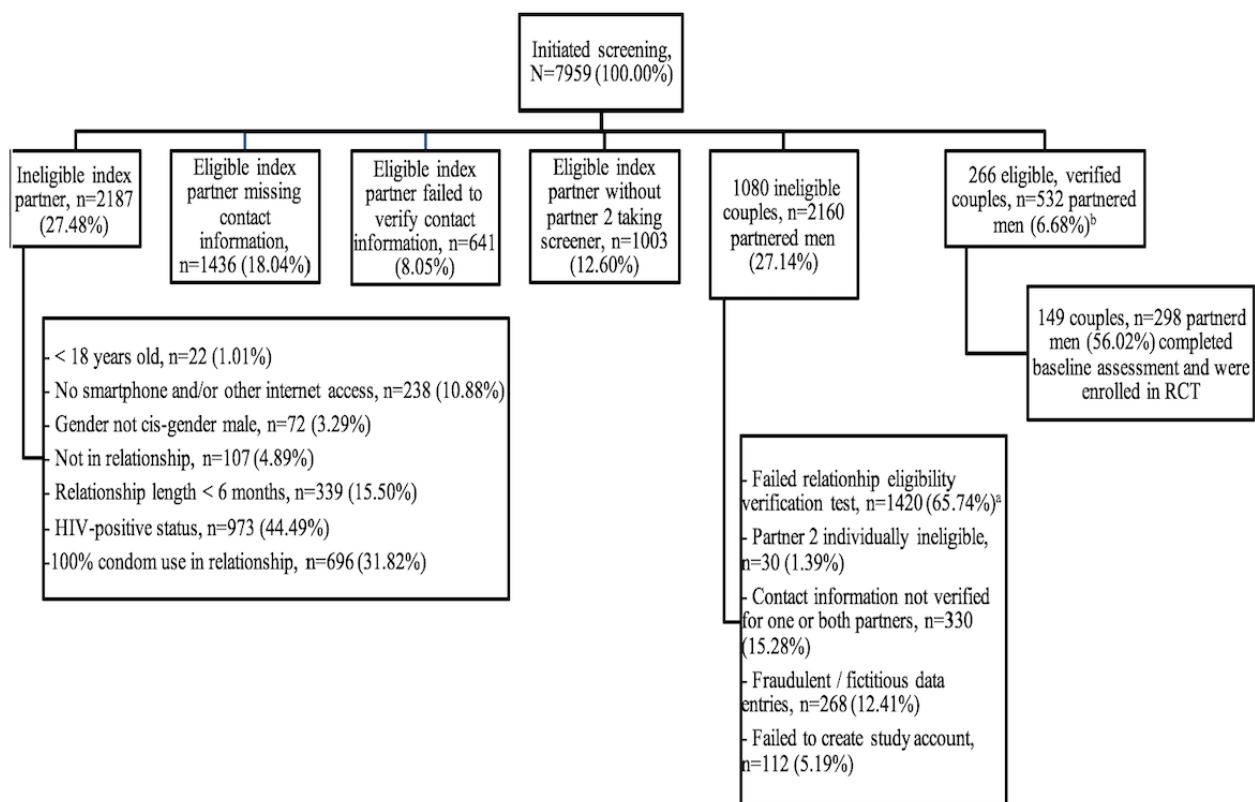
a Among the 1420 partnered men (representing 710 dyads), 1270 partnered men (representing 635 dyads) (89.44%) reported they had established a sexual agreement. Data about type of sexual agreement was not collected.

b All 266 eligible couples created an account on the study website.

Recruitment and Screening Procedures

Targeted advertisements were placed on Facebook to recruit same-sex male couples over the course of 6 months; findings from these campaigns have been previously described and published (blinded). The advertisements targeted English-speaking adult males living in the United States (≥18 years) who had an interest in men and one of these relationship statuses: married, engaged, domestic partnership, civil union, or in a relationship. Each advertisement included a picture of a male couple with a brief study descriptor and a Web link that led interested individuals to the study introductory website. The study introductory website included webpages for the electronic consent document; eligibility screener; inputting and verifying contact information; and an embedded, electronic algorithm that automatically determined study eligibility at the individual and couple levels. The study introductory website was integrated with SurveyGizmo, a Health Insurance Portability and Accountability Act–compliant Web-based survey tool and database server, to collect and store data for the consent and eligibility screener. On the basis of our prior work leading to this pilot RCT, the electronic algorithm—embedded within the Web-based screener—was developed and used to verify whether both partners of the couple were in a relationship with one another and had met all the eligibility criteria.

After providing consent and completing the screener for individual-level eligibility, potential participants (ie, index partner) were then prompted to provide their own and their partners' contact information (eg, email and mobile phone); we refer to this participant as the index partner of the couple. At this point in time, the partner of the index partner (ie, partner 2) would receive an email invitation to join the couples-based study that contained a weblink to the study introductory website so he may follow the same procedures for individual-level eligibility, consent, and inputting and verifying contact information. Each individual who provided consent and passed the individual-level eligibility criteria was asked to verify his contact information. Once an individual entered his contact information, he was sent a passcode to his email address and text on his smartphone. He was then asked to enter these passcodes into the study introductory website to verify his contact information.

Once partner 2 completed the same Web-based screening procedures, contact information and screener items from both partners were used to automatically match and evaluate whether they were in a relationship together (ie, couple) and met the additional couple-level eligibility criteria for enrollment in the pilot RCT. This process is described in the following sections.

Eligibility Criteria

Each partner of the couple—independently—had to meet the following individual-level eligibility criteria to participate in the study: (1) self-report as cis-gender male, (2) aged at least 18 years, (3) be in a current sexual relationship with a main partner for 6 or more months, (4) self-report as HIV negative or unknown serostatus, (5) have had CAS with the primary partner within previous 6 months, (6) self-report no recent history of intimate partner violence or coercion within the

previous year, (7) own a smartphone, and (8) have an alternate method to access the internet (eg, computer).

Couples with one or both partners who did not meet one or more of these criteria were individually ineligible for the study and were automatically informed after completion of the electronic screener. For instance, index partners who self-reported living with HIV received a message thanking them for their interest in the study and that they were ineligible to participate; because of being ineligible, his partner (ie, partner 2) would not have received a study invitation by email. The same ineligibility message was emailed to both partners of couples in instances where they were deemed ineligible and/or did not pass the relationship verification test (see the following sections). Thus, in addition to meeting the eligibility criteria, couples also had to pass the couple-level eligibility criteria through verification and validation tests to enroll for the pilot RCT.

Verification of Couples' Relationships and Validity of Their Data

After screener data were received from both partners, verification of the couples' relationship (ie, couple-level eligibility criteria) was done automatically through the electronic algorithm by evaluating and comparing each partner's response to 5 screener items and using predetermined decisions rules of acceptable responses (see Multimedia Appendix 1). Couples who received a score of 5 on 5 passed the verification test; all other scores were categorized as the couple not passing the verification test. Once a couple was deemed eligible with a verified relationship, we then manually conducted validity checks of their corresponding screener data on a case-by-case basis. Data validity checks consisted of evaluating the following information: repetition of same Internet Protocol (IP) address, use of suspicious participant name(s), presence of duplicate email or fictitious email addresses, back-to-back screener entries, presence of unique data responses to other screener items. For instance, back-to-back screener entries from the same IP address were permitted for a couple as long as all other benchmarks for validation passed. All other instances were flagged as fraudulent and were investigated further by contacting the potential participant/couple for clarification.

Enrollment and Randomization Procedures for Pilot Trial

All couples had to provide consent, pass eligibility and verification criteria, and post hoc validation tests to enroll into the pilot RCT. Through the electronic screener system, consented, eligible, and verified couples were then randomly assigned a unique enrollment ID containing a 4-digit, 2-letter combination that ended with either .01 or .02 to represent the specific partner in each relationship (eg, 1572SP.01 and 1572SP.02). A 4:4 block allocation was electronically generated and used to randomly assign couples' enrollment IDs to 1 of 2 eHealth conditions: an information-only control website or the interactive intervention website. Random assignment was double blinded; however, couples may have guessed which condition they were assigned once they completed the baseline assessment and were granted access to the rest of the eHealth website.

XSL·FO

RenderX

Once the ID was assigned, each partner received an email with instructions to log into the study trial website via a computer to create a profile. After creating a study profile, each partner then proceeded to the assigned eHealth study trial website to complete a 45-min baseline assessment electronically. Follow-up assessments occurred 3 and 6 months after baseline and took approximately 45 min to complete. Each participant was sent up to 2 reminders (email and text) about completing each follow-up assessment. Each participant who completed an assessment was compensated with an electronic gift card incentive worth US $25 for his time.

Description of the Electronic Health HIV Prevention Toolkit Intervention

The toolkit intervention involved participation at both individual and couple levels. At the individual level, participants experienced the interactive website in a directed, sequential fashion before being able to use the website with their relationship partner (ie, couple level). This dual-level intervention design was based on our formative work with 13 same-sex male couples who used the intervention as designed and provided feedback in focus groups (n=7 from Miami, FL, and n=6 from Atlanta, GA), whereas the content and design of the activities and videos were based from our qualitative findings with 29 couples from the metro areas of Detroit, MI, (n=13) and Atlanta, GA, (n=16) [20,22,37]. The vast majority of partnered men from these formative phases stated they wanted an opportunity to read the content, participate in the activities, and have time to digest the material before discussing and comparing their responses to the activities with their relationship partner, including the establish of an SA.

At the individual level, the intervention directed participants through a sequence of instructional and educational videos and modules about evidence-based HIV prevention strategies, communication tips, and SAs. In addition, 3 different activities were also embedded in this series of modules: the creation of a relationship timeline, identification and selection of relationship values, and establishment of an SA via a menu of options arranged by category (see the following sections). After completion of the baseline assessment, each participant was prompted to watch a brief, introductory, 1-min video about the purpose of toolkit intervention and how to use and navigate the website before proceeding to the relationship timeline and value instructional videos and activities. Next, participants were asked to read and review educational content about evidence-based HIV prevention strategies, followed by content on SAs, which included a video that offered suggestions of ways to bring up agreements in the relationship along with some common communication tips (eg, active listening). The last individual-level module was the agreement builder activity with an accompanying instructional video encouraging individuals to begin creating the SA they would like to have with their relationship partner.

Once both partners used the toolkit intervention as directed and added items to their agreement, they were then prompted (via text and email) to sign back into the toolkit intervention website as a couple. Using the toolkit intervention as a couple differed from when participants used it as individuals in important ways.

First, the couple were shown their responses to the relationship timeline and value activities in a comparative fashion, which allowed partners to compare their responses and talk about where they differed and how they were similar. These activities served the purpose to *prime* partners to think about the fond memories they created (to date) and what they valued most about being in a relationship with one another collectively, before considering their future via an agreement. Then, the couple was shown a video about constructive communication tips (eg, negotiation) before proceeding to the agreement builder finalization activity. Similar to the other 2 activities, couples could also see—to a degree—how their individual selection of agreement items compared with one another as these pending items were arranged into 3 groupings: definitively wanted, potentially wanted with need to discuss, and did not want with discussion. Couples then negotiated which items they wanted to accept and place into their agreement or reject and place in the trash bin. Once all items were resolved, each partner would finalize his agreement by entering his unique password to the toolkit intervention.

Once a couple finalized their agreement, they could view all content, activities, and videos freely. Furthermore, the interactive website included a searchable resource center database (Sexual Health Resource Center) that allowed participants to find relevant sexual health resources in the United States and the option to download an app of a simplified version of the toolkit intervention that contained a blueprint of the couples finalized agreement, the ability to SMS/text within the relationship (ie, between partners), and the Sexual Health Resource Center. The educational content, videos, and activities were not available on the app.

HIV-Prevention Content

This educational module included text that described available evidence-based HIV prevention strategies, including female and male condoms, PrEP, nonoccupational postexposure prophylaxis, individual HIV/STI testing, and couple's HIV/STI testing.

Sexual Agreement Content

Another educational module focused on SAs. The content included an overarching definition of an SA along with different types of agreements that exist within the broader LGBTQ community (eg, closed, open with guidelines, and open without guidelines). Additional text drew from the extant literature about male couples' SAs to describe how common agreements are among male couples in the United States, what might motivate some to form an SA, the potential benefits of establishing an agreement in the relationship, whether agreements change over time, and the importance of communicating about the agreement in the relationship [17].

Relationship Timeline Activity

Participants could choose up to ten milestone life events that occurred throughout their relationship. Some examples of the events on the timeline activity included firsts such as *first kiss, first time I met his family (or he met my family), first big purchase together*. Each event was dated by the participant,

which was then automatically placed chronologically in the visual format of a timeline.

Relationship Values Activity

Participants could choose up to 5 items that represented what they valued most about in a relationship with their current partner. Some examples of values presented in this activity were *trusting each other*, *commitment to help our relationship grow*, *accepting our differences*, *feeling sexually satisfied with one another*, and *counting on each other*.

Agreement Builder Activity

Participants could choose as many items as they wanted in their agreement. A total of 96 items were organized in 5 different categories: wellness (20 items; eg, testing for HIV, eating healthy, and supporting each other in our health goals), social etiquette (9 items; eg, holding hands with partner in public and having profiles on social media websites/apps), sex with my partner (23 items; eg, bottoming without condoms with partner, giving or getting head with partner, and group sex play as a couple), sex with other people (22 items; eg, topping with condoms with others, kissing with others, and sexting with others), and drugs (22 items; eg, alcohol with sex, ecstasy without sex, and erectile dysfunction medications). All 5 categories also included the option for participants to *create their own* and add details to each selected item.

Sexual Health Resource Center

This searchable database presented participants with contact and operational information about HIV/STI testing locations throughout the United States by zip code, testing modality preference (eg, individual, CHTC, and over the counter), appointment type (eg, walk-in and appointment required), and cost (eg, free and sliding fee). Locations of pharmacies were also included and searchable by zip code.

Information-Only Control Condition

Couples assigned to the information-only control condition also received an interactive website that contained the same HIV prevention content and Sexual Health Resource Center, along with access to download a similar app as the intervention group sans the blueprint of an agreement.

Measures

All participants, regardless of the study arm, were asked to complete the 3-month and 6-month follow-up questionnaires. The content of follow-up surveys matched the content of the baseline survey, except follow-up surveys also collected information on the formation of, type of, and adherence to an SA. In the event that couples ended their relationship, each partner was still asked to complete their participation throughout the 6 months to collect remaining data and receive their incentives. All data from baseline and follow-up assessments were deidentified, anonymized, and stored on secured servers and password-protected computers.

Outcome Variables

The present analysis focuses on 2 outcomes: (1) establishment of an SA and (2) adherence to a SA. Data for these outcomes were collected at 3- and 6-month follow-up assessments from all participants, regardless of the study arm.

Independent Variables

The baseline assessment captured participants' demographic (eg, state of residence, age, race, ethnicity, sexual orientation, education, employment, and health insurance regular primary provider) and relationship characteristics (eg, relationship length, type, status, and cohabitation) via categorical or dichotomous responses. A number of common sexual behavior items (eg, CAS by partner type) and measures about HIV/STI testing were also captured.

A variety of relationship dynamics were also assessed by using validated instruments for trust [65], relationship commitment [66], relationship satisfaction [67], relationship sexual satisfaction [68], intimacy [69], communication patterns [70], communal confidence [71], use of communal coping strategies to reduce HIV threat [71], preferences for sexual health outcomes [71], HIV social support scale [72], HIV-negative couples' perceptions of severity of HIV infection [71], investment in an SA [24], and preferences for general lifestyle outcomes [71]. Perceptions of local stigma [73], perceptions of gay-related stigma [71], and internalized homophobia [74] were also assessed. HIV-related dyadic measures developed for GBMSM in a relationship [71] offer a quantitative way to assess transformation of motivation and communal coping of CIT. All scales were assessed at all 3 time points (baseline, 3 months, and 6 months), except investment in an SA, which was assessed only at the 3- and 6-month time points. Details about the scales used to capture couples' relationship dynamics are provided in Multimedia Appendix 2.

Analyses

Descriptive statistics were used to summarize cohort characteristics and relationship dynamic variables for the entire cohort by trial arm and by establishment of SA. Dyadic data were calculated for couples if there were no missing values from either partner.

For continuous variables, couple-level mean variables were generated by taking the averaged value from both partners' scores, whereas within-dyad variables (couple-level differences) were generated by taking the absolute difference between 2 partners' scores. Missing values were assigned if either or both partners did not provide a score. Categorical dyadic variables were generated based on whether both partners had the same or different answers. For example, dyadic ethnicity was categorized to 3 levels: both Hispanic, 1 Hispanic, and neither Hispanic. Furthermore, 2-sample *t* tests and chi-square tests were used to evaluate differences between the intervention and control arms for couple-level continuous and categorical independent variables, respectively. To examine the association between relationship dynamics and establishment of an SA, we performed multilevel logistic regression analyses with random intercept for couples to account for correlations of repeated measurements of relationship dynamics at months 3 and 6 and reported the odds ratio (OR) of establishment of an SA and the corresponding 95% confidence interval. All analyses were performed using statistical software R 3.5.2.

Results

Aim 1: Feasibility to Screen, Enroll, and Retain an Eligible and Consented Sample

As shown in Figure 1, 7959 individuals initiated screening resulting in 27.48% (2187/7959) of index partners being ineligible at the individual level; the remaining index partners were eligible at the individual level, but 18.04% (1436/7959) did not provide any contact information, 8.05% (641/7959) failed to verify their contact information, and 12.60% (1003/7959) did not have their partners (ie, partner 2) take the screener. The remaining screeners represented both partners of the couple, with 27.13% (2160/7959) being ineligible at the couple level among other reasons. Overall, 532 partners

representing 266 couples passed the eligibility, verification, and validation screening process and were invited to participate in the pilot RCT via email invitation. Of these 266 couples, 149 (56.0%) were enrolled in the pilot RCT as indicated in their creation of a required study profile and completion of the baseline assessment.

Figure 2 shows retention rates for the 6-month pilot trial at the individual and couple levels. Overall, 71.5% (213/298) of partnered men were retained at the end of the 6-month pilot trial. Retention rates at the 3-month assessment were 67.6% (92/136) of partnered men in the intervention arm and 77.2% (125/162) of partnered men in the control arm ($P=.07$). Retention rates at the 6-month assessment were 72.1% (98/136) of partnered men in the intervention arm and 71.0% (115/162) of partnered men in the control arm (nonsignificant).

Figure 2. Retention rates of pilot randomized controlled trial.
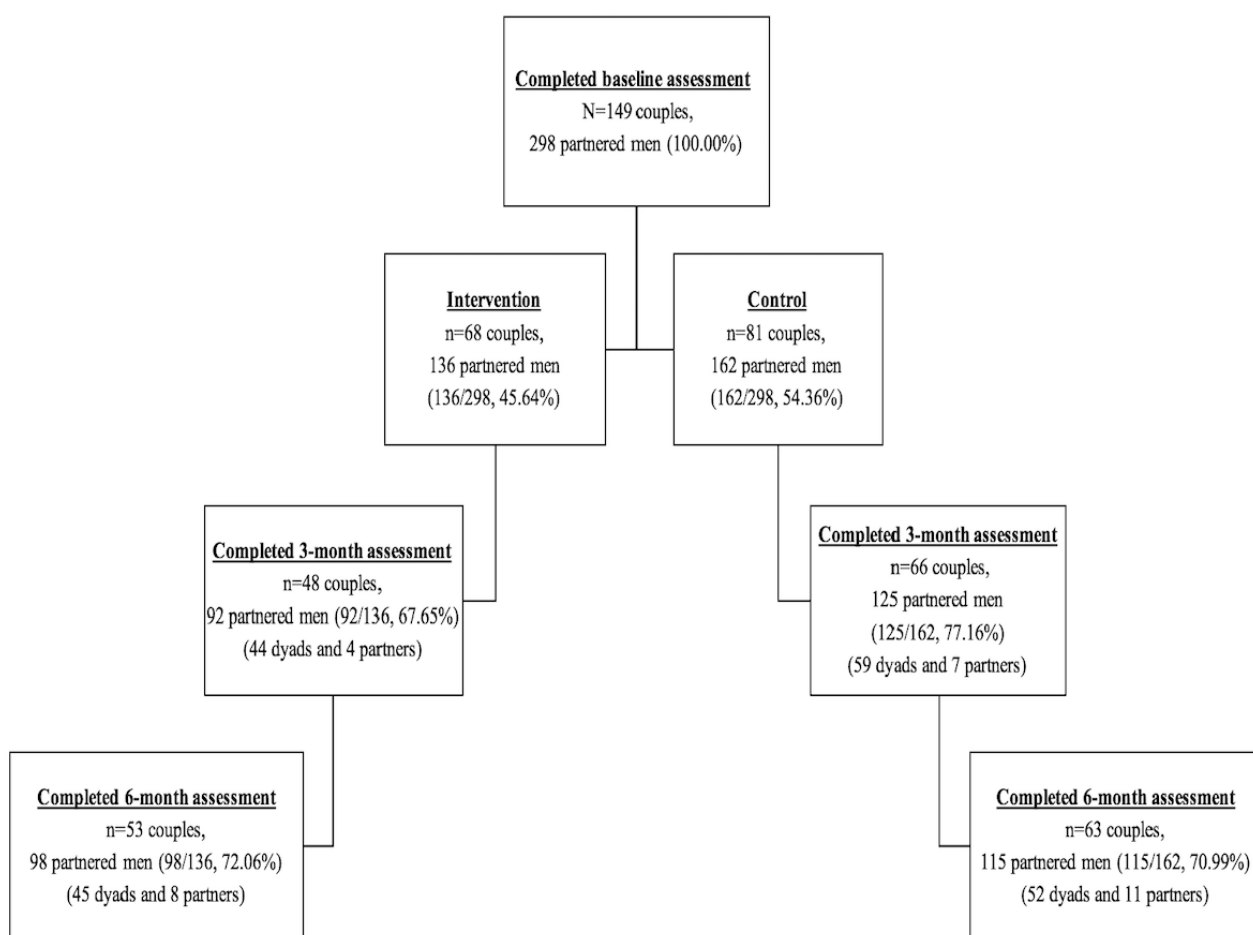
Sample Characteristics

Baseline characteristics and relationship dynamics for the total cohort and by study arm are provided in Tables 1 and 2,

respectively. Given randomization and allocation procedures were used, any differences in baseline characteristics and dynamics are the result of chance rather than bias.

Table 1. Cohort baseline demographic and relationship and sexual behavior characteristics by trial arm.

Characteristics	Cohort	Intervention	Control	P value
Couple-level characteristics[a]				
Ethnicity, n (%)				.64
Both Hispanic	11 (7.5)	5 (7.5)	6 (7.5)	
One Hispanic	26 (17.7)	14 (20.9)	12 (15.0)	
Neither Hispanic	110 (74.8)	48 (71.6)	62 (77.5)	
Race, n (%)				.47
Both white	121 (81.2)	53 (77.9)	68 (83.9)	
Multiracial or other race	28 (18.8)	15 (22.1)	13 (16.1)	
Sexual orientation, n (%)				.68
Both gay	140 (94.0)	65 (95.6)	75 (92.6)	
1 bisexual, 1 gay	9 (6.0)	3 (4.4)	6 (7.4)	
Education attainment, n (%)				.36
Both bachelor's degree or higher	39 (26.5)	18 (26.5)	21 (26.6)	
One bachelor's degree or higher	49 (33.3)	19 (27.9)	30 (38.0)	
Neither have at least bachelor's degree	59 (40.1)	31 (45.6)	28 (35.4)	
Employment status, n (%)				.84
Both employed	98 (65.8)	46 (67.7)	52 (64.2)	
One employed	34 (22.8)	14 (20.6)	20 (24.7)	
Neither employed	17 (11.4)	8 (11.8)	9 (11.1)	
Health insurance, n (%)				.77
Both have	112 (75.2)	53 (77.9)	59 (72.8)	
One has	27 (18.1)	11 (16.2)	16 (19.8)	
Neither has	10 (6.7)	4 (5.9)	6 (7.4)	
Regular general physician/MD, n (%)				.34
Both have	85 (57.1)	42 (61.8)	43 (53.1)	
One has	42 (28.2)	19 (27.9)	23 (28.4)	
Neither has	22 (14.8)	7 (10.3)	15 (18.5)	
US region of residence[b] , n (%)				.16
Northeast	14 (9.4)	5 (7.3)	9 (11.1)	
Midwest	40 (26.8)	24 (35.3)	16 (19.8)	
South	53 (35.6)	20 (29.4)	33 (40.7)	
West	37 (24.8)	18 (26.5)	19 (23.5)	
Two regions, long-distance[c]	5 (3.3)	1 (1.5)	4 (4.9)	
Relationship type, n (%)				.14
Monogamy	130 (87.8)	56 (82.4)	74 (92.5)	
Open	8 (5.4)	6 (8.8)	2 (2.5)	
Discrepant reports	10 (7.8)	6 (8.8)	4 (5.0)	
Relationship status[d] , n (%)				.08
Long-term oriented	85 (57.1)	45 (66.2)	40 (49.4)	
Boyfriends	41 (27.5)	13 (19.1)	28 (34.6)	
Partners reported differently	23 (15.4)	10 (14.7)	13 (16.1)	

XSL•FO
RenderX

Characteristics	Cohort	Intervention	Control	P value
Relationship length in years, mean (SD)	3.1 (3.66)	3.8 (4.4)	2.6 (2.7)	.04
Age difference between partners in years, mean (SD)	3.6 (3.9)	2.9 (3.2)	4.1 (4.4)	.07
Ever had an HIV test, n (%)				.82
Both have	108 (72.5)	51 (75.0)	57 (70.4)	
One has	34 (22.8)	14 (20.6)	20 (24.7)	
Neither has	7 (4.7)	3 (4.4)	4 (4.9)	
Ever had an STD[e] test, n (%)				.53
Both have	88 (59.1)	40 (58.8)	48 (59.3)	
One has	39 (26.2)	20 (29.4)	19 (23.5)	
Neither has	22 (14.8)	8 (11.8)	14 (17.3)	
Has had sex with a casual MSM[f] partner in prior 3 months, n (%)				.11
Both have	14 (9.4)	10 (14.7)	4 (4.9)	
One has	39 (26.2)	18 (26.5)	21 (25.9)	
Neither has	96 (64.4)	40 (58.8)	56 (69.1)	
Individual-level characteristics[g]				
Age (years, range: 18-58), mean (SD)	27.8 (7.16)	28.1 (7.33)	27.6 (7.03)	.49
Average number of condomless anal sex episodes with partner in prior 3 months, mean (SD)				
Insertive role	8.4 (14.0)	6.3 (11.9)	10.2 (15.4)	.02
Receptive role	7.9 (13.1)	4.9 (7.7)	10.4 (16.0)	<.001
Insertive and receptive in same episode	1.7 (5.8)	0.6 (1.7)	2.6 (7.6)	<.01
Average number of casual MSM partners in prior 3 months (n=68), mean (SD)	3.8 (6.1)	3.8 (6.3)	3.7 (6.0)	.69
Average number of anal sex episodes with casual MSM partner(s) in prior 3 months (n=67), mean (SD)	0.6 (1.7)	0.3 (0.7)	1.0 (2.5)	.11
Average number of condomless anal sex episodes with casual MSM partner(s) in prior 3 months (n=17), mean (SD)				
Insertive role	1.4 (2.2)	1.0 (0.7)	1.8 (3.2)	.5
Receptive role	2.7 (7.2)	0.6 (1.3)	5.1 (10.2)	.2

[a]Cohort, intervention, and control included 149, 68, and 81 couples, respectively.

[b]States and territories not represented: Guam, US Marshall Islands, Alaska, New Mexico, New Hampshire, Maine, Mississippi, and North Dakota.

[c]9 couples were in a long-distance relationship, 4 of whom resided in states within the same US region, whereas 5 couples had partners living in states in 2 different regions (Colorado and Illinois, Florida and Massachusetts, Michigan and Washington, Pennsylvania and Virginia, and Pennsylvania and Wisconsin).

[d]Long-term oriented was classified as couples who had both partners self-reporting one of the following: being married, engaged, had a commitment ceremony, or in a domestic partnership. Boyfriend category included couples who had both partners self-reporting as boyfriends, in a relationship, or none of the above. Discrepant reports represented couples in which one partner reported an option within the long-term oriented classification and the other partner reported an option within the boyfriend classification.

[e]STD: sexually transmitted disease.

[f]MSM: men who have sex with men.

[g]Cohort, intervention, and control included 298, 136, and 162 men, respectively.

Table 2. Cohort baseline relationship dynamics by trial arm.

Relationship dynamic	Cohort (298 men, 149 couples)	Intervention (136 men, 68 couples)	Control (162 men, 81 couples)	P value
Dyadic trust scale, mean (SD)				
Individual score	4.27 (0.76)	4.28 (0.74)	4.25 (0.77)	.74
Score difference between partners	0.61 (0.61)	0.64 (0.59)	0.59 (0.63)	.68
Investment model scale for relationship commitment, mean (SD)				
Individual score	5.13 (0.80)	5.15 (0.81)	5.11 (0.79)	.68
Score difference between partners	0.71 (0.68)	0.68 (0.67)	0.74 (0.68)	.59
Relationship satisfaction, mean (SD)				
Individual score	4.34 (0.73)	4.29 (0.74)	4.38 (0.71)	.29
Score difference between partners	0.57 (0.58)	0.60 (0.56)	0.55 (0.60)	.61
Miller social intimacy scale, mean (SD)				
Individual score	8.63 (1.07)	8.63 (1.06)	8.62 (1.08)	.96
Score difference between partners	0.90 (0.97)	0.88 (0.89)	0.92 (1.04)	.77
Avoidance and withdrawal communication pattern, mean (SD)				
Individual score	3.60 (1.53)	3.63 (1.50)	3.57 (1.55)	.73
Score difference between partners	1.19 (0.91)	1.08 (0.91)	1.29 (0.90)	.16
Constructive communication pattern, mean (SD)				
Individual score	6.57 (1.82)	6.34 (1.90)	6.77 (1.73)	.04
Score difference between partners	1.56 (1.28)	1.59 (1.25)	1.53 (1.32)	.78
Couple's communal confidence, mean (SD)				
Individual score	27.60 (5.00)	27.53 (5.09)	27.66 (4.93)	.82
Score difference between partners	3.68 (3.22)	3.82 (3.20)	3.57 (3.25)	.63
Communal coping strategies to reduce HIV threat, mean (SD)				
Individual score	4.06 (0.89)	4.10 (0.83)	4.02 (0.94)	.40
Score difference between partners	0.90 (0.81)	0.76 (0.68)	1.01 (0.89)	.07
Preferences for general lifestyle outcomes, mean (SD)				
Individual score	23.08 (3.67)	23.24 (3.87)	22.95 (3.51)	.38
Score difference between partners	3.16 (2.53)	3.28 (2.66)	3.06 (2.42)	.60
Preferences for sexual health outcomes, mean (SD)				
Individual score	31.61 (4.33)	31.66 (4.05)	31.57 (4.57)	.87
Score difference between partners	4.22 (4.40)	3.79 (4.01)	4.58 (4.69)	.28
HIV social support scale, mean (SD)				
Individual score	3.32 (0.38)	3.35 (0.36)	3.29 (0.40)	.22
Score difference between partners	0.38 (0.30)	0.39 (0.30)	0.37 (0.30)	.73
HIV-negative couples' perceptions of severity of HIV infection, mean (SD)				
Individual score	3.78 (0.76)	3.79 (0.73)	3.76 (0.78)	.75
Score difference between partners	0.71 (0.55)	0.68 (0.53)	0.74 (0.56)	.51
Sexual satisfaction with the relationship, mean (SD)				
Individual score	3.78 (0.89)	3.69 (0.91)	3.85 (0.87)	.12
Score difference between partners	0.69 (0.63)	0.78 (0.76)	0.61 (0.49)	.09
Perceptions of local stigma, mean (SD)				
Individual score	3.91 (0.94)	3.92 (0.99)	3.91 (0.90)	.93

XSL•FO
RenderX

Relationship dynamic	Cohort (298 men, 149 couples)	Intervention (136 men, 68 couples)	Control (162 men, 81 couples)	P value
Score difference between partners	0.82 (0.65)	0.78 (0.65)	0.85 (0.65)	.51
Perceptions of gay-related stigma, mean (SD)				
Individual score	4.18 (0.66)	4.27 (0.63)	4.10 (0.67)	.03
Score difference between partners	0.56 (0.47)	0.52 (0.51)	0.59 (0.43)	.34
Internalized homophobia, mean (SD)				
Individual score	1.66 (0.54)	1.68 (0.60)	1.64 (0.49)	.50
Score difference between partners	0.51 (0.48)	0.50 (0.51)	0.52 (0.45)	.82

Aim 2: Use and Acceptability of Toolkit Intervention

Over the period of 6 months, participants in the intervention arm logged into their eHealth toolkit an average of 13.42 times (range 1-38) compared with participants in the control arm who used their information-only website an average of 4.48 times (range 1-23). In total, 64.1% (191/298) participants downloaded the accompanying study app onto their smartphone: 65.4% (89/136) participants in the intervention arm (89 men representing 63 couples) and 63.0% (102/162) of participants in the control arm (102 men representing 74 couples). Differences were noted by arm with respect to whether one or both partners of the couple downloaded the app onto their smartphone. Specifically, a higher proportion of couples in the intervention arm (26/63 dyads, 41%) had both partners download the app compared with those in the control arm (28/74 dyads, 38%).

With respect to the acceptability of the eHealth HIV prevention toolkit, participants in the intervention arm also provided data about their perceptions of how easy it was to use various components of it, ranging from navigating the interactive website to using the agreement builder activity (Table 3). Participants reported, on average, that using different aspects of the intervention was *easy* for most items assessed across both time points. They also perceived downloading the accompanying smartphone app and using the Sexual Health Resource Center on the app was slightly less than *easy*, falling somewhere between *neither difficult nor easy* and *easy* across both time points.

Participants further reported how often they thought they would use an activity like the agreement builder with their partner in their relationship. As shown in Table 3, their responses varied at both assessment time points. About 38.0% (38/100) of participants thought they would use this type of activity on an as-needed basis, whereas between 28% and 32% of participants reported they would use this type of activity at a regular interval (ie, every 3-4 months, every 6 months, or yearly) in the relationship with their partner. In contrast, between 19% to 26% of the participants were not sure about how often they would use this type of activity, and about 8% of participants chose never.

Table 3. Acceptability data among participants in the intervention arm, by assessment time point.

Acceptability item	3-month assessment	6-month assessment
Item stem: How easy was it for you to..., mean (SD)		
Navigate the different sections of the toolkit website?	4.10 (0.91)	3.94 (1.04)
Use the sexual health center on the toolkit website?	4.04 (0.84)	3.93 (0.97)
Download the toolkit app onto your smartphone?	3.42 (1.03)	3.49 (1.23)
Use the Sexual Health Resource Center on your smartphone app?	3.44 (1.03)	3.53 (1.15)
Use the agreement builder activity—by yourself—to identify what items you wanted in an agreement with your partner?	3.94 (0.95)	4.05 (0.94)
Negotiate and finalize the items you wanted in the agreement with your partner?	3.88 (1.01)	3.90 (1.07)
Item: Now that you have experienced the agreement builder activity, how often do you think you would use this type of activity while in your current relationship with [partner's first name/nickname]?, n (%)		
Every 3-4 months	18 (18.0)	9 (8.7)
Every 6 months	8 (8.0)	15 (14.4)
Every 12 months	2 (2.0)	9 (8.7)
On as-needed basis	38 (38.0)	40 (38.5)
I'm not sure	26 (26.0)	19 (18.3)
Never	8 (8.0)	12 (11.5)
Item asked at 6 months: Please share any suggestions and/or thoughts that you may have about your experience of using the toolkit intervention. (Participant age, US state of residence, relationship length, agreement type), n		
"Surveys were quite lengthy" (27, CT, 3.3 years, closed agreement)	—[a]	23
"This helped me understand my relationship better. Going through the toolkit every few months made me realize how much things change in relationships over the course of six months." (20, IN, 6 months, closed agreement)	—	21
"Too many agreement items… felt overwhelmed by the choices." (30, CA, 3.2 years, open agreement)	—	17
"My partner and I liked the idea of the toolkit, but we weren't sure how often we would use it. It would be nice to have more to do [with it] over time." (39, TN, 4.3 years, open agreement)	—	14
"Since the last time I used this, me and my partner's relationship has gotten stronger and I believe by reading these questions and answering them has helped us communicate and work on building a brighter future for each other. So I want to say thank u so very much." (29, OR, 4.6 years, closed agreement)	—	12

[a]Not applicable.

Aim 3: Establishment, Type, and Adherence to a Sexual Agreement

Table 4 provides data about the proportion of couples who established an SA, the type of agreement formed, and whether they adhered to the agreement by trial assessment time point (ie, at 3 and 6 months). Among couples who had both partners provide data, almost two-thirds (63.4%) had established an SA at the 3-month assessment, with a nonsignificantly higher proportion of couples in the intervention arm (29/42, 69%) forming one compared with those in the control arm (35/59, 59%; P=.40). At the 6-month assessment, 63.4% of couples had established an SA, with a significantly higher proportion of couples in the intervention arm (32/43, 74%) forming one compared with those in the control arm (27/50, 54%; P<.05). In each arm at both time point assessments, the remaining proportion of couples did not establish an SA.

For both assessment time points, a nonsignificantly higher proportion of couples in the control arm reported having a closed agreement than couples in the intervention arm (3 month: 33/35, 94% vs 21/29, 72%, P=.07; 6 months: 24/27, 89% vs 26/32, 81%, P=.87). In contrast, a nonsignificantly higher proportion of couples in the intervention arm reported having an open agreement containing guidelines than those in the control arm (3 months: 4/29, 14% vs 1/34, 3%, P=.07; 6 month: 2/32, 6% vs 1/27, 4%, P=.87). Similarly, a nonsignificantly higher proportion of couples in the intervention arm had partners who disagreed about their agreement type than those in the control arm (3 months: 4/29, 14% vs 1/34, 3%, P=.07; 6 months: 4/32, 13% vs 2/27, 7%, P=.87). Although couples' type of agreement did not significantly differ by trial arm at either assessment time point, there was a trend at the 3-month assessment, with more couples in the intervention arm having formed an open SA with guidelines compared with those in the control condition (P=.07).

XSL•FO
RenderX

Table 4. Couples' establishment, type, and adherence to a sexual agreement by trial arm and assessment time point.

Aspect of sexual agreement	3-month assessment				6-month assessment			
	Intervention, n (%)	Control, n (%)	Difference, %	*P* value	Intervention, n (%)	Control, n (%)	Difference, %	*P* value
Establishment	42 (100)	59 (100)	—a	.4	43 (100)	50 (100)	—	.05
Yes	29 (69)	35 (59)	9.8		32 (74)	27 (54)	20.4	
No/did not concur	13 (31)	24 (41)	−9.8		11 (26)	23 (46)	−20.4	
Type	29 (100)	35 (100)	—	.07	32 (100)	27 (100)	—	.87
Closed	21 (72)	33 (94)	−21.9		26 (81)	24 (89)	−7.6	
Open with guidelines	4 (14)	1 (3)	10.9		2 (6)	1 (4)	2.6	
Disagreed about type	4 (14)	1 (3)	10.9		4 (13)	2 (7)	5.1	
Adherence	29 (100)	34 (100)	—	>.99	32 (100)	27 (100)	—	.4
Yes, by both partners	27 (93)	31 (91)	1.9		30 (94)	23 (85)	8.6	
No, by at least one partner	2 (7)	3 (9)	−1.9		2 (6)	4 (15)	−8.6	

a

Among couples who established an SA with both partners providing data, 92.1% had adhered to their agreement at the 3-month assessment, with a slightly nonsignificantly higher proportion of couples adhering to theirs in the intervention arm (27/29, 93%) compared with those in the control arm (31/34, 91%; *P*>.99). At the 6-month assessment, a nonsignificantly higher proportion of couples in the intervention arm self-reported adhering to their agreement (30/32, 94%) compared with those in the control arm (23/27, 85%; *P*=.40). The remaining proportion of couples, in each arm at both time point assessments self-reported not adhering to their agreement.

Aim 4: Composition and Investment in the Sexual Agreement

When using the agreement builder exercise, couples in the intervention arm, on average, included 18 items in their agreement (range 3-56). The types of items couples had in their agreement varied (Table 5). Overall, couples added more items about wellness than any other category; in contrast, items about drug use were the least included. With respect to HIV prevention, which included items in the wellness, sex with partner, and sex with others categories, 38% (11/29) of couples included regular testing of STIs; 28% (8/29) for regular testing of HIV; 31% (9/29) for topping without condoms with partner; 45% (13/29) for bottoming without condoms with partner; and 28% (8/29) specified sex or no sex with other/casual men who have sex with men partners.

Table 5. Couples' average and range of number of items included in their sexual agreements by agreement category.

Item included	Wellness	Sex with partner	Sex with others	Social etiquette	Drugs
Average number of items	9.24	4.58	0.97	2.06	1.21
Range of number of items	3-17	0-17	0-13	0-8	0-5

A number of couples also included items aimed at strengthening and affirming their relationship; these items were located in the wellness and social etiquette categories of the agreement builder. Specifically, 76% (22/29) of couples included talking to/listening to each other; 66% (19/29) had sharing hobbies; 93% (27/29) for going on dates together; 93% (24/26) for going on vacations together; 45% (13/29) included career/education/job support; and 55% (16/29) had being affectionate with partner in public, holding hands in public, and/or had public recognition of relationship.

In addition, many of the couples included items about health-promotive behaviors. For example, 86% (25/29) of couples included exercising more; 86% (25/29) had eating healthier; 76% (22/29) for managing stress; 31% (9/29) had

medical, dental, and eye check-ups; and 55% (16/29) included supporting each other in their health goals.

Multimedia Appendix 3 provides data about participants' investment in the SA and within-dyad score differences for Sexual Agreement Investment Scale (SAIS) [24]. On average, participants for the entire cohort and by study arm were between *very* and *extremely* invested in the SA they created with their relationship partner. Participants were also committed to it, satisfied with it, and valued the SA as noted by their averaged scores. No significant differences for SAIS were found for individual and within-dyad scores between the 2 trial arms.

XSL•FO

RenderX

Aim 5: Odds Ratio of Establishing a Sexual Agreement Relative to Couples' Relationship Dynamics

Multimedia Appendix 4 describes results from multilevel logistic regression that modeled the OR that couples established an SA via self-reported averaged relationship dynamic scores (ie, couples' mean and absolute mean difference between partner's scores) adjusting for months of assessment and trial arm. Given the exploratory nature of the pilot RCT, we used a P value of .10 and less to detect whether a potentially meaningful (ie, a signal) difference was noted between the 2 trial arms for predicting couples' establishment of an SA over time.

After controlling for averaged couple score of constructive communication, the OR of establishing an SA for couples in the intervention group versus couples in the control group was 2.33 (95% CI 0.86-6.31; P=.09). Similar results were found when controlling for averaged couple scores of preferences for sexual health outcomes (OR 2.23, 95% CI 0.85-5.89; P=.10), perceived gay-related stigma (OR 2.53, 95% CI 0.95-6.75; P=.06), and internalized homophobia (OR 2.26, 95% CI 0.84-6.10; P=.10).

When controlling for averaged within-dyad score for relationship commitment, the OR of establishing an SA for couples in the intervention group versus couples in the control group was 2.24 (95% CI 0.84-5.97; P=.10). Similar ORs for intervention group versus the control group were found when controlling for within-dyad score differences of sexual satisfaction (OR 2.31, 95% CI 0.86-6.16; P=.09), social intimacy (OR 2.30, 95% CI 0.87-6.10; P=.09), avoidance and withdrawal communication pattern (OR 2.34, 95% CI 0.84-6.50; P=.10), constructive communication pattern (OR 2.30, 95% CI 0.86-6.16; P=.10), communal coping strategies to reduce HIV threat (OR 2.26, 95% CI 0.84-6.10; P=.10), preferences for sexual health outcomes (OR 2.34, 95% CI 0.88-6.24; P=.09), HIV social support (OR 2.31, 95% CI 0.85-6.26; P=.09), perceived local stigma (OR 2.40, 95% CI 0.89-6.47; P=.08), perceived gay-related stigma (OR 2.38, 95% CI 0.88-6.43; P=.09), and internalized homophobia (OR 2.33, 95% CI 0.88-6.20; P=.09).

When controlling for trial group assignment, the odds of establishing an SA increased by 101% for each unit increase in couples averaged dyadic trust score (OR 2.01, 95% CI 0.94-4.32; P=.07). Similar results were found when controlling for trial group assignment for couples averaged scores of relationship satisfaction (OR 3.08, 95% CI 1.45-6.55; P<.01), social intimacy (OR 1.88, 95% CI 1.07-3.32; P=.03), constructive communication (OR 1.46, 95% CI 1.08-1.96; P=.01), communal confidence (OR 1.13, 95% CI 1.01-1.26; P=.03), communal coping strategies to reduce HIV threat (OR 4.22, 95% CI 2.04-8.73; P<.001), and perceptions of severity of HIV infection (OR 1.91, 95% CI 0.95-3.83; P=.07). In addition, the odds of establishing an SA decreased by 79% for each unit increase in couples averaged perceived local stigma score (OR 0.21, 95% CI 0.11-0.43; P<.001) after controlling for trial group assignment; a similar result was also found for perceived gay-related stigma (OR 0.35, 95% CI 0.16-0.79; P=.01).

Aim 5 Differences in Couples' Relationship Dynamics Relative to Adherence to a Sexual Agreement Over Time

The majority of couples adhered to their SA at 3- and 6-month assessments (see Table 4). Small sample sizes of nonadherence to an SA inhibit our ability to meaningfully assess whether relationship dynamics were associated with this outcome over time and by trial arm.

Discussion

The findings from this pilot RCT suggest the feasibility and acceptability of an eHealth HIV prevention toolkit intervention to encourage establishment and adherence to an SA among seroconcordant negative male couples.

Feasibility

A little more than half of couples (149/266 dyads, 56.0%) who could have enrolled did enroll by following the required steps (ie, create a profile on the study website and complete the baseline assessment). It is possible some men may have found these steps to be cumbersome and/or changed their minds about participating after the eligibility and consent portions of the study. In addition, it is also possible that the decision to participate in the study may be linked to relationship dynamics: those with poor communication may opt to not enroll in a study for male couples. For a future trial, modified enrollment steps could be used to simplify the procedures and to help increase the likelihood of couples following through with the necessary components to participate in the trial. First, a Zoom or phone meeting might help inform eligible participants of what is involved for participating in the trial and lead to higher follow-through rates of enrollment. This added step of enrollment has been implemented in an mHealth HIV testing RCT with GBMSM and has led to higher enrollment rates [75]. Second, the added step of requiring participants to create a user profile for the toolkit could be shortened by using data collected from their responses to the eligibility screener and consent. Specifically, the study website portal could automatically generate a user profile for each partner in an eligible, consented, verified male couple. This change would allow men to complete less information, take less time, and simplify the process by having them choose which contact information method they would like to verify (email address or text for mobile number vs both) and a security question to allow them to reset their password.

Some men also reported that the assessments were too time consuming. It is further possible that participants may have also perceived the compensation to be inadequate for the time it required for them to complete each assessment. These possibilities may help explain the retention rates observed for the pilot trial. Several improvements could be made for a future trial. Future assessments could be shortened by preventing overlap of measures across scales. For instance, the Relationship Satisfaction scale [67] could be eliminated as the Relationship Satisfaction subscale in the Investment Model [66] captures similar information about this dynamic. A subscale, instead of the complete scale, could also be used if it aligns with the

theoretical framework of the intervention and overall study. For example, to capture and assess changes in mutual constructive communication patterns over time (3 items), this subscale could be used instead of the entire Communication Patterns Scale (3+8 items) inclusive of the Avoidance and Withdrawal subscale, thereby eliminating 8 questions [70]. As smartphone use continues to increase among the general US population [76], the likelihood that future participants will take their assessments on a smartphone will increase. As such, it will become increasingly important to reduce the amount of time it takes for participants to complete an assessment (eg, from 45 to 20 min) to help increase retention rates. Moreover, a larger study could provide more funds to compensate participants for their time (eg, US $40 vs US $25), which may also help with improving retention. These changes, collectively, could help improve the feasibility of a future, larger RCT of this intervention as other procedures were found to be acceptable without issue (eg, randomization using block allocation, double blinded).

Acceptability and Use

Overall, participants reported high acceptability of the toolkit intervention. This was particularly true for navigating and using the different components of the interactive website and for individually selecting and then negotiating and finalizing the creation of an SA with their relationship partner. Their acceptability was slightly lower for downloading the accompanying smartphone app and using the Sexual Health Resource Center on the app. It is possible that participants may have had connectivity issues in downloading and/or while using the app, thereby influencing their attitudes toward this part of the toolkit. It is also possible that participants may have perceived the app to be too simplistic and questioned the need for it given the stark contrast of what the app offered compared with the interactive website. Moreover, the items used to assess participants' acceptability of this couples-based intervention may not have captured all key elements and/or their attitudes about it. Future digital health, couples-based interventions may want to consider using the Health Information Technology Usability Evaluation Scale, a customizable usability evaluation instrument that includes subscales of impact, perceived usefulness, ease of use, and user control [77].

Furthermore, a future iteration of this couples-based HIV/STI prevention intervention toolkit may need to be offered in a variety of formats to further increase reach, access, and acceptability among the target population. It is possible that some of the eligible and consented couples who chose not to enroll (eg, 117/266), did so because they realized the intervention could only be used on a laptop of desktop computer as it was not optimized for smaller screen devices. Moving forward, the toolkit may need to be delivered on a single responsive website that would work across all types of internet-connected devices, including smartphones, tablets, laptops, and desktop computers. It could also be offered on an app for smartphones and tablets, which would allow researchers to send reminders and notifications directly to participants (eg, time to complete an assessment). Future research with the target population is warranted to help decide whether one or both of these options for the next iteration of the toolkit ought to be offered.

Acceptability of the agreement builder activity must also be considered. Overall, participants liked the agreement builder activity and how they experienced and used it (solo followed by as a couple). They also provided feedback about how often they thought they would use it over time. About one-third of the participants thought they would use this activity on a regular basis (ie, at some interval), one-third of them perceived they would use it on an as-needed basis, and one-quarter of them were unsure; few of them said they would never use this type of activity. Similar to the importance of being tested for HIV/STIs at a regular interval (eg, every 3, 6, or 12 months), we believe using the agreement builder activity at a regular interval would be beneficial for the couple. SAs are fluid and could change over time to reflect partners' and couples' evolving needs. This type of activity would allow couples to revisit and change their agreement, and it would also provide couples with opportunities to help improve their understanding about behaviors they wish to agree to engage in and not engage in (ie, within-couple concordance). Findings from a recent study with male couples from Boston, Atlanta, and Chicago support this idea. The authors reported weak-to-moderate concordance on couples' agreements guidelines that pertained to having sex outside of the relationship and for specific sexual behaviors they allowed or disallowed to occur [78]. Although we do not think couples ought to be forced into these types of conversations, a toolkit could be programmed to periodically check in with each partner of the couple to assess their overall satisfaction with the agreement and whether their sexual health and relationship needs have changed from when they first created their agreement or from their last check-in. A future version of this activity could provide this kind of check-in mechanism, either preprogrammed or by a time interval (eg, quarterly) set by both partners of the couple.

Participants mentioned another area of the SA builder activity that warrants attention. Some perceived the agreement builder activity contained too many items for them to consider for their SA (Table 3). In addition, approximately one-quarter to one-third of couples included HIV/STI prevention items in their SA (Table 5), and 28% specified whether sex was permitted with casual GBMSM partners. In its current form, the agreement builder activity enabled couples to choose and select items for their SA from a menu consisting of 5 categories with a total of 96 items. This approach, although deemed to be acceptable in our formative work leading to the pilot trial, may have diminished the focus on HIV/STI prevention and overwhelmed some of the partners/couples given the array of choices. It is also possible that some of the couples may have perceived their risk for HIV/STIs to be low and opted to not include any items about prevention. Prior research has found that couples perceived their risk for HIV and other STIs to be generally low, in part because of their beliefs that being in a relationship—by virtue—incorrectly reduces their risk or protects them from HIV/STIs [79]. One possible solution to encourage couples to include HIV/STI prevention items in their SA is to restructure, streamline, and simplify the agreement builder activity. First, an electronic algorithm could be embedded in the activity to prompt each partner of the couple to answer a brief set of questions to gauge the kind of sexual relationship they would want and the types of sexual behaviors they would prefer to

engage in. Their responses to these questions could then automatically generate and place HIV/STI prevention items in their agreement for a more directed approach. Furthermore, the agreement builder activity could be broken down into several segments for couples to complete over time and not in one sitting. For example, once a couple decides which HIV/STI prevention items to include in their SA, they could then be prompted to revisit the agreement builder activity to focus on a different area that they deem to be important, such as strengthening and affirming their relationship. Changing the agreement builder activity is these ways (ie, algorithm, directed, and staggered) may help encourage couples to use the toolkit over time and simplify the process of building an agreement that meets their prevention and relationship needs (while lessening their feelings of being overwhelmed by too many choices).

Sexual Agreement Outcomes

The preliminary impact of the eHealth HIV prevention toolkit intervention on couples' establishment and adherence to an SA was also assessed. Compared with couples in the control arm, more couples in the intervention arm established an SA over time. Although a significant difference for establishing an SA was found at the 6-month follow-up between the 2 trial arms, the pilot trial was not adequately powered as we were more interested in obtaining point estimates and trends. These findings show initial promise for the toolkit intervention to help encourage couples who did not have an SA to establish one. However, there may be other possibilities that influenced couples to establish an SA, either apart (for couples in either trial arm) or in addition to using the toolkit (intervention arm only). Prior research has described that for some couples, certain circumstances or experiences (eg, events with others and influences from peers) may have led them to forming an SA [22]. It is also possible that couples established an SA as part of their natural progression in the relationship [19,31] and to enhance or improve an aspect of their relationship (eg, trust and intimacy) [21]. Future couples-based research that includes the establishment of an SA in the intervention would benefit to include an evaluation item to assess what influenced couples to form an agreement in their relationship.

A number of common relationship dynamics (eg, constructive communication, intimacy, and communal coping strategies to reduce HIV threat) at the averaged couple level were positively associated with couples establishing an SA—in general and over time. Similar findings were noted for lower averaged partner score differences being positively associated with couples establishing an SA. These findings align with what prior research with male couples has highlighted [6,26,80,81]: including and bolstering relationship dynamics along with sexual identity affirmation in couples-based interventions is critically important for HIV/STI prevention. It should be noted that findings from this trial suggest men's perceptions about how much stigma there is for being gay in their local community and for being in a same-sex relationship may play an important role in HIV/STI prevention with male couples by decreasing their odds of establishing an SA. Specifically, as scores of the averaged couple level and differences between partners increase for these measures, the odds of a couple establishing an

agreement decrease between 65% and 79%. Limited research has investigated the role that male couples' living and social environment(s) may have toward their risk for HIV/STIs [80,82,83], particularly with respect to internalized and perceived stigma. Further research is warranted to examine the ways in which stigma may impact male couples' relationships and efforts related to HIV/STI prevention.

With respect to adherence, fewer couples in the intervention arm broke their SA over time compared with couples in the control arm. Differences between the 2 trial arms were nonsignificant for both follow-up time points. Sample size constraints prevented our ability to quantitatively assess and meaningfully detect whether any differences in relationship dynamics existed between couples who broke their agreement compared with those who adhered to their agreement. A future trial with a larger sample size and longer follow-up time period (eg, 12 or 18 months) may provide a greater likelihood to assess any differences between couples who adhered to and did not adhere to their agreement, as had been found in a recent longitudinal study with male couples [84]. In addition, nonadherence to an SA may be defined differently between partners of the couple, which could influence how they might report about it. Recent research with male couples has found partner's reports on what components and behaviors their agreement included did not always align [17,78], which could in turn affect their understanding of the agreement and their report of adherence. As such, better measurements are needed to improve detection of agreement breaks by considering the different components (eg, emotional and sexual) of a couples' agreement.

Limitations

This pilot RCT has several limitations. A convenience sample was recruited by placement of targeted advertisements on Facebook, thereby limiting the generalizability of the study's findings as not all partnered men may use Facebook and those who do may not respond to advertisements about participating in HIV prevention or relationship research studies. Second, establishment and adherence to an SA were assessed by self-reporting. Social desirability bias may have influenced participants' responses to these survey items, thereby potentially affecting the study's outcome findings. The study also did not include serodiscordant and seroconcordant positive male couples or partnered transgender individuals (eg, transmen)—other populations who are in need of accessible, couples-based HIV/STI prevention interventions. A future iteration of the toolkit should include the biomedical (eg, Undetectable=Untransmissible and TasP), behavioral, and relational needs of serodiscordant and seroconcordant positive male couples [85] and transgender individuals and their relationship partners. Despite these limitations, findings from this pilot study showed promise for encouraging couples to establish and adhere to their SAs to warrant continuation of this research for HIV/STI prevention. A future trial of the updated toolkit with a larger sample size would provide sufficient power to detect effects and changes over time to assess whether establishing and adhering to an SA could enhance HIV/STI prevention efforts for male couples.

Conclusions

Our findings demonstrate strong evidence for the acceptability and feasibility of the eHealth toolkit as a brief, stand-alone, couples-based HIV/STI prevention intervention. These findings support the need to update the toolkit and evaluate it in a larger clinical trial powered for efficacy. Moreover, this intervention could be combined and/or supplemented with other couples-based HIV/STI prevention interventions such as CHTC to emphasize the importance of improving couple's relationship functioning—via agreements—for HIV/STI prevention. To date, most current and upcoming couples-based HIV/STI prevention interventions for male couples have focused on outcomes of HIV/STI testing, condom use, PrEP, and/or ART and less so on outcomes of SA formation and adherence. This intervention helps to fill this gap in couples-based HIV/STI prevention services for male couples.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screener items with accompanying decision rules used for couple verification test.
[DOC File , 229 KB - formative_v4i2e16807_app1.doc]

Multimedia Appendix 2
Description of relationship dynamics assessed.
[DOC File , 244 KB - formative_v4i2e16807_app2.doc]

Multimedia Appendix 3
Couples' investment in a sexual agreement, by trial arm and assessment time point.
[DOC File , 245 KB - formative_v4i2e16807_app3.doc]

Multimedia Appendix 4
Results from multilevel logistic regression to predict the odds ratio of a couple establishing a sexual agreement over time for trial arm and relative to their averaged relationship dynamic score.
[DOC File , 295 KB - formative_v4i2e16807_app4.doc]

Multimedia Appendix 5
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1527 KB - formative_v4i2e16807_app5.pdf]

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. Burton J, Darbes LA, Operario D. Couples-focused behavioral interventions for prevention of HIV: systematic review of the state of evidence. AIDS Behav 2010 Feb;14(1):1-10 [FREE Full text] [doi: 10.1007/s10461-008-9471-4] [Medline: 18843530]
4. Mashaphu S, Burns JK. Couples-based interventions in the context of HIV discordance. S Afr J Psychiatr 2017;23:1009 [FREE Full text] [doi: 10.4102/sajpsychiatry.v23i0.1009] [Medline: 30263183]
5. Purcell DW, Mizuno Y, Smith DK, Grabbe K, Courtenay-Quick C, Tomlinson H, et al. Incorporating couples-based approaches into HIV prevention for gay and bisexual men: opportunities and challenges. Arch Sex Behav 2014 Jan;43(1):35-46 [FREE Full text] [doi: 10.1007/s10508-013-0205-y] [Medline: 24233328]
6. Hoff CC, Campbell CK, Chakravarty D, Darbes LA. Relationship-based predictors of sexual risk for HIV among MSM couples: a systematic review of the literature. AIDS Behav 2016 Dec;20(12):2873-2892. [doi: 10.1007/s10461-016-1350-9] [Medline: 27048237]

XSL•FO
RenderX

7.   Jiwatram-Negrón T, El-Bassel N. Systematic review of couple-based HIV intervention and prevention studies: advantages, gaps, and future directions. AIDS Behav 2014 Oct;18(10):1864-1887 [FREE Full text] [doi: 10.1007/s10461-014-0827-7] [Medline: 24980246]

8.   Crepaz N, Tungol-Ashmon MV, Vosburgh HW, Baack BN, Mullins MM. Are couple-based interventions more effective than interventions delivered to individuals in promoting HIV protective behaviors? A meta-analysis. AIDS Care 2015;27(11):1361-1366 [FREE Full text] [doi: 10.1080/09540121.2015.1112353] [Medline: 26608175]

9.   Mendelsohn JB, Calzavara L, Daftary A, Mitra S, Pidutti J, Allman D, et al. A scoping review and thematic analysis of social and behavioural research among HIV-serodiscordant couples in high-income settings. BMC Public Health 2015 Mar 13;15:241 [FREE Full text] [doi: 10.1186/s12889-015-1488-9] [Medline: 25885027]

10.   El-Bassel N, Gilbert L, Witte S, Wu E, Hunt T, Remien RH. Couple-based HIV prevention in the United States: advantages, gaps, and future directions. J Acquir Immune Defic Syndr 2010 Dec;55(Suppl 2):S98-101 [FREE Full text] [doi: 10.1097/QAI.0b013e3181fbf407] [Medline: 21406997]

11.   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]

12.   Lewis MA, Gladstone E, Schmal S, Darbes LA. Health-related social control and relationship interdependence among gay couples. Health Educ Res 2006 Aug;21(4):488-500. [doi: 10.1093/her/cyh075] [Medline: 16459342]

13.   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]

14.   Goldenberg T, Stephenson R. 'The more support you have the better': partner support and dyadic HIV care across the continuum for gay and bisexual men. J Acquir Immune Defic Syndr 2015 May 1;69(Suppl 1):S73-S79 [FREE Full text] [doi: 10.1097/QAI.0000000000000576] [Medline: 25867781]

15.   Berg CA, Upchurch R. A developmental-contextual model of couples coping with chronic illness across the adult life span. Psychol Bull 2007 Nov;133(6):920-954. [doi: 10.1037/0033-2909.133.6.920] [Medline: 17967089]

16.   Gamarel KE, Revenson TA. Dyadic adaptation to chronic illness: the importance of considering context in understanding couples' resilience. In: Skerrett K, Fergus K, editors. Couple Resilience. Dordrecht: Springer; 2015:83-105.

17.   Rios-Spicer R, Darbes L, Hoff C, Sullivan PS, Stephenson R. Sexual agreements: a scoping review of measurement, prevalence and links to health outcomes. AIDS Behav 2019 Jan;23(1):259-271. [doi: 10.1007/s10461-018-2212-4] [Medline: 29959719]

18.   Mitchell JW. Characteristics and allowed behaviors of gay male couples' sexual agreements. J Sex Res 2014;51(3):316-328 [FREE Full text] [doi: 10.1080/00224499.2012.727915] [Medline: 23514544]

19.   Hoff CC, Beougher SC. Sexual agreements among gay male couples. Arch Sex Behav 2010 Jun;39(3):774-787 [FREE Full text] [doi: 10.1007/s10508-008-9393-2] [Medline: 18686027]

20.   Mitchell JW, Lee J, Woodyatt C, Bauermeister J, Sullivan P, Stephenson R. Perceived challenges and rewards of forming a sexual agreement among HIV-negative male couples. Arch Sex Behav 2016 Aug;45(6):1525-1534 [FREE Full text] [doi: 10.1007/s10508-016-0701-y] [Medline: 26964794]

21.   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]

22.   Mitchell JW, Lee J, Woodyatt C, Bauermeister J, Sullivan P, Stephenson R. Illuminating the context and circumstances of male couples establishing a sexual agreement in their relationship. Am J Mens Health 2017 May;11(3):600-609 [FREE Full text] [doi: 10.1177/1557988316655528] [Medline: 27334670]

23.   Hoff CC, Chakravarty D, Darbes L, Neilands TB. Studying the motivations behind sexual agreements: a first look at the Motivations Behind Agreement (MBA) scale for male couples. J Sex Res 2019;56(6):718-727. [doi: 10.1080/00224499.2018.1489490] [Medline: 30028213]

24.   Neilands TB, Chakravarty D, Darbes LA, Beougher SC, Hoff CC. Development and validation of the sexual agreement investment scale. J Sex Res 2010 Jan;47(1):24-37 [FREE Full text] [doi: 10.1080/00224490902916017] [Medline: 19396645]

25.   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]

26.   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]

27.   Mitchell JW. Between and within couple-level factors associated with gay male couples' investment in a sexual agreement. AIDS Behav 2014 Aug;18(8):1454-1465 [FREE Full text] [doi: 10.1007/s10461-013-0673-z] [Medline: 24327185]

28.   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]

XSL•FO

RenderX

29.   Parsons JT, Starks TJ, DuBois S, Grov C, Golub SA. Alternatives to monogamy among gay male couples in a community survey: implications for mental health and sexual risk. Arch Sex Behav 2013 Feb;42(2):303-312 [FREE Full text] [doi: 10.1007/s10508-011-9885-3] [Medline: 22187028]

30.   Parsons JT, Starks TJ, Gamarel KE, Grov C. Non-monogamy and sexual relationship quality among same-sex male couples. J Fam Psychol 2012 Oct;26(5):669-677. [doi: 10.1037/a0029561] [Medline: 22906124]

31.   Mitchell JW. Aspects of gay male couples' sexual agreements vary by their relationship length. AIDS Care 2014;26(9):1164-1170 [FREE Full text] [doi: 10.1080/09540121.2014.882491] [Medline: 24512593]

32.   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 2013 Oct;42(7):1337-1345 [FREE Full text] [doi: 10.1007/s10508-012-9985-8] [Medline: 22875716]

33.   Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]

34.   Mitchell JW. Gay male couples' attitudes toward using couples-based voluntary HIV counseling and testing. Arch Sex Behav 2014 Jan;43(1):161-171 [FREE Full text] [doi: 10.1007/s10508-013-0211-0] [Medline: 24213623]

35.   Rendina HJ, Breslow AS, Grov C, Ventuneac A, Starks TJ, Parsons JT. Interest in couples-based voluntary HIV counseling and testing in a national U.S. sample of gay and bisexual men: the role of demographic and HIV risk factors. Arch Sex Behav 2014 Jan;43(1):149-159 [FREE Full text] [doi: 10.1007/s10508-013-0226-6] [Medline: 24277109]

36.   Starks TJ, Doyle KM, Shalhav O, John SA, Parsons JT. An examination of gay couples' motivations to use (or forego) pre-exposure prophylaxis expressed during couples HIV testing and counseling (CHTC) sessions. Prev Sci 2019 Jan;20(1):157-167. [doi: 10.1007/s11121-018-0892-7] [Medline: 29651646]

37.   Mitchell JW, Lee J, Woodyatt C, Bauermeister J, Sullivan P, Stephenson R. HIV-negative male couples' attitudes about pre-exposure prophylaxis (PrEP) and using PrEP with a sexual agreement. AIDS Care 2016 Aug;28(8):994-999 [FREE Full text] [doi: 10.1080/09540121.2016.1168911] [Medline: 27055001]

38.   Malone J, Syvertsen JL, Johnson BE, Mimiaga MJ, Mayer KH, Bazzi AR. Negotiating sexual safety in the era of biomedical HIV prevention: relationship dynamics among male couples using pre-exposure prophylaxis. Cult Health Sex 2018 Jun;20(6):658-672 [FREE Full text] [doi: 10.1080/13691058.2017.1368711] [Medline: 28872441]

39.   Eisinger RW, Dieffenbach CW, Fauci AS. HIV viral load and transmissibility of HIV infection: undetectable equals untransmittable. J Am Med Assoc 2019 Feb 5;321(5):451-452. [doi: 10.1001/jama.2018.21167] [Medline: 30629090]

40.   Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, HPTN 052 Study Team. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011 Aug 11;365(6):493-505 [FREE Full text] [doi: 10.1056/NEJMoa1105243] [Medline: 21767103]

41.   Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, HPTN 052 Study Team. Antiretroviral therapy for the prevention of HIV-1 transmission. N Engl J Med 2016 Sep 1;375(9):830-839 [FREE Full text] [doi: 10.1056/NEJMoa1600693] [Medline: 27424812]

42.   Rodger A, Cambiano V, Brunn T, Vernazza P, Collins S, Corbelli G, PARTNER Study Group. Risk of HIV Transmission Through Condomless Sex in MSM Couples With Suppressive Art: the Partner2 Study Extended Results in Gay Men. In: Proceedings of the 22nd International AIDS Conference. 2018 Presented at: AIDS'18; July 23-27, 2018; Amsterdam.

43.   John SA, Starks TJ, Rendina HJ, Grov C, Parsons JT. Should I convince my partner to go on pre-exposure prophylaxis (PrEP)? The role of personal and relationship factors on PrEP-related social control among gay and bisexual men. AIDS Behav 2018 Apr;22(4):1239-1252 [FREE Full text] [doi: 10.1007/s10461-017-1835-1] [Medline: 28634660]

44.   Mimiaga MJ, Closson EF, Kothary V, Mitty JA. Sexual partnerships and considerations for HIV antiretroviral pre-exposure prophylaxis utilization among high-risk substance using men who have sex with men. Arch Sex Behav 2014 Jan;43(1):99-106 [FREE Full text] [doi: 10.1007/s10508-013-0208-8] [Medline: 24243002]

45.   Gamarel KE, Golub SA. Intimacy motivations and pre-exposure prophylaxis (PrEP) adoption intentions among HIV-negative men who have sex with men (MSM) in romantic relationships. Ann Behav Med 2015 Apr;49(2):177-186 [FREE Full text] [doi: 10.1007/s12160-014-9646-3] [Medline: 25124457]

46.   Underhill K. Intimacy, condom use, and pre-exposure prophylaxis (PReP) acceptability among men who have sex with men (MSM) in primary partnerships: a comment on Gamarel and Golub. Ann Behav Med 2015 Apr;49(2):151-153 [FREE Full text] [doi: 10.1007/s12160-014-9651-6] [Medline: 25245137]

47.   Mitchell JW, Stephenson R. HIV-negative partnered men's willingness to use pre-exposure prophylaxis and associated factors among an internet sample of US HIV-negative and HIV-discordant male couples. LGBT Health 2015 Mar;2(1):35-40 [FREE Full text] [doi: 10.1089/lgbt.2014.0092] [Medline: 26790016]

48.   Hoff CC, Chakravarty D, Bircher AE, Campbell CK, Grisham K, Neilands TB, et al. Attitudes towards PrEP and anticipated condom use among concordant HIV-negative and HIV-discordant male couples. AIDS Patient Care STDS 2015 Jul;29(7):408-417 [FREE Full text] [doi: 10.1089/apc.2014.0315] [Medline: 26057304]

49.   Brooks RA, Nieto O, Landrian A, Donohoe TJ. Persistent stigmatizing and negative perceptions of pre-exposure prophylaxis (PrEP) users: implications for PrEP adoption among Latino men who have sex with men. AIDS Care 2019 Apr;31(4):427-435 [FREE Full text] [doi: 10.1080/09540121.2018.1499864] [Medline: 30021456]

50. 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]

51. Gamarel KE, Darbes LA, Hightow-Weidman L, Sullivan P, Stephenson R. The development and testing of a relationship skills intervention to improve HIV prevention uptake among young gay, bisexual, and other men who have sex with men and their primary partners (We Prevent): Protocol for a randomized controlled trial. JMIR Res Protoc 2019 Jan 2;8(1):e10370 [FREE Full text] [doi: 10.2196/10370] [Medline: 30602433]

52. Newcomb ME, Macapagal KR, Feinstein BA, Bettin E, Swann G, Whitton SW. Integrating HIV prevention and relationship education for young same-sex male couples: a pilot trial of the 2GETHER intervention. AIDS Behav 2017 Aug;21(8):2464-2478 [FREE Full text] [doi: 10.1007/s10461-017-1674-0] [Medline: 28083833]

53. Starks TJ, Dellucci TV, Gupta S, Robles G, Stephenson R, Sullivan PS, et al. A pilot randomized trial of intervention components addressing drug use in couples HIV testing and counseling (CHTC) with male couples. AIDS Behav 2019 Sep;23(9):2407-2420. [doi: 10.1007/s10461-019-02455-2] [Medline: 30887190]

54. Stephenson R, Freeland R, Sullivan SP, Riley E, Johnson BA, Mitchell J, et al. Home-based HIV testing and counseling for male couples (Project Nexus): a protocol for a randomized controlled trial. JMIR Res Protoc 2017 May 30;6(5):e101 [FREE Full text] [doi: 10.2196/resprot.7341] [Medline: 28559225]

55. Stephenson R, Suarez NA, Garofalo R, Hidalgo MA, Hoehnle S, Thai J, et al. 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 Aug 31;6(8):e170 [FREE Full text] [doi: 10.2196/resprot.7884] [Medline: 28860107]

56. 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]

57. Martinez O, Wu E, Frasca T, Shultz AZ, Fernandez MI, Rios JL, et al. Adaptation of a couple-based HIV/STI prevention intervention for Latino men who have sex with men in New York City. Am J Mens Health 2017 Mar;11(2):181-195 [FREE Full text] [doi: 10.1177/1557988315579195] [Medline: 25846772]

58. Martinez O, Fernandez MI, Wu E, Carballo-Diéguez A, Prado G, Davey A, et al. A couple-based HIV prevention intervention for Latino men who have sex with men: study protocol for a randomized controlled trial. Trials 2018 Apr 5;19(1):218 [FREE Full text] [doi: 10.1186/s13063-018-2582-y] [Medline: 29622045]

59. Wu E, El-Bassel N, McVinney LD, Hess L, Remien RH, Charania M, et al. Feasibility and promise of a couple-based HIV/STI preventive intervention for methamphetamine-using, black men who have sex with men. AIDS Behav 2011 Nov;15(8):1745-1754 [FREE Full text] [doi: 10.1007/s10461-011-9997-8] [Medline: 21766193]

60. Nguyen LH, Tran BX, Rocha LE, Nguyen HL, Yang C, Latkin CA, et al. A systematic review of eHealth interventions addressing HIV/STI prevention among men who have sex with men. AIDS Behav 2019 Sep;23(9):2253-2272 [FREE Full text] [doi: 10.1007/s10461-019-02626-1] [Medline: 31401741]

61. Sullivan PS, Carballo-Diéguez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, et al. Successes and challenges of HIV prevention in men who have sex with men. Lancet 2012 Jul 28;380(9839):388-399 [FREE Full text] [doi: 10.1016/S0140-6736(12)60955-6] [Medline: 22819659]

62. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: a systematic review. J Med Internet Res 2014 May 26;16(5):e134 [FREE Full text] [doi: 10.2196/jmir.3393] [Medline: 24862459]

63. 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]

64. Mitchell JW. The use of technology to advance HIV prevention for couples. Curr HIV/AIDS Rep 2015 Dec;12(4):516-522 [FREE Full text] [doi: 10.1007/s11904-015-0290-8] [Medline: 26412083]

65. Larzelere RE, Huston TL. The dyadic trust scale: toward understanding interpersonal trust in close relationships. J Marriage Fam 1980 Aug;42(3):595-604. [doi: 10.2307/351903]

66. Rodrigues D, Lopes D. The investment model scale (IMS): Further studies on construct validation and development of a shorter version (IMS-S). J Gen Psychol 2013;140(1):16-28. [doi: 10.1080/00221309.2012.710276] [Medline: 24837343]

67. Hendrick SS. A generic measure of relationship satisfaction. J Marriage Fam 1988 Feb;50(1):93-98. [doi: 10.2307/352430]

68. Stulhofer A, Busko V, Brouillard P. Development and bicultural validation of the new sexual satisfaction scale. J Sex Res 2010 Jul;47(4):257-268. [doi: 10.1080/00224490903100561] [Medline: 19629836]

69. Miller RS, Lefcourt HM. The assessment of social intimacy. J Pers Assess 1982 Oct;46(5):514-518. [doi: 10.1207/s15327752jpa4605_12] [Medline: 16367633]

70. Futris TG, Campbell K, Nielsen RB, Burwell SR. The communication patterns questionnaire-short form: a review and assessment. Fam J 2010;18(3):275-287. [doi: 10.1177/1066480710370758]

71. 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]

72. 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]

73. Herek GM, Glunt EK. Identity and community among gay and bisexual men in the AIDS era: Preliminary findings from the Sacremento Men's Health Study. In: Herek GM, Greene BA, editors. AIDS, Identity, and Community: The HIV Epidemic and Lesbians and Gay Men. Newbury Park: Sage; 1995:55-84.

74. Wagner G, Brondolo E, Rabkin J. Internalized homophobia in a sample of HIV+ gay men, and its relationship to psychological distress, coping, and illness progression. J Homosex 1996;32(2):91-106. [doi: 10.1300/j082v32n02_06] [Medline: 9010828]

75. Mitchell J, Lammert S, Danh, Horvath K. Optimizing enrollment methods: lessons learned from an mHealth, pilot HIV prevention randomized controlled trial with sexual minority men. Int J STD AIDS 2019 Jun;30(7):703-706. [doi: 10.1177/0956462419831854] [Medline: 30890120]

76. Pew Research Center. 2019 Jun 13. Mobile Technology and Home Broadband 2019 URL: https://www.pewinternet.org/2019/06/13/mobile-technology-and-home-broadband-2019/ [accessed 2020-01-03]

77. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: Validation study. JMIR Mhealth Uhealth 2018 Jan 5;6(1):e4 [FREE Full text] [doi: 10.2196/mhealth.8851] [Medline: 29305343]

78. Sharma A, Garofalo R, Hidalgo MA, Hoehnle S, Mimiaga MJ, Brown E, et al. Do male couples agree on their sexual agreements? An analysis of dyadic data. Arch Sex Behav 2019 May;48(4):1203-1216. [doi: 10.1007/s10508-019-1391-z] [Medline: 30895490]

79. 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]

80. Stachowski C, Stephenson R. Homophobia and communal coping for HIV risk management among gay men in relationships. Arch Sex Behav 2015 Feb;44(2):467-474 [FREE Full text] [doi: 10.1007/s10508-014-0417-9] [Medline: 25614049]

81. Mustanski B, Newcomb ME, Clerkin EM. Relationship characteristics and sexual risk-taking in young men who have sex with men. Health Psychol 2011 Sep;30(5):597-605 [FREE Full text] [doi: 10.1037/a0023858] [Medline: 21604883]

82. Fergus S, Lewis MA, Darbes LA, Butterfield RM. HIV risk and protection among gay male couples: the role of gay community integration. Health Educ Behav 2005 Apr;32(2):151-171. [doi: 10.1177/1090198104271964] [Medline: 15856614]

83. Feinstein BA, Bettin E, Swann G, Macapagal K, Whitton SW, Newcomb ME. The influence of internalized stigma on the efficacy of an HIV prevention and relationship education program for young male couples. AIDS Behav 2018 Dec;22(12):3847-3858 [FREE Full text] [doi: 10.1007/s10461-018-2093-6] [Medline: 29569000]

84. Darbes LA, Chakravarty D, Neilands TB, Beougher SC, Hoff CC. Sexual risk for HIV among gay male couples: a longitudinal study of the impact of relationship dynamics. Arch Sex Behav 2014 Jan;43(1):47-60 [FREE Full text] [doi: 10.1007/s10508-013-0206-x] [Medline: 24233329]

85. Mitchell JW, Lee J, Godoy F, Asmar L, Perez G. HIV-discordant and concordant HIV-positive male couples' recommendations for how an eHealth HIV prevention toolkit for concordant HIV-negative male couples could be improved to meet their specific needs. AIDS Care 2018 Jun;30(sup2):54-60 [FREE Full text] [doi: 10.1080/09540121.2018.1465527] [Medline: 29848043]

Abbreviations

ART: antiretroviral treatment
CAS: condomless anal sex
CHTC: couple's HIV testing and counseling
CIT: couples interdependence theory
eHealth: electronic health
GBMSM: gay, bisexual, and other men who have sex with men
IP: Internet Protocol
LGBTQ: lesbian, gay, bisexual, transgender, and questioning/queer
mHealth: mobile health
OR: odds ratio
PrEP: pre-exposure prophylaxis
RCT: randomized controlled trial
SA: sexual agreement
SAIS: Sexual Agreement Investment Scale
STI: sexually transmitted infection
TasP: treatment as prevention

Original Paper

A Weekly, Evidence-Based Health Letter for Caregivers (90Second Caregiver): Usability Study

Athena Milios[1,2], BSc; Patrick McGrath[1,2], PhD; Hannah Baillie[2]

[1]Department of Psychiatry, Dalhousie University, Halifax, NS, Canada
[2]Centre for Research in Family Health, IWK Health Centre, Halifax, NS, Canada

Corresponding Author:
Athena Milios, BSc
Centre for Research in Family Health, IWK Health Centre
5980 University Ave
Halifax, NS
Canada
Phone: 1 9025790269
Email: athena.milios1@gmail.com

Abstract

Background: Informal caregivers are family members or close friends who provide unpaid help to individuals with acute or chronic health conditions so that they can manage daily life tasks. The greatest source of health information is the internet for meeting the needs of caregivers. However, information on the internet may not be scientifically valid, it may be written in language that is difficult to read, and is often in very large doses. 90Second Caregiver is a health letter whose aim is to disseminate knowledge to caregivers in a user-friendly, weekly format, in order to improve their wellbeing.

Objective: The main objective was to test a sample of 90Second Caregiver health letters in order to assess their usability and to optimize the design and content of the health letters.

Methods: Usability research themes were assessed using semi-structured phone interviews, incorporating the Think Aloud method with retrospective questioning.

Results: Usability was assessed in the context of five main themes: understandability and learnability, completeness, relevance, and quality and credibility of the health letter content, as well as design and format. Caregivers generally provided positive feedback regarding the usability of the letters. The usability feedback was used to refine 90Second Caregiver in order to improve the design and content of the series. Based on the results of this study, it may be of maximum benefit to target the series towards individuals who are new to caregiving or part-time caregivers, given that these caregivers of the sample found the letters more useful and relevant and had the most positive usability experiences.

Conclusions: The findings assisted in the improvement of the 90Second Caregiver template, which will be used to create future health letters and refine the letters that have already been created. The findings have implications for who the 90Second Caregiver series should be targeting (ie, newer or part-time caregivers) in order to be maximally impactful in improving mental health and wellbeing-related outcomes for caregivers, such as self-efficacy and caregiving knowledge. The results of this study may be generalizable to the examination of other electronic health information formats, making them valuable to future researchers testing the usability of health information products. In addition, the methods used in this study are useful for usability hypothesis generation. Lastly, our 90Second delivery approach can generate information useful for a set of similar products (eg, weekly health letters targeted towards other conditions/populations).

XSL•FO
RenderX

Introduction

Background and Rationale

Informal caregivers are family members or close friends who provide unpaid help to individuals with acute or chronic health conditions so that they can manage daily life tasks [1]. Although informal caregiving is better than paid caregiving for the mental and physical well-being of the individuals receiving the care, it can negatively affect the well-being of the caregivers themselves [1]. Caregivers have been found to have high levels of stress, depression, and risk for mortality. They are less likely to preventatively manage their own health, that is, to take care of themselves through engaging in healthy lifestyle behaviors such as exercise, healthy diet, and proper sleep hygiene [1,2]. A meta-analysis found that caregivers, compared with noncaregivers, had higher stress levels and depressive symptoms, lower self-efficacy, and poorer general subjective well-being [2,3]. Self-efficacy is the belief that one is capable and competent to manage situations. Self-efficacy in caregivers is associated with a lower risk of caregiver burnout and psychological distress and higher care-recipient well-being [4].

The greatest source of health information for caregivers is the internet. However, there are many barriers facing caregivers given the amount of scientific knowledge available to the public and the difficulty in interpreting dense (low readability) and lengthy (containing information in large doses) scientific articles [5,6]. Furthermore, much of the information on the internet is not evidence based or scientifically valid [5,6]. Caregivers of those with mental and physical health problems are very busy and stressed and need a trustworthy source of easy-to-read, concise, accurate, and evidence-based information presented in manageable portions. Exposure to inaccurate, misleading, outdated, or vague information could be detrimental to caregiver and patient health-related outcomes [6].

90Second Caregiver is a health letter that aims to disseminate knowledge to caregivers in a user-friendly, weekly format to improve their self-efficacy, increase their health-related caregiving knowledge (on a general level, with the goal of them being able to apply this general knowledge to disease-specific caregiving dilemmas), and promote healthy coping behaviors. The letters are all developed using credible scientific sources, such as academic journals and/or government agencies.

Usability research involves the participants using and evaluating a product or service, such as Web-based electronic health (eHealth) apps, websites, or health documents. Usability studies aim to detect usage-related difficulties and improve the design of health-related services and products [7].

Objectives

The main objective was to pilot test a sample of 90Second Caregiver health letters to obtain data regarding their usability and to optimize the delivery, design, and content utility of the health letters. Usability was assessed in terms of design and format, understandability and learnability, completeness, practical relevance, and quality and credibility of the health letter content. The secondary objectives were to assess factors that participants liked and disliked about the format and content, to determine which additional topics participants would like to see in the future, to assess their interest in becoming subscribers in the future, and to determine what improvements and changes the caregivers would like to see regarding the design and format or content of the health letters.

Methods

Recruitment

After the research ethics board's approval was obtained, participants eligible for the study were recruited through our health center's volunteer service and the Brain Injury Association of Nova Scotia. A sample size calculation was not performed because most usability problems are discovered by the first 5 participants [8].

The exclusion criteria were as follows: (1) individuals who were not primary caregivers, (2) participants who did not speak English, (3) participants who did not have access to a computer or email or internet, and (4) participants who did not have a phone. The 90Second health letters were delivered by an email link to a Web page.

The first page of the letters had 3 components: 200 words of evidence-based, plain-language health information on a focused issue related to caregiving (the main body), 100 to 150 words of actionable suggestions, and a license-free graphic relating to the topic that supports the main message of each letter (located above the main body). The second page contained a 7-item assessment tool. Each item was rated on a scale from 1 to 5 to generate a total score. There was also an explanation of the total score. The third page was a personal account of a caregiver on the issue (up to 300 words). A 3-item rating of the health letter followed. Finally, links to additional resources for caregivers were provided.

The questions in the 7-item assessment tool capture a single construct (either behavior or attitudes, not simply factual knowledge) based on the central concept of each health letter. The 7 questions can be answered on a scale from 1 (false) to 5 (true) to generate a total score ranging from 7 to 35. Some items can be reverse scored. Items must measure a single construct, should be short (less than 10 words), and not have long or unusual words. Double negatives are not permitted. For example, "I am proud of my role and abilities as a caregiver" is an item in the stigma health letter assessment tool.

The *Rate our Health Letter* scale contains 3 questions/items: "Did you find this health letter helpful?," "Could you relate to the content of this health letter?," and "Would you recommend this letter to a friend or organization?" The response options are yes, not applicable, and no. Then, the reader is asked if they have suggestions for future topics.

Participant Characteristics

The sample included 10 caregivers who evaluated 2 health letters each. The first 5 caregivers evaluated one set of (2) health letters (anxiety and depression), and the second 5 caregivers evaluated another set of (2) health letters (hope and caregiver stigma).

The caregivers' demographic information is shown in Table 1. The population was a diverse sample of caregivers, caring for individuals with various mental and physical health conditions.

Table 1. Demographic characteristics of participants.

Participant	Age (years)	Sex	Length of time spent caregiving	Occupation	Highest level of education	Hours spent caregiving per week	Nature of loved one's health condition
SL	25	Female	6 months	Nurse	BSc[a]	3	Cancer
SK	52	Male	2 years	Professor	PhD[b]	20	Cancer (breast)
AB	24	Male	1 year	Student	BSc	3	Irritable bowel syndrome
RM	37	Female	3 years	Economist	MA[c]	60	Cancer (leukemia)
ME	43	Female	10 years	Community relations and television producer	MA	3	Epilepsy
KL	34	Female	3 years	Senior policy analyst	Community college diploma	3	Acquired brain injury (traumatic)
JM	57	Female	30 years	Accountant	Postgraduate diploma	25	15q duplication syndrome (neurodevelopmental disorder)
DM	53	Female	4 years	Nurse	Community college (nursing) diploma	60	Acquired brain injury
WM	61	Female	6 years	Administrative assistant	High school	80	Acquired brain injury (anoxic)
CM	77	Female	26 years	Research scientist	PhD	80	Acquired brain injury (traumatic)

[a]BSc: Bachelor of Science.

[b]PhD: Doctor of Philosophy.

[c]MA: Master of Arts.

Procedure

First, caregivers who had expressed an interest were sent an email with some background information regarding the study and what was involved. The information and consent form was attached to this email. Participants were asked to respond with the dates and times that they would be available for the phone interview.

At the beginning of the phone interview, informed consent was obtained. Participants were then asked questions, regarding (1) age and sex, (2) length of time as a caregiver, (3) current occupation, (4) level of education completed, (5) the nature of their care recipient's health condition, and (6) length of time spent caregiving per week (Table 1).

Participants were emailed PDF copies of the two 90Second Caregiver health letters that they were to assess and encouraged to vocalize their thought processes as they read them, which is known as the concurrent think aloud technique. The think aloud portion of the interview involved the interviewer recording the report of the thinking process as each participant reflected on the content and format of the sample of health letters. The participants were encouraged to express their understanding, beliefs, attitudes, and expectations regarding the sample of health letters that they were reading. The think aloud method has been shown to be the most effective in detecting usability problems, which is why it was chosen for this study [9].

After they read each health letter, the participants were asked a series of questions to assess its design, format, and content to expand and complement the think aloud results (Textbox 1) [10]. The caregivers' opinions regarding various aspects of the design and content, including ease of use, ease of learning, completeness, practical relevance, usefulness, quality, and credibility were elicited [11,12].

Textbox 1. Semistructured interview questions. (Caregivers were asked questions 1-9 twice, once for each health letter they read. Then, they were asked questions 10-15 after both the letters were read.)

1. What is your impression regarding the purpose of the health letter?

2. What is the first section you would read to get started?

3. How do you feel about the way the information was presented/formatted? (Is the text big enough, do you like the pictures, the order of the material, etc) [13].

4. How do you feel about the way the information is written? (Writing level, understanding, readability, and unfamiliar terms) [14].

5. Could we do anything to make this topic easier or more enjoyable to read (illustrations, more explanations, short video clips, etc.)?

6. Was the amount of information included enough/not enough, in other words, how complete did you feel each health letter was in covering the topic?
   i.   Was there anything you found yourself wanting to know that was not included?

   ii. Was there any part of the health letter that you thought was unnecessary or should be removed?

   iii. How did you find the length of each health letter? Would you prefer it longer or shorter?

7. How useful and relevant did you find the information?

8. What knowledge have you learned/gained? What did you already know?

9. What did you like/dislike about the content of each health letter [15]?
   i.   The main body of text?

   ii. The tips/suggestions?

   iii. The self-assessment?

   iv. The personal account?

10. How satisfied are you with the overall quality and reliability of what you read?

11. What was the best part of the health letters? The worst?

12. If you could change anything (about either the design/format or content), what would it be?

13. What additional topics would you like to see if you were a subscriber to the series?

14. Do you have any suggestions for specific companies, associations, or agencies that you would like to see sponsoring the 90Second Caregiver Series?

15. Would you be interested in subscribing to this series in the future?

The questions were developed by performing a literature search regarding the most important aspects of usability [7]. For example, Lund (2001) found that ease of use, understandability, learnability, usefulness, and overall satisfaction are the most important elements of usability, which is why these themes were incorporated into the question design for this study [11]. The larger list of questions was then narrowed down to a more concise set of questions to eliminate redundancies and ambiguities and reduce the number of questions. Face and content validities of the interview questions were reviewed by a set of psychology researchers to ensure the various domains of usability were adequately covered and to ensure the questions were clear and practical [16].

The think aloud technique was also used to assess how participants answered the assessment questions in each health letter. Caregivers provided feedback on whether the questions in the assessment captured the concept that they were supposed to measure. Hope, caregiver stigma, depression, and anxiety were the topics of the 4 health letters used.

After they completed the phone interview, the participants were sent a thank-you email, with their signed information and consent form and their Can $10 Amazon gift card.

Data Analysis

The data were deidentified (names and contact information removed) and transcribed. Thematic analysis was performed to analyze the transcripts: frequencies of emerging usability categories and themes from the concurrent think aloud data and the retrospective semistructured usability questions were analyzed to assess the outcome measures of the study (ie, caregiver satisfaction and opinions regarding the main themes of usability—understandability and learnability, completeness, relevance, usefulness, and quality and credibility of the health letter content—as well as design and format) [10].

Usability problems and improvements suggested by participants were also recorded, and their frequencies were analyzed. Modifications were made to 90Second Caregiver based on the participant feedback (see Tables 2 and 3). Combining think aloud data with retrospective questionnaire data is the most complete way to understand the usability experiences of participants [17-19].

Table 2. Summary of the usability themes of the Anxiety and Depression health letters' content and feedback and changes made.

Content theme	Feedback and changes	Comments by caregivers
Understandability and learnability	• The term "full-blown" when used to refer to anxiety and depressive disorders was removed. • The title Background Information was changed to Resources in the template.	"There were no unfamiliar terms to me. It was easy to read and comprehend and left me with no questions." [AB]
Completeness	• The main body of both health letters was made more concise, and more information was added about how to access resources while accounting for constraints that caregivers face (eg, time constraints and financial). The main body was shortened so that it fit entirely on the first page of each letter.	"I would like to know more about resources, such as group therapy that might be at a better rate, or maybe something online that is free?" [SL]
Relevance	• Participants all found the content of both the anxiety and depression health letters very relevant. SL liked the focus on self-care in the depression letter. She found the assessment statements in both letters very relevant as well.	"I liked that it was relevant to the broader caregiving community." [RM]
Usefulness	• "Recognize your boundaries" was added to the SMART[a] Tips section in the Depression health letter. Information about making social media connections with other caregivers was added to the main body, as suggested by RM.	"I found the letter very useful because caregiver anxiety is so common and so overlooked; it is nice to have some resources." [SL]
Quality and credibility	• The average response was *very satisfied* when caregivers were asked about the overall quality and reliability of the health letters. • The references increased the quality and reliability of the letters [AB].	"I found the background information increased the credibility of the facts. I liked that the links were relevant to papers published in recent years." [AB]

[a]SMART: specific, measurable, attainable, realistic, and time bound.

Table 3. Summary of the usability themes of the Keeping Hope and Stigma health letters' content and feedback and changes made.

Content theme	Feedback and changes	Comments by caregivers
Understandability and learnability	• The definition of stigma was clarified in the Stigma letter. • The terms in the bullet list on page 2 of the Keeping Hope letter were bolded to stand out more and be more learnable and memorable.	"I learned that I need to educate others and get my story out there more to reduce the stigma." [KL]
Completeness	• Overall, participants liked the length and completeness of the letters. • JM found the Stigma letter too negative. The letters were all edited to increase the use of positive, optimistic, and empowering statements, particularly in the SMART[a] Tips sections. • WM, CM, and JM wanted the point about connecting with a higher power, in the main body of the Keeping Hope letter, to be removed. This feedback was implemented.	"I found the letters provided a nice quick snippet of a little bit of information; even though they were short, there were some nice messages to take away." [KL]
Relevance	• JM felt that the letters should be made more targeted to specific types of caregivers so that they can be more relevant to peoples' needs such as her own, given her son's condition is quite rare. • In the SMART Tips section of the Keeping Hope letter, CM did not find the "Make a positives and negatives list" and the "Set short-term goals" points to be relevant to caregivers of individuals with acquired brain injury because she felt it would be easier to just cope with things as they come along instead of risking overthinking about the future.	"It was very relevant because it touches on areas of hope that would have been especially helpful when I was in the most burdensome part of my caregiving experience." [DM]; "The content should be made more relevant and targeted. How can you help someone who is in a completely different situation than somebody else?" [JM]
Usefulness	• Additional links were added to the end of the health letters under the Resources section of the Keeping Hope and Stigma letters to make them more useful for caregivers needing more information on a specific topic.	"It was useful overall, but I would like to see contacts for if a person needed help with something." [WM] (referring to the Keeping Hope letter)
Quality and credibility	• The average response was *very satisfied* when caregivers were asked about the overall quality and reliability of the health letters.	—[b]

[a]SMART: specific, measurable, attainable, realistic, and time bound.

[b]Not applicable.

XSL•FO
RenderX

Results

Content

All caregivers made positive comments about the health letter content (see Tables 2 and 3), including the main body (first page), the image, the specific, measurable, attainable, realistic, and time-bound (SMART) Tips, the assessment, the Personal Account, and the Resources. Some minor, content-specific suggestions regarding each section were made. For example, RM felt that making major life changes during a stressful and tumultuous caregiving period would not be realistic for most caregivers. She felt that the actionable suggestion in the Anxiety letter "Limit caffeine and alcohol" would be a poor suggestion. However, she also acknowledged that this would depend on the severity and nature of the illness of the care recipient.

JM and DM both did not like the first suggestion in the SMART Tips section of the Keeping Hope letter, "Accept your situation and your role in it." They found it too abrasive and confrontational.

In the Keeping Hope letter, JM and DM also did not like the last point in the SMART Tips section, "Identify and use your supports," and the last bullet point in the first page (main body) section "Connecting with your outside support system...engage in community activities." JM felt that these were not attainable or realistic because they are outside of the caregiver's control: "You can try to connect with them, but if they aren't there for you, they're not, and you should go find support in other places, where people can relate, and they're not scared of your situation." JM said "you can do that if you have time, and if you live in a place where there are community activities, but a lot of people don't. You need the time, the energy, the resources."

JM also did not like the suggestion of the Keeping Hope letter: "Accept the things you cannot change, such as the course of your loved one's illness." JM found it too simplistic, negative, and "bossy" because "there's some things you can change, but you don't know it until you have enough information. This is why it is critical to persist to inform yourself as much as possible. That means finding people who can help you." JM, WM, and CM did not like the suggestion of the Keeping Hope letter about connecting with a higher power because they felt that this would not be useful or relevant for many caregivers (Table 3).

AB, RM, and SK suggested modifying the font and colors of the figure in the Depression health letter and simplifying the figure's text to improve its readability. ME suggested changing the title of the Background Information section to Resources to increase its understandability.

SL suggested the signs and symptoms of anxiety be placed on the first page of the Anxiety health letter. DM suggested that caregiver stigma be defined more clearly in the Stigma health letter.

Overall, caregivers were very pleased with the completeness of the health letters. However, a common theme regarding the completeness of the letters was that caregivers wanted more information on solutions, such as more coping skills, treatments, and resources for anxiety and depression (SL, RM, AB, and ME). Participants felt that the health letters should overall be more actionable. For example, RM said "Don't tell people what they already know. Identify the issue and provide a solution!"

Furthermore, AB and SK suggested that both the Anxiety and the Depression letters deemphasize the role of antidepressant medications as a treatment. They did not want the health letter to make it seem like the first treatment option for anxiety or depression in caregivers is medication. In the list of treatment options, instead of antidepressants being listed first, talk therapies, *then,* caregiver support groups, and *finally,* antidepressant medications were listed.

SL suggested adding in more details to the Personal Account section about the name and age of each caregiver.

Participants found the health letter content quite useful. Although some caregivers did not find certain content relevant to their situations, they recognized that it might be relevant to caregivers more generally. For example, RM found the Anxiety letter content not relevant to her personally but stated that it would be relevant within the "broader caregiving community." In addition, in the SMART Tips section of the Depression letter, RM suggested adding a tip about boundaries.

To make the content of the assessment more relevant in the Anxiety letter, several participants suggested including more response options in the scale. Instead of having just 4 response options, participants suggested adding an option between *not at all* and *several days*. An in-between rating was also suggested for the "Rate our Newsletter" section, between the *yes* and *no* options (Textbox 2).

Of all of the caregivers, JM and CM were the least satisfied (they selected the *moderately satisfied* option) with the usefulness of the health letters they read (the Keeping Hope and Stigma letters) because they did not think we could make a health letter that is relevant and useful for all types of caregivers, given that caregiver experiences are so diverse: "Trying to produce newsletters for a one-size-fits-all is going to be tough" (CM). It is also relevant to note that JM and CM were the most burdened caregivers in the sample based on the length of time they had spent caregiving in years (see Table 1).

The average response for the quality and credibility of the health letter content was *very satisfied* (see Tables 2 and 3). The Resources section helped to increase the credibility of the health letter content. Participants appreciated seeing the additional references. Even if they indicated they would not actually use them, they found it important that they be included. KL reported that she was extremely satisfied with the letter quality. She found the health letters better than any of the other materials she had read before.

Textbox 2. Changes made to the design and format (ie, to the template) of the series.

- The font of the letters was changed to Arial (from Hoefler text; based on the suggestion by KL).

- The title, subtitle, and section headings were changed from blue and black to dark red (to increase consistency of the colors throughout the letter).

- The font of the titles was reduced to size 40 and bolded. The subtitle fonts were increased to size 18 and bolded (to reduce the size discrepancy between the title and subtitle and simplify the layout; based on the suggestions by CM, KL, AB, and WM).

- The Call to Action title was changed to SMART (specific, measurable, attainable, realistic, and time bound) Tips, bolded, and placed in all capital letters. Its font was increased to size 16.

- Each suggestion in the SMART Tips sections was bolded and increased to size 14. The black text explaining each suggestion was changed to font size 11 (not bolded).

- The entire letter template was changed to single spacing.

- The main body of text was made more concise, so that it all fit onto the first page.

- The title of the Background Information section was changed to Resources.

- The Rate Our Newsletter section title was changed to Rate Our Health Letter.

- Not applicable (N/A) was added as an option between yes and no in the Rate Our Health Letter section.

- The yes, N/A, and no response options in the Rate Our Health Letter section were adjusted so that they lined up for each of the 3 questions in this section.

- The first name and age of the caregiver were added in to each Personal Account (based on the suggestion by SL).

Design and Format

Participants generally provided positive feedback about the design and the layout. Participants were satisfied with the length of the letters; 3 pages was the optimal length suggested by the majority of participants.

CM and KL suggested the design and layout of the health letters be simplified (Textbox 2). CM suggested the fonts be more consistent on the first page. This was in reference to the size discrepancy between the large red "90SECOND CAREGIVER" title and the small blue subtitle beneath it. AB, WM, CM, and DM suggested that the "90SECOND CAREGIVER" title be bolded, that the blue subheading beneath the title be a larger font, and that the SMART Tips section (on the first page) be more actionable. KL suggested that the font of the health letters be increased and changed to a more readable font, such as Arial (Textbox 2).

CM and WM felt that the SMART Tips should be on the left-hand side of the page at the beginning and that the main body should be underneath. On the contrary, SK suggested that the SMART Tips section be moved more to the right to make it clearer to the reader that the main body section should be read first.

RM, ME, and WM found that the SMART Tips provided a good summary of the content of each health letter, which was helpful because the reader could read this first to decide if they want to read the entire letter (they may not want to read it all if they do not find it relevant for their needs). ME also felt that having more of the main body section content in point form would make the material more readable, increasing its learnability.

All changes to the template of the series (as opposed to minor, content-specific changes) were implemented only upon consultation and agreement among the principal investigator, the editor in chief, and the principal scientist of the series (Textbox 2).

Likes and Dislikes (Secondary Objective)

Regarding the format and content, participants did not like that only 4 response options were present in the rating scale for the Anxiety health letter. Overall, participants liked that each main body section was comprehensive, providing a good overview of each topic and referring to real-world statistics. However, some participants mentioned they would like the main body section to be more oriented toward solutions to anxiety and depression rather than explaining the problem itself. Participants also enjoyed reading the Personal Account section. RM stated, "it is always interesting to hear other peoples' stories." Participants appreciated the fact that references were included because this increased the overall quality and credibility of the health letters.

RM did not like the suggestions in the Depression letter (relating to sleep, social activities, and self-care) because she thought they would be too unrealistic and unattainable for caregivers in situations such as her own, caring for an acutely and severely ill child with extended, frequent hospital stays. However, RM was one of the most burdened caregivers in the sample (60 hours of caregiving per week), so her situation may not be generalizable to most caregivers.

Overall, the SMART Tips and the Personal Account were the sections of the letters that participants liked the most. KL stated that the Personal Account "validates how people are feeling and what they are dealing with."

Regarding the main body section of the Keeping Hope letter, JM said: "it is burdensome for caregivers to be reading such negative content. There are too many negative elements that aren't helpful." She also did not like the assessments, because she did not feel they could add anything in terms of addressing caregiver burden and the burden of stigma (particularly in the Stigma health letter). However, JM enjoyed the Personal Account section and the SMART Tips, particularly the location of the SMART Tips within the health letters, and its clarity. She

also had quite a few specific suggestions to help improve each SMART Tips point in both health letters that she read. For example, she felt that the first point in the SMART Tips section of the Stigma letter "Educate others" might not be specific enough or attainable because she believes that most people are not open and willing to listen, and this advice would be more difficult for introverted caregivers to follow. She also felt the second point in the SMART Tips of the Stigma letter, "Don't be Shy...Make sure to bring your loved one to gatherings" was too simplistic and unrealistic (ie, there are too many barriers to caregivers actually implementing this advice). She felt that

telling a caregiver who is introverted "don't be shy" is like telling someone with depression to just "get over it."

JM and DM both did not like the graphic in the Keeping Hope letter because they felt it was too dark and pessimistic.

Additional Topics Suggested (Secondary Objective)

Textbox 3 lists all the additional topics that participants mentioned that they would like to see if they were subscribers to the series in the future. Furthermore, all the caregivers, except for 1 (CM), expressed an interest in becoming subscribers to the 90Second Caregiver series in the future.

Textbox 3. Additional topics suggested.

- Caregivers of the sandwich generation
- Caregiving and finances
- Caregiving for children
- Perfectionism in caregivers
- Insomnia in caregivers
- Faith in caregivers
- Social isolation in caregivers
- Caregiver support groups or the importance of a support system in caregivers
- Caregiver burnout
- How to actively participate in your loved one's health care
- How to balance caregiving with being a parent
- How to deal with setbacks or relapse in recovery
- How to empower and educate yourself as a caregiver
- How to engage your loved one (the care recipient) in their health care
- Finding resources for help as a caregiver
- How to handle negativity as a caregiver, especially in the care recipient
- Supporting an adult survivor of acquired brain injury, for example, with independence, romantic relationships, and workplace discrimination

Discussion

Principal Findings

The main outcome of the study was to assess and consequently improve the usability of 90Second Caregiver based on caregiver input. Usability in this study was assessed in terms of the themes of understandability, learnability, completeness, practical relevance, usefulness, and quality and credibility of the health letter content, as well as design and format. Findings reinforced the fact that 90Second Caregiver is a very user-friendly, learnable, and useful series of health letters. Caregivers were very satisfied with both the content and design of the series. Participants found the reading level acceptable (eg, no unfamiliar terms), and they found the content provided an excellent summary of each health letter topic. The majority of participants found the information useful and relevant to their needs as caregivers, and they were satisfied with the content's credibility.

Some letter-specific changes as well as template changes were suggested and implemented based on the feedback data of participants. On the basis of participant feedback, the content

of the health letters will be changed to specify that the letters should (1) contain specific, measurable, attainable, realistic, and time-bound tips; (2) be objective, empowering, positive, and optimistic; (3) avoid overly *bossy* language or polarizing topics; (4) avoid suggesting changes that require significant financial investment or travel time; and (5) ensure that the main body section is solution oriented instead of explanation based. Textbox 2 provides a summary of the changes implemented to the template of the series based on participant feedback and the principles of persuasive design, such as making sure that the SMART tips are fitting suggestions that can be used successfully by caregivers and that the health letters are visually attractive in addition to being trustworthy [20].

A relationship was observed between the perceived usefulness of the letters and the burden of caregiving in this sample. Caregivers who were more burdened (RM, JM, and CM), based on the length of time they spent caregiving per week and/or the length of time in years that they had been caregiving, tended to be more critical of the health letter content, finding it less useful and relevant. For example, RM was one of the most burdened caregivers of the sample, spending 60 hours a week caregiving,

so her situation may not necessarily be the most representative of a typical caregiver or subscriber to 90Second Caregiver.

If this finding is confirmed with a larger sample, it may be best to target the current letters toward individuals who are relatively new to caregiving or part-time caregivers. In addition, a different approach may be needed for caregivers who are more burdened because the caregivers who were overly critical tended to find the content of the letters less useful and feasible. This new approach could be needed because these caregivers may simply be too burnt out from the demands of caregiving to truly appreciate, retain, and apply the health knowledge in the letters, or they may be so experienced in their caregiving role that they find the letters redundant. To better address the needs of these caregivers, a new approach could involve pairing the 90Second Caregiver letters with another distance-delivered intervention, such a Web-based weekly stress-management health letter.

Many participants commented that they appreciated being included in this project and provided with the opportunity to vocalize their feedback and sharing how it was related to their specific caregiving experiences.

Limitations and Strengths

One limitation of this study is that social desirability bias might have influenced the way that participants responded to the interview questions. In other words, participants may have overemphasized the positive features of the health letters and omitted certain usability issues or factors that they disliked about the letters to please and impress the interviewer [21]. This potential bias was minimized by verbalizing to participants to not worry about *hurting the investigator's feelings* and encouraging them to give their honest thoughts during the interview. Confirmation bias, the tendency to interpret data in a way that confirms the researchers' pre-existing beliefs about the usability of the health letters, might have also been a risk to the validity of the results [21]. This potential bias was addressed by having a second researcher review the usability themes and categories that emerged from the thematic analysis of the interview transcripts.

Another limitation was that not all the changes proposed by the caregivers were able to be implemented because of inconsistencies in the feedback provided [13]. A larger sample size may have clarified some of these inconsistencies, but it was not possible to implement it in this study because of time constraints.

Another limitation was that more than half of the participants had a university qualification and/or worked in a health-related field. This could explain why some of the participants were overly critical of both the design and the content of the letters.

A significant strength of this study was that triangulation (ie, both the concurrent think aloud method and retrospective questioning) were used to assess caregivers' usability experiences [16,22]. Triangulation is a means to assess outcomes from multiple perspectives to ensure that the usability findings are reliable (ie, consistent, reproducible, and repeatable) and valid (trustworthy, credible, and accurate). Another strength was that caregivers in our sample represented a broad range of ethnic and sociodemographic groups, with a variety of caregiving backgrounds and experiences.

Conclusions

In conclusion, the findings showed that health information for caregivers is most usable when it is delivered in a solution-oriented (as opposed to a fact-based) manner. Incorporating principles of SMART goals may also be useful to improve usability.

The results of this study may be the building blocks for the examination of other eHealth information formats, making them valuable to future researchers testing the usability of health information products. The approach that we took in designing the 90Second Caregiver letters, and in subsequently testing their usability, has 3 main benefits. First, it appears to be useful for examining the format and delivery of health information (which is an understudied domain in health research). Second, it is useful for usability hypothesis generation (however, our present 90Second Caregiver design and delivery approach may not be useful for highly burdened caregivers or those who have been caregiving for many years). Third, our approach can generate information that is useful for a similar set of products (ie, health letters targeted toward other conditions/populations, such as 90Second Parent or 90Second Cannabis), even though we used a relatively small sample from our larger repertoire of health letters for this study.

Conflicts of Interest

PM may commercialize 90Second Caregiver. The authors have no other conflicts of interest to report, financial or otherwise.

References

1. Navaie-Waliser M, Feldman PH, Gould DA, Levine C, Kuerbis AN, Donelan K. When the caregiver needs care: the plight of vulnerable caregivers. Am J Public Health 2002 Mar;92(3):409-413. [doi: 10.2105/ajph.92.3.409] [Medline: 11867321]
2. Crellin NE, Orrell M, McDermott O, Charlesworth G. Self-efficacy and health-related quality of life in family carers of people with dementia: a systematic review. Aging Ment Health 2014;18(8):954-969 [FREE Full text] [doi: 10.1080/13607863.2014.915921] [Medline: 24943873]

3. Pinquart M, Sörensen S. Differences between caregivers and noncaregivers in psychological health and physical health: a meta-analysis. Psychol Aging 2003 Jun;18(2):250-267. [doi: 10.1037/0882-7974.18.2.250] [Medline: 12825775]

4. Merluzzi TV, Philip EJ, Vachon DO, Heitzmann CA. Assessment of self-efficacy for caregiving: the critical role of self-care in caregiver stress and burden. Palliat Support Care 2011 Mar;9(1):15-24. [doi: 10.1017/S1478951510000507] [Medline: 21352614]

5. Higgins O, Sixsmith J, Barry M, Domegan C. A Literature Review on Health Information- Seeking Behaviour on the Web: A Health Consumer and Health Professional Perspective. Stockholm: ECDC; 2011.

6. Tonsaker T, Bartlett G, Trpkov C. Health information on the internet: gold mine or minefield? Can Fam Physician 2014 May;60(5):407-408 [FREE Full text] [Medline: 24828994]

7. Hundert AS, Huguet A, Green CR, Hewitt AJ, Mushquash CJ, Muhajarine N, et al. Usability testing of guided internet-based parent training for challenging behavior in children with Fetal Alcohol Spectrum Disorder (Strongest Families FASD). J Popul Ther Clin Pharmacol 2016;23(1):e60-e76. [Medline: 27115205]

8. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors 1992;34(4):457-468. [doi: 10.1177/001872089203400407]

9. Donker A, Markopoulos P. A comparison of think-aloud, questionnaires and interviews for testing usability with children. In: People and Computers XVI - Memorable Yet Invisible. London: Springer; 2002:305-316.

10. Charters E. The use of think-aloud methods in qualitative research: an introduction to think-aloud methods. Brock Educ J 2003;12(2):68-82. [doi: 10.26522/BROCKED.V12I2.38]

11. Lund AM. Measuring usability with the USE questionnaire. J Usability Int 2001;8(2):3-6 [FREE Full text]

12. Stinson J, McGrath P, Hodnett E, Feldman B, Duffy C, Huber A, et al. Usability testing of an online self-management program for adolescents with juvenile idiopathic arthritis. J Med Internet Res 2010 Jul 29;12(3):e30 [FREE Full text] [doi: 10.2196/jmir.1349] [Medline: 20675293]

13. Breakey VR, Warias AV, Ignas DM, White M, Blanchette VS, Stinson JN. The value of usability testing for Internet-based adolescent self-management interventions: 'Managing Hemophilia Online'. BMC Med Inform Decis Mak 2013 Oct 4;13:113 [FREE Full text] [doi: 10.1186/1472-6947-13-113] [Medline: 24094082]

14. Fonda SJ, Paulsen CA, Perkins J, Kedziora RJ, Rodbard D, Bursell S. Usability test of an internet-based informatics tool for diabetes care providers: the comprehensive diabetes management program. Diabetes Technol Ther 2008 Feb;10(1):16-24. [doi: 10.1089/dia.2007.0252] [Medline: 18275359]

15. Stinson JN, Petroz GC, Tait G, Feldman BM, Streiner D, McGrath PJ, et al. e-Ouch: usability testing of an electronic chronic pain diary for adolescents with arthritis. Clin J Pain 2006;22(3):295-305. [doi: 10.1097/01.ajp.0000173371.54579.31] [Medline: 16514331]

16. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]

17. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. Qual Health Res 1993;3(4):430-441. [doi: 10.1177/104973239300300403]

18. Jaspers MW. 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]

19. Johnson CD, Zeiger RF, Das AK, Goldstein MK. Task analysis of writing hospital admission orders: evidence of a problem-based approach. AMIA Annu Symp Proc 2006;2006:389-393 [FREE Full text] [Medline: 17238369]

20. Wozney L, Huguet A, Bennett K, Radomski AD, Hartling L, Dyson M, et al. How do eHealth programs for adolescents with depression work? A realist review of persuasive system design components in internet-based psychological therapies. J Med Internet Res 2017 Aug 9;19(8):e266 [FREE Full text] [doi: 10.2196/jmir.7573] [Medline: 28793983]

21. 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]

22. Golafshani N. Understanding reliability and validity in qualitative research. Qual Rep 2003;8(4):597-606 [FREE Full text] [doi: 10.4135/9781412985659.n2]

Abbreviations

eHealth: electronic health
SMART: specific, measurable, attainable, realistic, and time bound

XSL•FO
RenderX

Original Paper

Usability of the Turkish Translation of the Dutch Talking Touch Screen Questionnaire for Physical Therapy Patients With a Turkish Background: Qualitative Study

Marlies Welbie[1], PhD; Harriet Wittink[1], PhD; Sahin Bozkurt[1], MSc; Tugba Coban[1], MSc; Walter LJM Devillé[2,3,4], MD, PhD

[1]Research Group Lifestyle and Health, Research Center Healthy and Sustainable Living, Utrecht University of Applied Sciences, Utrecht, Netherlands

[2]Julius Centre for health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, Netherlands

[3]Faculty of Social and Behavioral Sciences, University of Amsterdam, Amsterdam, Netherlands

[4]Dutch Centre of Expertise on Health Disparities (Pharos), Utrecht, Netherlands

Corresponding Author:
Marlies Welbie, PhD
Research Group Lifestyle and Health
Research Center Healthy and Sustainable Living
Utrecht University of Applied Sciences
Postbus 12011
Utrecht, 3501 AA Utrecht
Netherlands
Phone: 31 638192100
Email: marlies.welbie@hu.nl

Abstract

Background: The Turkish translation of the Dutch Talking Touch Screen Questionnaire (TTSQ) has been developed to help physical therapy patients with a Turkish background in the Netherlands to autonomously elucidate their health problems and impairments and set treatment goals, regardless of their level of health literacy.

Objective: The aim of this study was to evaluate the usability of the Turkish TTSQ for physical therapy patients with a Turkish background with diverse levels of health literacy and experience in using mobile technology.

Methods: The qualitative Three-Step Test-Interview method was carried out to gain insight into the usability of the Turkish TTSQ. A total of 10 physical therapy patients participated. The interview data were analyzed using a thematic content analysis approach aimed at determining the accuracy and completeness with which participants completed the questionnaire (effectiveness), the time it took participants to complete the questionnaire (efficiency), and the extent to which the participants were satisfied with the ease of use of the questionnaire (satisfaction). The problems encountered by the participants in this study were given a severity rating, which was used to provide a rough estimate of the need for additional usability improvements.

Results: No participant in this study was able to complete the questionnaire without encountering at least one usability problem. A total of 17 different kinds of problems were found. On the basis of their severity score, 3 problems that should be addressed during future development of the tool were "Not using the navigation function of the photo gallery in Question 4 causing the participant to not see all presented response items;" "Touching the text underneath a photo in Question 4 to select an activity instead of touching the photo itself, causing the activity not to be selected;" and "Pushing too hard or tapping too softly on the touch screen causing the touch screen to not respond." The data on efficiency within this study were not valid and are, therefore, not reported in this study. No participant was completely satisfied or dissatisfied with the overall ease of use of the Turkish TTSQ. Two participants with no prior experience of using tablet computers felt that, regardless of what kinds of improvement might be made, it would just be too difficult for them to learn to work with the device.

Conclusions: As with the Dutch TTSQ, the Turkish TTSQ needs improvement before it can be released. The results of this study confirm the conclusion of the Dutch TTSQ study that participants with low levels of education and little experience in using mobile technology are less able to operate the TTSQ effectively. Using a Dutch speaking interviewer and Turkish interpreter has had a negative effect on data collection in this study.

XSL·FO
RenderX

Introduction

Background

In the past three decades, health care provision in the Netherlands has evolved from a paternalistic to a patient-centered care approach. Since 1995, the government has introduced a series of laws and regulations aimed at increasing the autonomy and self-determination of patients [1]. Even today, policy makers, institutions, and health care professionals strive to further develop shared decision making and self-management in patients. Patients are increasingly expected to behave as active partners in encounters with health care professionals [2]. Not all patients are able to take on such a role. An important undermining factor is inadequate health literacy [3-5], which applies to 36% of the Dutch population [6].

Health literacy is defined as the cognitive and social skills that determine the motivation and ability of individuals to gain access to, understand, and use information in ways which promote and maintain health [7]. The concept contains cognitive and noncognitive aspects [8]. Cognitive aspects are referred to as "the capacity to think" and comprise functional skills such as literacy, numeracy, and information processing. Noncognitive aspects are referred to as "the capacity to act" and comprise skills such as goal setting, making a plan, and taking action [9]. Having the capacity to think and to act are equally important preconditions for patients taking on a proactive role during encounters with health professionals [8]. The majority of health literacy interventions, however, are aimed at improving cognitive skills [10-18]. To create a successful health literacy intervention, developers should (1) try to best fit the needs of persons with inadequate health literacy by incorporating members of the target group into their design team and (2) focus on noncognitive, as well as cognitive, aspects of health literacy [11]. On the basis of the results of current research, the possibilities of training noncognitive skills are expected to be limited [9]. This may mean that interventions aimed at increasing "the capacity to act" should not be focused on *training* noncognitive skills but on *supporting* them. This was exactly what the initiators of the development of the Dutch Talking Touch Screen Questionnaire (TTSQ) had in mind [19].

The Dutch TTSQ has been developed to help Dutch physical therapy patients, regardless of their level of health literacy, to elucidate their health problems and impairments, and set treatment goals. A total of 10 low-literate persons were involved in the development process of the prototype. In this prototype, which runs on a tablet computer, plain language and self-explanatory scales were used, alternatives to text were offered (eg, audio, pictures, and clips), and easily accessible background information on the questionnaire's rationale was provided. The development of the prototype of the Dutch TTSQ was described in detail in the study by Cremers et al [19]. It was pretested for usability [20] and face validity [21]. The results of both studies were promising but showed the need for further development.

Alongside the Dutch version, a Turkish version was developed. Development of this was seen as a starting point for development of other language versions. The initiators started with the Turkish version because people with a Turkish background form the biggest minority group in the Netherlands (about 400,000 people, 2.3% of the total population) [22]. Approximately one-third of the Turkish people aged between 15 and 65 years in the Netherlands only went to primary school, compared with 6% of the Dutch majority population [23]. The proportion of Turkish people with low literacy and low health literacy is unknown but, as education and literacy are very strongly associated [24,25], one can assume that low literacy and low health literacy are overrepresented in the Turkish minority group. Most people with low literacy are not digitally skilled [26], and recent studies found ethnic and socioeconomic differences in the use of mobile technology [27,28]. Therefore, it is to be expected that a relatively large proportion of this target population has little experience of using mobile technology. This may be a complicating factor in the use of the Turkish version of the TTSQ.

Objective

The aim of this study was to test the prototype of the Turkish TTSQ within the physical therapy context to see which parts of the prototype needed adjustment to increase user-friendliness for physical therapy patients with a Turkish background, regardless of their level of health literacy or experience of operating mobile technology.

The research question underlying this study was "What is the usability of the prototype of the Turkish TTSQ for physical therapy patients with a Turkish background with diverse levels of health literacy and experience in using mobile technology?"

Methods

Design

A qualitative descriptive case study [29] was carried out. Data were collected and analyzed as in the study on ease of use of the Dutch version of the TTSQ [20]. Data on the way participants operated the Turkish TTSQ were collected through the Three-Step Test-Interview (TSTI) method [30]. This method includes both think-aloud and retrospective probing techniques.

Definitions

Usability was defined by the International Standards Organization as "the effectiveness, efficiency and satisfaction with which specified users can achieve goals in particular environments" [31].

Effectiveness is the accuracy and completeness with which users achieve certain goals [32]. In this study, rates and severity of problems were used as primary indicators of effectiveness.

Efficiency is the relation between the accuracy and completeness with which users achieve certain goals and the resources expended in achieving them [32]. In this study, task completion time was used as an indicator of efficiency.

Satisfaction is the users' comfort with and positive attitudes toward the use of a system [32]. In this study, participants were interviewed about their satisfaction with the ease of use of the Turkish TTSQ. Ease of use was defined as the degree to which the use of a particular system is free from effort [33].

Setting and Participant Selection

Recruitment took place in 12 primary care practices in deprived areas of Utrecht, the Netherlands. Potential participants were invited by their physical therapist to participate in this study. Researcher SB was a native Turkish speaker with a Turkish background and employed as a physical therapist in one of the recruiting practices. No other recruiting therapists had Turkish backgrounds or spoke Turkish. Each recruiting therapist shortly explained the goal of the study to potential participants and provided them with Turkish and Dutch versions of a flyer and information letter. The flyer contained a brief summary of the background and goal of the research project and an invitation to its readers to read more about the project in the accompanying information letter. Both versions of the flyer and information letter were written in plain language. If patients were interested in participating, their therapist asked permission to give their contact information to the researchers. If patients spoke and understood Dutch, researcher MW contacted them by telephone; otherwise, researcher SB contacted them. During the telephone conversation, the researchers invited questions, checked that patients understood what was being asked of them, and checked that inclusion criteria were met. Inclusion criteria were: aged 18 years or older, able to understand the Turkish language, and both parents born in Turkey. The sampling procedure was aimed at getting a sample of 6 to 12 participants, typical for formative usability testing of devices such as the TTSQ [34] because it would reveal the most important points needing improvement for further development of a tool without the risk of unnecessary expenditures [35]. Data collection was stopped when a good balance was reached in terms of age, gender, level of education, level of functional health literacy, and prior experience with using a tablet computer. Throughout the recruitment process, the recruiting physical therapists were constantly kept informed about the profiles of participants the researchers were looking for.

Content of the Turkish Talking Touch Screen Questionnaire

The prototype of the Turkish TTSQ (see Multimedia Appendix 1) is a direct translation of the Dutch TTSQ [19-21], which is described in detail in the methodological sections by Welbie et al 2018 and 2019 [20,21].

Translation of the Dutch TTSQ into Turkish was done by a native Turkish speaker who worked as a Turkish language teacher in the Netherlands. Comprehension of the translated text was tested by researcher TC, a native Turkish speaker with a Turkish background. She asked 7 non-Dutch–speaking women, who were born in Turkey and now lived in the Netherlands, to read the written text, listen to the spoken text in the Turkish TTSQ, and explain to her what they thought was meant by the questions and answer options. The 7 women had finished primary school at most and were following different kinds of courses (such as cooking and handicraft) together at a mosque

in Utrecht. The 7 Turkish female testers had no problems understanding both spoken and written text. An overview of all types of screens is given in Screenshots 1 to 16 in the Multimedia Appendix 1. The 8 questions of the questionnaire can be found in Multimedia Appendix 1: screenshots 2, 3, 4, 7, 9, 11, 12, and 13.

Data Collection and Procedures

Data collection took place at the physical therapy practice or the participant's home, depending on the preference of the participant. Researchers MW and SB were present. Researcher MW was in the lead during the interviews. She communicated in Dutch during the whole meeting. Researcher SB functioned as an observer as well as an interpreter when participants spoke Turkish. As an interpreter-researcher, SB did not interfere in the conversation but solely acted as an intermediary. Participants spoke Dutch, Turkish, or a mixture of both languages, depending on their preference and abilities. At the end of the interview, researcher SB asked complementary questions if some information was lacking. When SB asked these questions in Turkish, he directly translated them and later the answers given by the participants into Dutch so that researcher MW could closely follow what was said.

The following data-gathering steps were taken according to the TSTI method [30]:

- *Step 1:* All participants were observed by researchers MW and SB while they were completing the Turkish TTSQ. During the completion of the questionnaire, they thought out loud. When participants spoke Turkish or used some Turkish words, researcher SB took on the role of interpreter and translated the text into Dutch. This step was aimed at collecting observational data on the usability of the Turkish TTSQ. The data collected consisted of 2 types: (1) observations of participants' behavior and (2) think-aloud data. A video recording was made of this interview step. The video camera was aimed at the tablet computer and the hands of the participant while operating the screen. In addition, both researchers MW and SB took real-time notes for use during the following steps of the interview as well as for later analysis. The researchers wrote their notes down on hard copies of screenshots of the Dutch TTSQ, which were printed next to the identical screens of the Dutch questionnaire, so researcher MW was able to read the question and answer options in Dutch. Researchers MW and SB noted problems with operating the tablet computer, including using the touch screen, navigating through the questionnaire, understanding the task given in each screen, selecting response items, using the correction function, and use of the stop and help buttons.
- *Step 2:* Researcher MW interviewed each participant after they had finished completing the Turkish TTSQ. Data collection during this step was exclusively focused on filling possible gaps and checking the observational data collected during step 1. An audio recording was made of this interview step.
- *Step 3:* During step 3 of the TSTI, researcher MW conducted a semistructured interview aimed at eliciting experiences and opinions of participants. At the end of the

interview, researcher SB asked complementary questions, if he felt it was necessary, to get complete and rich data. When participants encountered problems in operating the Turkish TTSQ, they were asked what they thought the exact nature and possible cause of each type of problem was. In addition, they were asked how they tried to overcome the problem and if they had suggestions for making it easier to operate the Turkish TTSQ at this point. Afterwards, the participants were questioned about their satisfaction regarding the overall ease of use of the Turkish TTSQ. The participants were encouraged to report feelings, express opinions, state preferences, and make recommendations. An audio recording was made of this interview step.

When the interview was finished, demographic data, data on self-reported experience with using a tablet computer, self-reported health, and functional health literacy measured with the Set of Brief Screening Questions-Dutch version (SBSQ-D) [36] were collected (see Tables 1 and 2). The SBSQ-D is the Dutch version of Chew's SBSQ. This tool consists of the following 3 statements: "How often do you have someone help you read hospital materials?" "How confident are you filling out medical forms by yourself?" and "How often do you have problems learning about your medical condition because of difficulty understanding written information?" The combined item-responses result in a subjective health literacy score [37,38]. The SBSQ-D was conducted orally by researcher SB who translated the statements into Turkish if necessary.

Analyses

Data were analyzed using a thematic content analysis approach [39]. A total of 4 types of data were analyzed, which were as follows: (1) video recordings of the completion of the questionnaire, (2) field notes of the observed participant behavior, (3) transcriptions of the Dutch spoken text within the video and audio recordings, and (4) background information regarding educational level, level of literacy, age, gender, and prior experience using a tablet computer.

Only the *Dutch* spoken text within the interviews was transcribed. After transcription, researcher TC listened closely to the recordings while looking at the transcriptions of the Dutch spoken text. When she disagreed with the translation made by researcher SB during the interview, she added what she thought was a more accurate translation to the transcript in a different color. Afterwards, researcher TC and SB sought consensus on the most accurate translation.

Researcher MW started the coding process by coding step 1 of the interview directly on the video recordings, using MAXQDA 12 (VERBI Software). This was partly an inductive and partly a deductive process. The deductive process consisted of using the descriptions of the 13 usability problems found in the ease of use study of the Dutch TTSQ [20] as codes. The inductive process comprised open coding of new problems, statements of the participants about the cause of these problems, and the way they thought these problems could be avoided in the future. In addition, statements of participants about satisfaction regarding the ease of use of the Turkish TTSQ were coded, and completion times were registered. After researcher MW finished coding step 1 for 1 interview, she checked from the transcription

of steps 2 and 3 of that interview whether the problems were described and spoken about in a way congruent with her analysis of step 1. If not congruent, she watched the video again to see if her initial coding for step 1 needed adjustment. In addition, she coded the statements participants made during steps 2 and 3 about the causes of problems during completion of the Turkish TTSQ and the ways they thought these problems could be avoided. She also coded all statements of participants about satisfaction with ease of use of the Turkish TTSQ.

Directly after coding all 3 parts of an interview, researcher MW made a descriptive summary of that interview. Each summary contained information on whether or not the questionnaire was fully completed; if, when and why the stop function was used; if, when and why the help function was used and whether this was effective; the kinds of problems that occurred with the operation; the completion times; and all emerging themes regarding satisfaction with ease of use of the questionnaire. The themes emerging in the summaries were supplemented with related field notes and information regarding educational level, health literacy level, age, gender, and experience in using mobile technology. Afterwards, researcher MW compared this summary with that made at the end of the interview to check for inconsistencies. If any were found, she looked at all related data again to see if her interpretation and coding of what had happened and was said during the interview needed adjustment.

As the last step of the content analysis, researcher SB took on the role of peer debriefer to test the emerged hypotheses and see if they were reasonable and plausible to him. To get a good understanding of how the hypotheses emerged, researchers MW and SB looked at the summaries, codes, and raw data (transcripts and videos) together. During their conversation, they constantly and explicitly reflected on the influence their Turkish and Dutch backgrounds might have had on their views on the data and whether or not this made their interpretations of the data differ at any point.

As a next step, researcher MW extracted the observed usability problems from the summaries. MW reanalyzed the video recordings to see how many times each problem had occurred in total and per participant. After a full overview of problems had emerged, she categorized the problems as low, medium, serious, or critical as described by Nielsen and Loranger [40]. The scoring method was described in detail in Welbie et al [20]. Nielsen and Loranger recommend tackling only serious and critical problems during the development of a digital tool because those of low and medium severity are not worth tackling from a cost-benefit perspective. Serious and critical problems, however, can be so disruptive that they make users stop using a tool or prevent them from even starting to use it [40].

During the whole course of the study, procedures, coding, analysis steps, and interpretation decisions were discussed with researchers HW and WD.

Transcripts were made in the Dutch language. Only quotes used in this paper were translated from Dutch into English by researcher MW and checked by researcher HW, who is a bilingual speaker.

Ethics

No external funding was received by the Utrecht University of Applied Sciences to conduct this study. This study was registered with the medical ethics committee of the Academic Medical Centre of Amsterdam, which declared that it does not fall under the scope of the "Medical Research Involving Human Subjects Act." The study was conducted according to the principles of the Declaration of Helsinki [41]. All participants provided written informed consent. The participants' names used in this study are all fictitious to protect their privacy.

Results

Study Population

A total of 10 physical therapy patients were included in this study. Characteristics of the study population can be found in Tables 1 and 2.

Table 1. Characteristics of study population (n=10).

Characteristics	Value
Age (years), mean (range)	53 (35-74)
Gender, n	
Male	6
Female	4
Level of education, n	
Low[a]	4
Moderate[b]	4
High[c]	2
Functional health literacy level measured with Set of Brief Screening Questions-Dutch version [36]	
Adequate	5
Inadequate	5
Prior experience operating a tablet computer, n	
Yes	5
No	5

[a]Low: none or at most finished primary education.

[b]Moderate: lower secondary education, (upper) secondary education, or post-secondary nontertiary education (including vocational education).

[c]High: tertiary education (bachelor's degree or higher).

Table 2. Characteristics per participant.

Pseudonym	Gender (F: female and M: male)	Age (years)	Level of education	Functional health literacy level measured with Set of Brief Screening Questions-Dutch version [36]	Self-reported health status	Prior experience using a tablet computer
Meryem	F	74	Low[a]	Inadequate	Poor	No
Mert	M	71	Low[a]	Inadequate	Poor	No
Ceyda	F	65	Low[a]	Inadequate	Satisfactory	No
Gizem	F	44	Low[a]	Inadequate	Poor	Yes
Memhet	M	59	Moderate[b]	Inadequate	Good	No
Berat	M	38	Moderate[b]	Adequate	Satisfactory	Yes
Elif	F	40	Moderate[b]	Adequate	Good	Yes
Eren	M	48	Moderate[b]	Adequate	Good	No
Imraam	M	52	High[c]	Adequate	Good	Yes
Onur	M	35	High[c]	Adequate	Good	Yes

[a]Low: none or at most finished primary education.

[b]Moderate: lower secondary education, (upper) secondary education, or postsecondary nontertiary education (including vocational education).

[c]High: tertiary education (bachelor's degree or higher).

Effectiveness

Of the 10 participants, 2 managed to complete the questionnaire fully. Both had prior experience with operating tablet computers (see Table 3). Ceyda (age 65 years) and Meryem (age 74 years) left all questions open, and Mehmet (age 59 years) stopped completing the questionnaire at question 5. All 3 were inexperienced in operating tablet computers. Inexperienced Eren (age 48 years) and Mert (age 71 years) and experienced Imraam (age 52 years), Elif (age 40 years), and Gizem (age 44 years) went through the whole questionnaire but unintentionally left 1 or more parts incomplete.

Table 3. Prior experience with using a tablet computer in comparison with ability to fully complete the Turkish Talking Touch Screen Questionnaire.

Population	Not fully completed, n	Fully completed, n
No prior experience using a tablet computer (n=5)	5	0
Prior experience using a tablet computer (n=5)	3	2
Total population (n=10)	8	2

Unintentionally Unanswered (Parts of) Questions

Inexperienced Eren (age 48 years) and Mert (age 71 years) and experienced Imraam (age 52 years), Elif (age 40 years), and Gizem (age 44 years) failed to fully complete the Turkish TTSQ because they failed to select answer options and/or unintentionally skipped questions because of problems such as tapping on the text underneath a photograph instead of on the photograph itself and by double-tapping on the next button (see problems 1, 3, 4, and 5 in Table 4). None of the participants noticed they had failed to select answer options or skipped questions while they were completing the questionnaire.

Table 4. Frequency and severity of problems encountered during the completion processes for all participants.

Problem	Number of participants who encountered the problem	Number of times the problem occurred	Severity rating
1. Accidentally skipping a screen by double-tapping the "next" button	2	2	Low
2. Double-tapping an answering option causing activation and deactivation of the answer of choice	0	0	__a
3. Skipping a screen by accidentally touching the next button with the palm of the hand	0	0	—
4. Not using the navigation function of the photo gallery in question 4 causing the participant to not see all response items	5	22	Serious
5. Touching the text under a photo in question 4 to select an activity, instead of touching the photo itself, causing the activity not to be selected	4	10	Critical
6. Not able to see whether or not a selected answer is activated (not accentuated enough)	1	1	Low
7. Not knowing how to get to the next screen	0	0	—
8. Pushing too hard or tapping too softly on the touch screen so that it does not respond	8	19	Serious
9. Not able to correct a wrong answer	3	3	Medium
10. Not reading the text above the photos in question 5, causing the participant to continue the task given in question 4	1	1	Low
11. Not noticing that the multiple numeric rating scale "effort" scores in question 8 are related to different activities, which in error results in identical scores for different activities	1	1	Low
12. Mistakenly scoring the mirror image in the body chart in question 2	1	1	Low
13. Scoring (serial) questions that do not apply to the participants' situation (forced by the software)	0	0	—
14. Using navigation function question 4 to try to get to the next screen.	2	2	Low
15. Not knowing how to enter an answer into the TTSQ[b]	2	2	Medium
16. Not being aware of the existence of the "help" function	2	2	Medium
17. Entering more than one answer into an NRS[c] causing the TTSQ to select only the last entered answer	2	2	Low
18. Activating the "stop" function accidentally by touching it with the palm of the hand holding the tablet	1	2	Medium

[a]This problem was found in the study on the Duth Talking Touch Screen Questionnaire [20], not in this study.

[b]TTSQ: Talking Touch Screen Questionnaire.

[c]NRS: Nummeric Rating Scale

Stopped Completing Prematurely

Inexperienced Ceyda (age 65 years) read the first question "Do you have pain" (see Multimedia Appendix 1, screenshot 2 "Pain"). She was very doubtful about what answer would be right because her pain had decreased since her first physical therapy visit. She gave back the Turkish TTSQ to the researcher without answering the question because she was not able to decide on her answer, and skipping the question was not a possibility. Afterwards, during interview step 3, she told the researcher that she did not know that the red square with "yes" in it and the green square with "no" in it were "buttons," which she could have tapped to insert an answer.

Inexperienced Meryem (age 74 years) did not know what to do with the tablet. She read the first question and then spoke directly to researcher MW to give the answer. When the researcher asked her what she thought she should do next, she answered:

Well, I hope to benefit from the therapy. That's what I am going for [Meryem, age 74 years]

When the researcher then asked her if she had any idea what she should do with "the screen," she seemed to get somewhat nervous and almost whispered:

I don't know, I do not know what to say [Meryem, age 74 years]

Inexperienced Memhet (age 59 years) managed to get to question 4 without encountering any serious or critical usability problems. In this question, he was asked to select photographs of activities in which he was limited (see Multimedia Appendix 1, screenshot 7 "Activity 'lying'"). Memhet tapped on the text beneath the photographs most of the time instead of on the given

photographs. He did not notice that this was not sufficient to select the answering option and therefore, thought that he had selected far more photos than he actually had. In question 5, he was asked to select the 3 activities that were most important to him out of those he selected in answer to question 4 (see Multimedia Appendix 1, screenshot 9 "Most important activities"). As most of his answers had not actually been "selected," he only saw a fraction of his "activity selection." This confused him. He thought he had misunderstood the question. He did not know how to answer it. After he unsuccessfully tried to skip the question by tapping on the "next" button, he stopped completing the questionnaire by handing it back to the researchers.

Frequency and Severity of Problems Encountered

Even though 2 participants were able to complete the Turkish TTSQ fully (see Table 3), no participant completed it without encountering any problems. A complete overview of the frequency and severity of all problems encountered during operation of the Turkish TTSQ can be found in Table 4.

Efficiency

Because of the need to translate the "spoken out loud thoughts" of participants into Dutch, the completion time was lengthened. As a result, the collected data on efficiency were not valid and will not be reported in this paper.

Satisfaction

Positive Remarks

No participant was distinctly positive or negative about the overall ease of use of the Turkish TTSQ. Out of 10 participants, 5 made positive remarks on the way the user interface was designed and on the short completion time.

> These visual images are appealing and make it "come to life". [Imran, age 52 years]

Experienced Onur (age 35 years) was positive about the regular overviews of given answers, and inexperienced Eren (age 48 years) was positive about the short length of the questionnaire.

Recommendations for Improvement

Out of 10 participants, 9 formulated recommendations for improvement. Most mentioned recommendations were improved accentuation of the activated response items, give a complete overview of activities to choose from in answer to question 4, and shorten the instruction clips by limiting the information to the main issues.

Inexperienced Seyda (age 65 years) and Meryem (age 74 years) had trouble concentrating on the information in the introduction clip, as did others. However, they were not sure if limiting the amount of information or length of the clip was going to help them. They felt it would just be too difficult for them to learn to work with the Turkish TTSQ, regardless of improvements on its usability. They linked their lack of ability to comprehend and remember the instructions given on their lack of experience with operating tablet computers, their older age, and their health status.

Experienced Berat (age 38 years) recommended limiting the text above the overviews. For example, he suggested deleting the first sentence from the text: "On this screen you see all the activities that you selected in previous screens. These are the activities in which you are limited. Is that right?" (see Multimedia Appendix 1, screenshot 8 "Overview activities").

Some participants suggested adding more advanced options to the Turkish TTSQ. Experienced Onur (age 35 years) and Berat (age 38 years) recommended a swipe function for the screens that contained rows of activity photos. Experienced Elif (age 40 years) would have liked to see muscles in the body chart so she would be able to indicate the location of her pain more precisely. Like Elif, Onur and inexperienced Eren (age 48 years) also wanted to be able to indicate the locations of their complaints more precisely, but they suggested a function that would enable them to zoom in on a specific body part.

Discussion

Principal Findings

In all, 2 participants, who had prior experience with using tablet computers, managed to complete the questionnaire fully without leaving any parts unanswered. No participant in this study was able to complete the questionnaire without encountering a usability problem.

A total of 17 different kinds of problems were found. Three problems should be addressed during future development of the tool based on their severity score [40]: "Not using the navigation function of the photo gallery in Question 4 causing the participant to not see all presented response items," "Touching the text underneath a photo in Question 4 to select an activity instead of touching the photo itself causing the activity not to be selected," and "Pushing too hard or tapping too softly on the touch screen causing the touch screen to not respond."

No participant was distinctly satisfied or dissatisfied about the overall ease of use of the Turkish TTSQ. Positive remarks were mainly made on the user interface and the short completion time of the Turkish TTSQ. The most frequently made recommendations were improve accentuation of the activated response items, give a complete overview of activities to choose from in answer to question 4, and shorten the instruction clips by limiting the information to the main issues. Two inexperienced participants felt that, regardless of what improvements might be made, it would just be too difficult for them to learn to work with the device.

Strengths and Limitations

A strength of this study was the inclusion of 10 members of a target population that is generally "hard to reach" for researchers, including some of the most vulnerable subjects within this population [42]. This made it possible to both collect data from people who are rarely represented in research populations and, at the same time, gather knowledge about the effects strategic and methodological choices have on the quality of research in such populations. Researcher SH played an important role in the recruitment and data collection within this study. His Turkish background and his being a native Turkish speaker, combined with his network, status, and trustworthiness

as a physical therapist working in the community, may have had a positive influence on the willingness of potential candidates to participate in this study [43]. This hypothesis is reinforced by the fact that, although recruitment was done in 12 different physical therapy practices, 8 out of 10 participants were recruited in the practice where researcher SH was employed.

The positive effect researcher SH had on the sampling procedure may also have had a downside. Despite all efforts of the researchers to inform potential participants thoroughly and make sure that participation was done voluntarily, the authority of researcher SH as researcher and physical therapist [43] may have caused participants to agree to participate too quickly without really foreseeing what was being asked of them. The majority of the participants seemed to have "a lot on their plate" and were, therefore, not able to entirely focus on their tasks during the data collection process. A total of 8 out of 10 participants reported multiple health problems. One participant even ended the interview prematurely because it became too much for her because of her physical and mental state. Another participant, who reported 11 different kinds of health problems, told the researchers that his biggest problem was not even his health status but his poor financial situation. In hindsight, the researchers got the impression that, for some, participation in this study may have been too much to ask.

The bilingual research setting also brought some limitations to this study. Apart from the translation lengthening the completion time, 3 participants forgot to insert some of their answers during the completion process, although they *did* formulate their answers when thinking out loud. They all said they would not have forgotten this in a "real life" physical therapy setting where there would have been no observers or interpreters present and they would not have been asked to think out loud. A total of 3 other participants said that the translation limited their ability to concentrate on their task and thoughts. This may have caused participants to make more mistakes than they would have done had the whole interview been in the Turkish language.

Comparison With Prior Work

Although there is a considerable amount of overlap in the kind and severity of problems encountered in the current and Dutch TTSQ study [20], the participants of this study encountered different kinds of problems and were less able to complete the questionnaire fully than those in the Dutch TTSQ study. The explanation for this can be found in the fact that, compared with the Dutch study, the population of this study was less educated, had lower health literacy, and had less experience with using tablet computers. In this study, no participant was completely satisfied or dissatisfied with the overall ease of use of the Turkish TTSQ, although, in the Dutch TTSQ study, the participants were not only very satisfied but their expectations of ease of use of the tool were exceeded [20]. In contrast to the Dutch TTSQ study, not all Turkish participants had the sense of self-efficacy to be able to complete the Turkish TTSQ, no matter what improvements might be made. The results of the Dutch TTSQ study showed that participants with lower education and less experience in using mobile technology were

less able to operate it effectively [20]. This is confirmed by the results of this study.

Two earlier studies were found in which usability was part of the assessment of a direct translation of a Talking Touchscreen (TT) questionnaire, both published by Hahn et al [44,45]. In the 2003 study, the usability components "satisfaction" and "efficiency" were tested. In this study, 30 Spanish-speaking patients with cancer completed a TT which contained the Functional Assessment of Cancer Therapy-General (FACT-G) [46] and the Short Form-36 Health Survey [47]. A total of 50% (7/15) of the participants had lower than 7th grade education. Satisfaction with ease of use and efficiency were tested by presenting evaluation questions on the use of the TT followed by a short debriefing interview. What is noticeable about the satisfaction and efficiency results is that all 30 participants reported that they thought of the tool as "very easy" or "easy to use" and the completion "did not take too long," whereas 57% (8/15) of participants with less than 7th grade education and 14% (2/15) of the participants with more than 7th grade education preferred an interviewer orally conducting the questionnaire to use of the TT. Hahn et al [44,45] interpreted these results in a positive way and reported that *many patients either preferred using the touchscreen rather than having an interviewer ask the questions, or had no preference.* Although true for the more educated participants, the majority of the less educated participants did not prefer using the TT. Hahn et al [44,45] concluded their paper by stating that *the "Talking Touchscreen" will allow Latino patients with varying literacy skills to be included more readily in clinical trials, clinical practice research and QOL studies.* This conclusion may be too 1 dimensional, given the results they reported and the methods they used. In the other study by Hahn et al, published in 2010, only user satisfaction was tested [45]. In this study, 414 Spanish-speaking patients with cancer were included of which 213 had low levels of literacy. The tested touch screen system contained the FACT-G [46], SF-36 [47], and Standard Gamble Utility Questionnaire [48]. The methods used to test satisfaction about the ease of use were highly comparable with the earlier study of Hahn et al [44]. Looking at the quantitative results, one can conclude that, although satisfaction among the majority of the participants was high, low-literacy participants were less satisfied with the ease of use of the TT than were those with high literacy. It is hard to compare the results of the studies of Hahn et al with the results of this study because, although the participants in their studies could ask for assistance from the researchers during completion of the TT, participants in this study did not receive any help at all. In the 2003 study, 60% of participants received help from a researcher during completion of TT; how many received help in the 2010 study was not reported. It can be concluded that researchers in this study tested and reported the usability of their tool much more thoroughly. Although it is difficult to directly compare the results of the Hahn et al studies with the current studies because of differences in study setups and the detail in which results were reported, the results of both Hahn et al studies seem to confirm our findings that it is harder for less educated participants to use a TT than for higher educated participants.

Conclusions

Just like the Dutch TTSQ, the Turkish TTSQ needs improvement before it can be released. The results of this study confirm the conclusion of the Dutch TTSQ study that participants with low education and little experience in using mobile technology are less able to operate the TTSQ effectively. Although the methodology of this usability study was very thorough, using a Dutch-speaking interviewer and Turkish interpreter has had a negative effect on data collection.

Directions for Future Research

The aim of the project, of which this study is a part, is to create multiple language versions of the TTSQ to help Dutch physical therapy patients, regardless of their level of health literacy, to elucidate their health problems and limitations, and set treatment goals. The results of both usability studies of the TTSQ show that this should particularly be improved for the least skilled future users. Therefore, the logical next step is adapting and testing both language versions of the tool solely with inexperienced users who have low literacy. When the pretests show that future users at risk of exclusion are able to complete the Turkish and Dutch versions of the TTSQ fully without encountering serious or critical usability problems, pretests on response processes should be conducted to get a first impression of the face validity of both versions of the questionnaire [49]. In addition, the equivalence of both language versions should be tested using item response theory [50]. Dependent on the results of these response processes and item response theory studies, cultural adaptation of the Turkish TTSQ may be needed to avoid bias from cultural and linguistic effects on interpretation, retrieval, judgment, and response selection, which are the 4 phases of the response process as described by Tourangeau et al [51]. Both researchers and participants should communicate in Turkish in all future studies on the Turkish TTSQ to avoid the methodological problems encountered in this study. Recruitment of participants with a Turkish background should be done by intermediaries with Turkish backgrounds, rather than by the researchers themselves, to limit the chance of people agreeing to participate too easily without foreseeing the consequences of their participation. When the results of all pretests are satisfactory, the last step in research should be quantitative usability, validity, and reliability testing to produce generalizable data.

No data on levels of literacy, health literacy, or digital skills are available for the Turkish minority group in the Netherlands. Research should be done to get insight into these characteristics and into attitudes toward use of information and communication technology in general and of mobile health (mHealth) technology more specifically within this and other minority groups. Otherwise, these already disadvantaged groups may not be able to profit from the advantages of the use of mHealth and electronic health technologies [52-54]. This may add to the ongoing exacerbation of health inequalities in the Netherlands [55].

It is of great importance to keep striving for the development of TT questionnaires, which are user-friendly to low literacy minority patients who have not mastered the native language of the countries in which they are living in. Such tools will greatly facilitate data collection within these hard-to-reach populations. It will empower vulnerable patients who will be able to give their input to research and clinical practice. And because they will not need help or instructions from researchers or health care providers, it will reduce staff burden, costs, and interviewer bias. The use of TT questionnaires may also serve as a way to increase exposure of underserved populations to new technologies and contribute to information about the experiences of diverse populations with these technologies [56]. To get reliable and valid test results for the evaluations of these tools, researchers need to keep striving for research setups and methods that fit the needs and abilities of hard-to-reach populations. Publishing positive as well as negative results on usability, reliability, and validity and giving as much insight into evaluation methods, study contexts, and setups as possible will help researchers and developers in finding ways to accommodate hard-to-reach populations and contribute to the body of knowledge on inclusive design-oriented research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots Turkish Talking Touch Screen Questionnaire.
[PDF File (Adobe PDF File), 804 KB - formative_v4i2e14189_app1.pdf]

References

1. Schäfer W, Kroneman M, Boerma W, van den Berg M, Westert G, Devillé W, et al. The Netherlands: health system review. Health Syst Transit 2010;12(1):v-xxvii, 1. [Medline: 21132996]
2. Kooiker S, Hoeymans N. Zorgkennis. 2014. Citizens and Health. Theme Report: Public Health Future Exploration 2014 URL: https://www.zorgkennis.net/wp-content/uploads/2019/09/

ZK-kennisbank-Burgers-en-Gezondheid-Themarapport-Volksgezondheid-Toekomst-Verkenning-2014-2412.pdf [accessed 2019-12-06]

3. Rademakers J. Nivel. 2013 Oct. Knowledge Synthesis: The Dutch Patient and Healthcare User URL: https://www.nivel.nl/sites/default/files/bestanden/nederlandsepatientinbeeld.pdf [accessed 2019-12-06]

4. Rademakers J. Nivel. 2014 Sep. Knowledge Synthesis: Health skills Not Obvious for Everyone URL: https://www.nivel.nl/sites/default/files/bestanden/Kennissynthese-Gezondheidsvaardigheden-2014.pdf [accessed 2019-12-06]

5. Heijmans M, Waverijn G. Nivel. 2014. Health Skills of the Chronically Ill Important for Self-Management URL: https://www.nivel.nl/sites/default/files/bestanden/Gezondheidsvaardigheden_chronische_zieken_belangrijk_zelfmanagement.pdf [accessed 2019-12-06]

6. Heijmans M, Brabers A, Rademakers J. Nivel. 2018. Health Literacy in Nederland URL: https://nivel.nl/sites/default/files/bestanden/Gezondheidsvaardigheden_in_Nederland.pdf [accessed 2012-12-06]

7. Nutbeam D. Health promotion glossary. Health Promot 1986 May;1(1):113-127. [doi: 10.1093/heapro/1.1.113] [Medline: 10318625]

8. Rademakers J, Heijmans M. Beyond reading and understanding: health literacy as the capacity to act. Int J Environ Res Public Health 2018 Aug 7;15(8):pii: E1676 [FREE Full text] [doi: 10.3390/ijerph15081676] [Medline: 30087254]

9. WRR: Scientific Council for Government Policy. Knowing is Not Doing Yet. A Realistic Perspective on Reasonability URL: https://www.wrr.nl/publicaties/rapporten/2017/04/24/weten-is-nog-geen-doen [accessed 2019-12-06]

10. Visscher BB, Steunenberg B, Heijmans M, Hofstede JM, Devillé W, van der Heide I, et al. Evidence on the effectiveness of health literacy interventions in the EU: a systematic review. BMC Public Health 2018 Dec 29;18(1):1414 [FREE Full text] [doi: 10.1186/s12889-018-6331-7] [Medline: 30594180]

11. Heijmans M, Uiters E, Rose T, Hofstede J, Devillé W, Heide I, et al. Study on Sound Evidence for a Better Understanding of Health Literacy in the European Union (HEALIT4EU). Luxembourg: European Union; Jun 2015.

12. Mantwill S, Monestel-Umaña S, Schulz PJ. The relationship between health literacy and health disparities: a systematic review. PLoS One 2015;10(12):e0145455 [FREE Full text] [doi: 10.1371/journal.pone.0145455] [Medline: 26698310]

13. D'Eath M, Barry M, Sixsmith J. Rapid Evidence Reviews of Interventions for Improving Health Literacy. Stockholm: European Centre for Disease Prevention and Control; May 2012.

14. Dennis S, Williams A, Taggart J, Newall A, Denney-Wilson E, Zwar N, et al. Which providers can bridge the health literacy gap in lifestyle risk factor modification education: a systematic review and narrative synthesis. BMC Fam Pract 2012 May 28;13:44 [FREE Full text] [doi: 10.1186/1471-2296-13-44] [Medline: 22639799]

15. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Viera A, Crotty K, et al. Health literacy interventions and outcomes: an updated systematic review. Evid Rep Technol Assess (Full Rep) 2011 Mar(199):1-941. [Medline: 23126607]

16. Sheridan SL, Halpern DJ, Viera AJ, Berkman ND, Donahue KE, Crotty K. Interventions for individuals with low health literacy: a systematic review. J Health Commun 2011;16(Suppl 3):30-54. [doi: 10.1080/10810730.2011.604391] [Medline: 21951242]

17. DeWalt DA, Hink A. Health literacy and child health outcomes: a systematic review of the literature. Pediatrics 2009 Nov;124(Suppl 3):S265-S274. [doi: 10.1542/peds.2009-1162B] [Medline: 19861480]

18. Sanders LM, Federico S, Klass P, Abrams MA, Dreyer B. Literacy and child health: a systematic review. Arch Pediatr Adolesc Med 2009 Feb;163(2):131-140. [doi: 10.1001/archpediatrics.2008.539] [Medline: 19188645]

19. Cremers AH, Welbie M, Kranenborg K, Wittink H. Deriving guidelines for designing interactive questionnaires for low-literate persons: development of a health assessment questionnaire. Univers Access Inf Soc 2017;16(1):161-172. [doi: 10.1007/s10209-015-0431-2]

20. Welbie M, Wittink H, Westerman MJ, Topper I, Snoei J, Devillé WL. A mobile patient-reported outcome measure app with talking touchscreen: usability assessment. JMIR Form Res 2019 Sep 27;3(3):e11617 [FREE Full text] [doi: 10.2196/11617] [Medline: 31573909]

21. Welbie M, Wittink H, Westerman MJ, Topper I, Snoei J, Devillé WL. Using plain language and adding communication technology to an existing health-related questionnaire to help generate accurate information: qualitative study. J Med Internet Res 2018 Apr 23;20(4):e140 [FREE Full text] [doi: 10.2196/jmir.7940] [Medline: 29685873]

22. StatLine - CBS. Population; key figures URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37296NED/table?fromstatweb [accessed 2019-12-06] [WebCite Cache ID 774X1LU5c]

23. Portegijs W, van den Brakel M. Emancipation Monitor 2016. Netherlands: SCP Publications; 2016.

24. Buisman M, Houtkoop W. Low Literacy on the Map. Den Bosch: ECBO; 2014.

25. van der Heide I, Wang J, Droomers M, Spreeuwenberg P, Rademakers J, Uiters E. The relationship between health, education, and health literacy: results from the Dutch Adult Literacy and Life Skills Survey. J Health Commun 2013;18(Suppl 1):172-184 [FREE Full text] [doi: 10.1080/10810730.2013.825668] [Medline: 24093354]

26. Algemene Rekenkamer. 2016. Approach to Low Literacy URL: https://www.rekenkamer.nl/publicaties/rapporten/2016/04/20/aanpak-van-laaggeletterdheid [accessed 2019-12-06]

27. Arora S, Ford K, Terp S, Abramson T, Ruiz R, Camilon M, et al. Describing the evolution of mobile technology usage for Latino patients and comparing findings to national mHealth estimates. J Am Med Inform Assoc 2016 Sep;23(5):979-983 [FREE Full text] [doi: 10.1093/jamia/ocv203] [Medline: 26995564]

28. Hamilton EC, Saiyed F, Miller CC, Eguia A, Fonseca AC, Baum GP, et al. The digital divide in adoption and use of mobile health technology among caregivers of pediatric surgery patients. J Pediatr Surg 2018 Aug;53(8):1478-1493. [doi: 10.1016/j.jpedsurg.2017.08.023] [Medline: 28927983]

29. Yin RK. Case Study Research: Design and Methods. London: SAGE Publications Ltd; 2014.

30. Hak T, van DV, Jansen H. The three-step test-interview (TSTI): an observation-based method for pretesting self-completion questionnaires. Surv Res Methods 2008;2(3):143-150 [FREE Full text] [doi: 10.18148/srm/2008.v2i3.1669]

31. World Wide Web Consortium. Usability - ISO 9241 definition URL: http://www.w3.org/2002/Talks/0104-usabilityprocess/slide3-0.html [accessed 2018-07-18]

32. Frokjaer E, Hertzum M, Hornbaek K. Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated? In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. 2000 Presented at: CHI'00; April 1 - 6, 2000; The Hague, Netherlands p. 345-352. [doi: 10.1145/332040.332455]

33. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]

34. Sharp H, Rogers Y, Preece J. Interaction Design: Beyond Human–Computer Interaction. Second Edition. New York: John Wiley & Sons Inc; 2007.

35. Dumas JS, Redish JC. A Practical Guide To Usability Testing. Exeter: Intellect Ltd; 1999.

36. Fransen MP, van Schaik TM, Twickler TB, Essink-Bot ML. Applicability of internationally available health literacy measures in the Netherlands. J Health Commun 2011;16(Suppl 3):134-149. [doi: 10.1080/10810730.2011.604383] [Medline: 21951248]

37. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]

38. Chew LD, Griffin JM, Partin MR, Noorbaloochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. J Gen Intern Med 2008 May;23(5):561-566 [FREE Full text] [doi: 10.1007/s11606-008-0520-5] [Medline: 18335281]

39. Braun A, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

40. Nielsen J, Loranger H. Prioritizing Web Usability. Berkeley, California USA: New Riders; 2006.

41. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. J Am Med Assoc 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]

42. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. BMC Med Res Methodol 2014 Mar 25;14:42 [FREE Full text] [doi: 10.1186/1471-2288-14-42] [Medline: 24669751]

43. Joosten-Van ZE, Kocken P, Hoop de T. Onderzoek naar de effectiviteit van de inzet van allochtone zorgconsulenten in Rotterdamse huisartsenpraktijken in de zorg aan vrouwen van Turkse en Marokkaanse afkomst met stressgerelateerde pijnklachten. Rotterdam: GGD Rotterdam e.o. 2004. Het project Bruggen Bouwen URL: http://www.sozawe.rotterdam.nl/Rotterdam/Openbaar/Diensten/GGD/Pdf/Bieb/Bruggenbouwennov2004.pdf [accessed 2019-12-06]

44. Hahn EA, Cellal D, Dobrez DG, Shiomoto G, Taylor SG, Galvez AG, et al. Quality of life assessment for low literacy Latinos: a new multimedia program for self-administration. J Oncol Manag 2003;12(5):9-12. [Medline: 14619989]

45. Hahn EA, Du H, Garcia SF, Choi SW, Lai J, Victorson D, et al. Literacy-fair measurement of health-related quality of life will facilitate comparative effectiveness research in Spanish-speaking cancer outpatients. Med Care 2010 Jun;48(6 Suppl):S75-S82. [doi: 10.1097/MLR.0b013e3181d6f81b] [Medline: 20473208]

46. 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. [doi: 10.1200/JCO.1993.11.3.570] [Medline: 8445433]

47. Alonso J, Prieto L, Antó JM. [The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): an instrument for measuring clinical results]. Med Clin (Barc) 1995 May 27;104(20):771-776. [Medline: 7783470]

48. von Neumann J, Morgenstern O. Theory of Games and Economic Behavior. New Jersey: Princeton University Press; 1953.

49. Collins D. Pretesting survey instruments: an overview of cognitive methods. Qual Life Res 2003 May;12(3):229-238. [doi: 10.1023/a:1023254226592] [Medline: 12769135]

50. Hahn EA, Kallen MA, Jacobs EA, Ganschow PS, Garcia SF, Burns JL. English-Spanish equivalence of the Health Literacy Assessment Using Talking Touchscreen Technology (Health LiTT). J Health Commun 2014;19(Suppl 2):285-301. [doi: 10.1080/10810730.2014.936567] [Medline: 25315599]

51. Tourangeau R, Edwards B, Johnson T, Wolter K, Bates N. Hard-to-Survey Populations. Cambridge: Cambridge University Press; 2014.

52. 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 Apr 19;4(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4928] [Medline: 27095507]

53.   Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2015 Sep 7(9):CD002098 [FREE Full text] [doi: 10.1002/14651858.CD002098.pub2] [Medline: 26343551]

54.   Henderson C, Knapp M, Fernández JL, Beecham J, Hirani SP, Cartwright M, Whole System Demonstrator evaluation team. Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. Br Med J 2013 Mar 20;346:f1035 [FREE Full text] [doi: 10.1136/bmj.f1035] [Medline: 23520339]

55.   Gheorghe M, Wubulihasimu P, Peters F, Nusselder W, Van Baal PH. Health inequalities in the Netherlands: trends in quality-adjusted life expectancy (QALE) by educational level. Eur J Public Health 2016 Oct;26(5):794-799. [doi: 10.1093/eurpub/ckw043] [Medline: 27085191]

56.   Yost KJ, Webster K, Baker DW, Jacobs EA, Anderson A, Hahn EA. Acceptability of the talking touchscreen for health literacy assessment. J Health Commun 2010;15(Suppl 2):80-92 [FREE Full text] [doi: 10.1080/10810730.2010.500713] [Medline: 20845195]

Abbreviations

FACT-G: Functional Assessment of Cancer Therapy-General
mHealth: mobile health
SBSQ-D: Set of Brief Screening Questions-Dutch version
TSTI: Three-Step Test-Interview
TT: Talking Touchscreen
TTSQ: Talking Touch Screen Questionnaire

XSL•FO
RenderX

Original Paper

Development and Field Evaluation of the INTER-ACT App, a Pregnancy and Interpregnancy Coaching App to Reduce Maternal Overweight and Obesity: Mixed Methods Design

Annick Bogaerts[1,2,3], MSc, PhD; Margriet Bijlholt[1,3], MSc; Lotte Mertens[2], MSc; Marijke Braeken[2,4], PhD; Bart Jacobs[2], MSc; Bert Vandenberghe[5], MSc; Lieveke Ameye[1], PhD; Roland Devlieger[1,6,7], PhD, MD

[1]Department of Development and Regeneration, University of Leuven, Leuven, Belgium

[2]Faculty of Health and Social Work, Research Unit Resilient People, University Colleges Leuven-Limburg, Leuven, Belgium

[3]Faculty of Medicine and Health Sciences, Centre for Research and Innovation in Care, University of Antwerp, Antwerp, Belgium

[4]Faculty of Rehabilitation Sciences, Rehabilitation Research Center, Biomedical Research Institute, Hasselt University, Hasselt, Belgium

[5]Meaningful Interactions Lab (Mintlab), University of Leuven, Leuven, Belgium

[6]Department of Obstetrics and Gynecology, University Hospitals Leuven, Leuven, Belgium

[7]Department of Obstetrics, Gynecology and Fertility, Gasthuiszusters Antwerpen, Campus Sint-Augustinus, Wilrijk, Belgium

Corresponding Author:
Annick Bogaerts, MSc, PhD
Department of Development and Regeneration
University of Leuven
O&N IV Herestraat 49 - bus 805
Leuven, 3000
Belgium
Phone: 32 16 19 31 27
Email: annick.bogaerts@kuleuven.be

Abstract

Background: The interpregnancy and pregnancy periods are important windows of opportunity to prevent excessive gestational weight retention. Despite an overwhelming number of existing health apps, validated apps to support a healthy lifestyle between and during pregnancies are lacking.

Objective: To describe the development and evaluation of the INTER-ACT app, which is part of an interpregnancy and pregnancy lifestyle coaching module, to prevent excessive weight gain in pregnancy and enhance optimal weight and a healthy lifestyle in the interpregnancy period.

Methods: A mixed methods design was used to identify the needs of health care providers and end users, according to 15 semistructured interviews, two focus groups, and two surveys. The user interface was evaluated in a pilot study (N=9).

Results: Health care providers indicated that a mobile app can enhance a healthy lifestyle in pregnant and postpartum women. Pregnant women preferred graphic displays in the app, weekly notifications, and support messages according to their own goals. Both mothers and health care providers reported increased awareness and valued the combination of the app with face-to-face coaching.

Conclusions: The INTER-ACT app was valued by its end users because it was offered in combination with face-to-face contact with a caregiver.

Introduction

An increasing number of women are obese at the start of pregnancy. Concurrently, one in three European pregnant

women has excessive gestational weight gain [1]. In particular, women with a high pregestational body mass index (BMI), young women (<20 years), single women, and women belonging to ethnic minority groups are at risk [2]. Adverse outcomes

associated with maternal obesity and excessive gestational weight gain include gestational hypertension, gestational diabetes mellitus, and large-for-gestational age infants [3]. Approximately half of women with excessive gestational weight gain do not return to their prepregnancy BMI before the next pregnancy. This increases prepregnancy obesity and is an important predictor for increased risks of pregnancy- and birth-related outcomes in the next pregnancy, including cesarean delivery, fetal overgrowth, and postnatal weight retention [3-6].

Face-to-face lifestyle intervention studies during pregnancy are effective to reduce gestational weight gain [7-9], but they are time-consuming with limited scalability, and no or minimal effects have been shown regarding relevant pregnancy outcomes [10-12]. Given the high impact of prepregnancy BMI, intervening early during the preconception period is essential [3]. Reaching the most vulnerable women and subsequently achieving adherence to a healthy lifestyle before becoming pregnant are of high priority [13].

The use of mobile health (mHealth) technology in the prevention, screening, and treatment of health-related issues is increasing, as is reflected by the ample offering of smartphone apps. On one hand, mHealth can offer easier access to individually-tailored support at a low cost. On the other hand, these apps are mostly not targeted at groups with specific needs, such as pregnant and postnatal (between pregnancies) women. Moreover, their effectiveness has not been tested in randomized controlled trials (RCTs) [14]. Results of the effectiveness of mHealth tools are scarce [15,16], but pioneering studies have shown promising results regarding intervention adherence, feasibility, and achieving an adequate pregnancy weight gain [17-19].

The aim of this study was therefore twofold. First, we aimed to develop an app to monitor and coach pregnant and postnatal women with focus on maternal weight, physical activity, healthy eating, and mental wellbeing. Second, we aimed to gather feedback on user experience (ie, usability, usefulness, and user acceptance). This app, called INTER-ACT, will be used in combination with four postnatal (interpregnancy) and three prenatal face-to-face coaching sessions. The ultimate aim of an RCT, in which this app is embedded, is to reduce the risk of gestational hypertension, gestational diabetes, cesarean section, and large-for-gestational-age infants in subsequent pregnancy among women who had excessive gestational weight gain in their previous pregnancy [19].

Methods

Overview

The INTER-ACT app targets women during the interpregnancy period, as well as pregnant women. The interpregnancy period is defined as the period between delivery and the start of a subsequent pregnancy.

The app was developed in three stages (Figure 1). First, a mixed methods design was used to gain insights into experiences with and views on perinatal lifestyle coaching from the perspective of health care providers and women/end users. Second, the app was designed by user-experience researchers and developed by the Belgium Campus ITversity in South Africa. Third, the app was evaluated in a qualitative field evaluation study. The three stages are elaborated below. A subsequent stage that is beyond the scope of this study involves embedding the app in a lifestyle intervention and evaluating it with an RCT design. The content of the face-to-face coaching is described elsewhere [19].

Stage I: Insights From Caregivers and End Users

Health Care Providers' Perspectives

We conducted semistructured interviews with a purposive sample of four general practitioners, three gynecologists, five midwives, and three dieticians (Table 1), who were selected according to their previous experience with obesity care in pregnant and postnatal women. A topic list was developed to gain insight into their experiences with and views on perinatal lifestyle coaching and their attitude towards technology-supported lifestyle coaching. In addition, two focus groups with a total of 16 midwives were conducted to explore their experiences with and views on perinatal lifestyle coaching and their attitude towards technology-supported lifestyle coaching in order to support data triangulation and achieve data saturation. All interviews and focus groups were audiotaped, transcribed, and analyzed thematically using open coding. The analysis of focus groups additionally included a peer debriefing with our researchers to control the interpretation of the results. Furthermore, 43 caregivers (Table 2) attending a symposium about lifestyle coaching in pregnant women were asked to respond to two open questions about their knowledge and skills regarding perinatal lifestyle coaching and potential gaps. Written informed consent was obtained from all participants, and confidentiality and anonymity were assured. Ethical approval was obtained from University Hospital Universitair Ziekenhuis Leuven, Belgium (B300201422650).

Figure 1. Flow chart of the three stages of app development (data collection, app development, and field evaluation). RCT: randomized controlled trial.
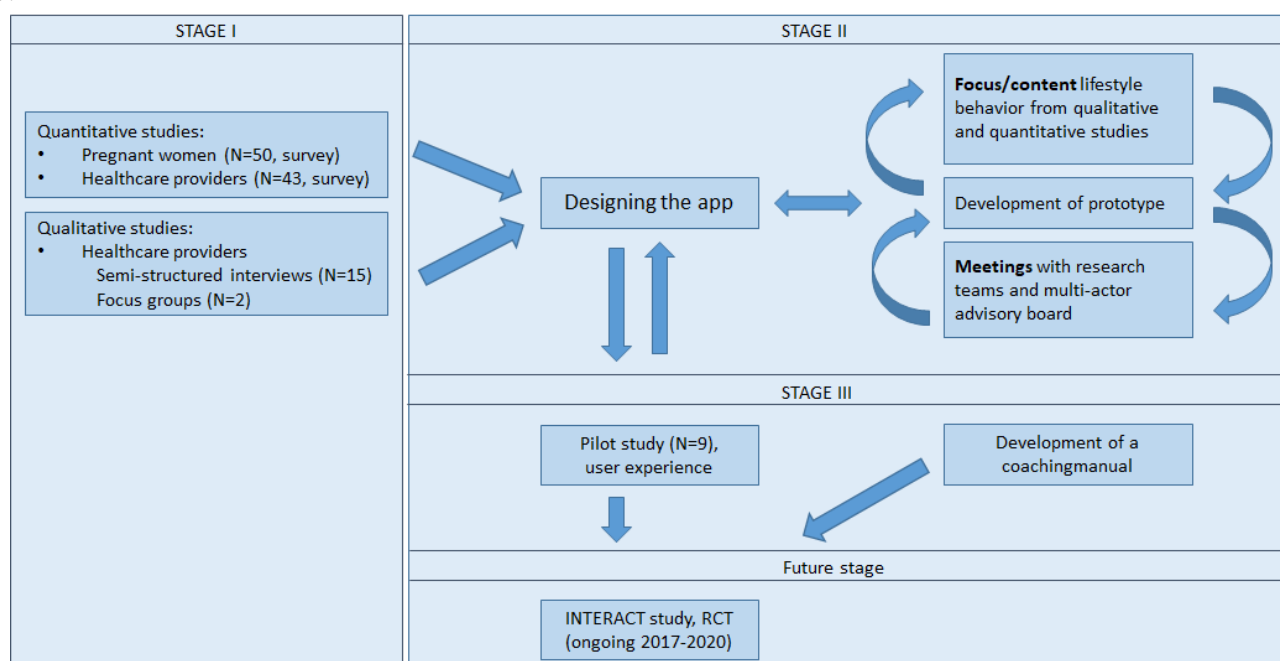
Table 1. Professions of health care providers participating in interviews and focus groups (N=31).

Profession	Frequency, n (%)
Midwife in primary care	13 (42)
Midwife in secondary care	5 (16)
Midwife in primary and secondary care	3 (10)
General practitioner	4 (13)
Gynecologist	3 (10)
Dietician	3 (10)

Table 2. Professions of health care providers participating in answering two open questions (N=43).

Profession	Frequency, n (%)
Midwife in primary care	3 (7)
Midwife in secondary care	16 (37)
Midwife (not in direct care)	9 (21)
Dietician	3 (7)
Nurse	3 (7)
Student	6 (14)
Teacher	3 (7)

End Users' Needs

We conducted a survey among 50 pregnant women between 12 and 42 weeks of pregnancy (Table 3) to explore their needs regarding technology-supported lifestyle coaching to optimize gestational weight gain. They were recruited from the waiting room before prenatal consultations in two nonuniversity hospitals. The inclusion criteria were as follows: sufficient fluency in spoken Dutch, age between 18 and 45 years, uncomplicated pregnancy between 12 and 42 weeks, and at least one prenatal consultation prior to the current consultation. The exclusion criteria were as follows: twin pregnancies, diagnosis of gestational diabetes or complications influencing physical activity or eating behavior. Ethical approval was obtained from University Hospital Universitair Ziekenhuis Leuven, Belgium (B243201628083).

Table 3. Characteristics of the survey participants (pregnant women) (N=50).

Characteristic	Value, n (%)
Gestational age (weeks)	
First trimester (0-14)	8 (16)
Second trimester (15-27)	10 (20)
Third trimester (28-40)	32 (64)
Gravidity	
Nullipara	33 (66)
Multipara	17 (34)
BMI[a] group	
Underweight (<18.5)	7 (14)
Normal weight (18.5-24.9)	23 (46)
Overweight (25-29.9)	12 (24)
Obesity class I (30-34.9)	4 (8)
Obesity class II (35-39.9)	1 (2)
Obesity class III (≥40)	2 (4)
Method of conception	
Spontaneous	44 (88)
Assisted reproduction	6 (12)
Age (years)	
18-24	5 (10)
25-29	23 (46)
30-34	16 (32)
35-39	5 (10)
40-44	1 (2)
Education	
Primary education	2 (4)
Secondary education	21 (42)
Bachelor's degree	15 (30)
Master's degree	13 (26)
Nationality	
Belgian	46 (92)
Dutch	2 (4)
Others	2 (4)
Marital status	
Married	30 (60)
Cohabiting	19 (38)
Single	1 (2)

[a]BMI: body mass index.

Stage II: App Development

The content of the app was based on the nutritional recommendations of the Superior Health Council of Belgium and the Institute of Medicine guidelines for gestational weight gain [20]. Additionall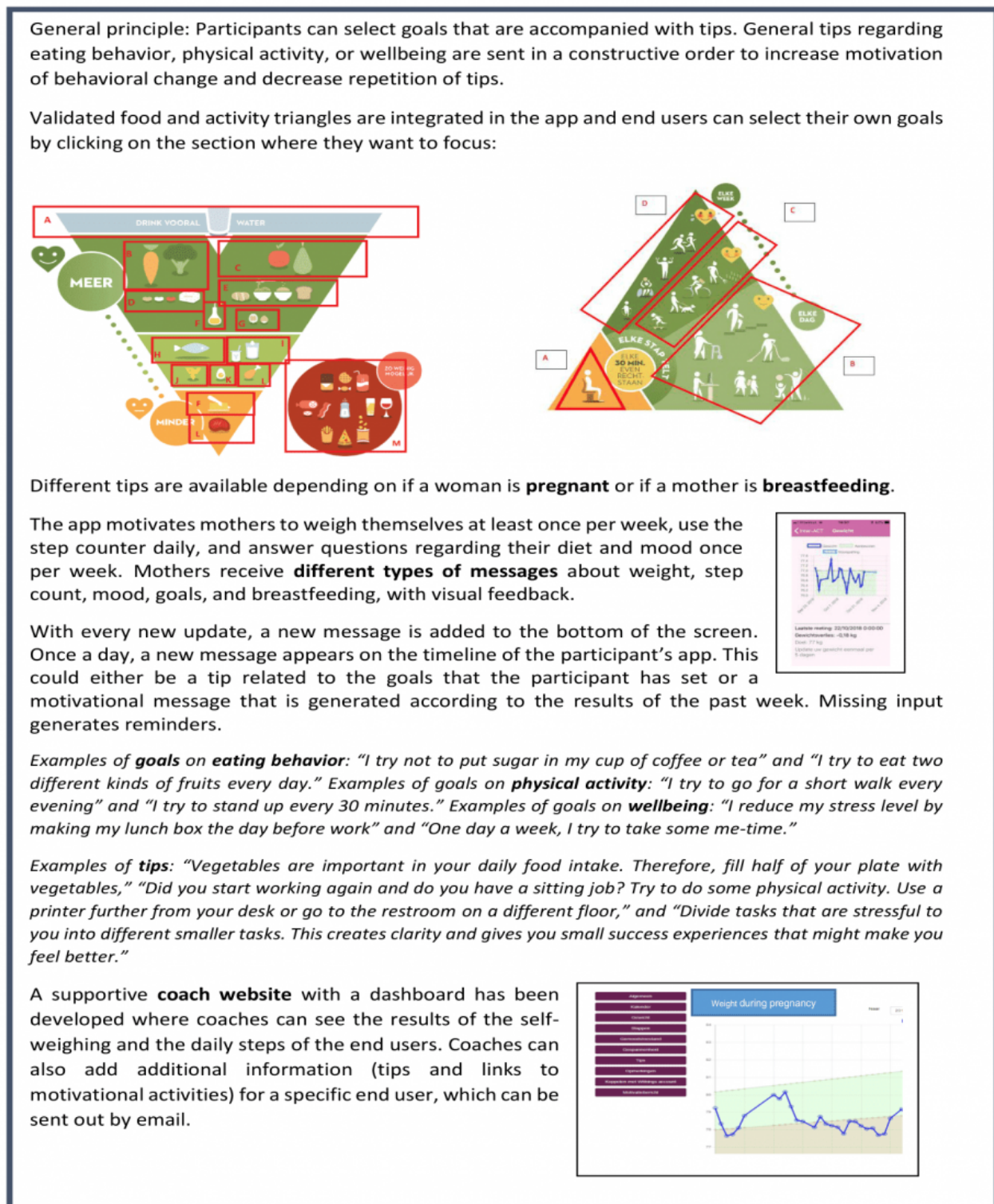y, guidelines from the Flemish Institute for Healthy Living and results from discussions with experts (clinicians, researchers, and policy makers) on the INTER-ACT external advisory board contributed to the content of the app. Furthermore, the principles of motivational interviewing techniques, goal setting, and positive messaging were incorporated in the app. User-experience researchers designed

the INTER-ACT app (Figure 2) according to usability heuristics, state-of-the-art insights from the domain of human-computer interaction, research experiences from previous mHealth projects and technologies [21], and results from the interviews and focus groups described in the first stage.

The participants could use INTER-ACT to monitor mental wellbeing, set goals on physical activity and healthy eating, and record progress on these goals. Additionally, Bluetooth connections were made with the Withings Go activity tracker (model WAM02; Withings, Issy-les-Moulineaux, France) and Withings Body+ weighing scale (model WBS05; Withings) in order to track physical activity and weight, respectively. Tips and motivating messages to support weight management, physical activity, healthy eating, and mental wellbeing were created, and an algorithm was developed to send these messages to the participants according to their input. Custom tips could be added by the researchers and sent to specific participants.

Figure 2. Description of the functionality of the INTER-ACT app and examples of messages.

General principle: Participants can select goals that are accompanied with tips. General tips regarding eating behavior, physical activity, or wellbeing are sent in a constructive order to increase motivation of behavioral change and decrease repetition of tips.

Validated food and activity triangles are integrated in the app and end users can select their own goals by clicking on the section where they want to focus:

Different tips are available depending on if a woman is **pregnant** or if a mother is **breastfeeding**.

The app motivates mothers to weigh themselves at least once per week, use the step counter daily, and answer questions regarding their diet and mood once per week. Mothers receive **different types of messages** about weight, step count, mood, goals, and breastfeeding, with visual feedback.

With every new update, a new message is added to the bottom of the screen. Once a day, a new message appears on the timeline of the participant's app. This could either be a tip related to the goals that the participant has set or a motivational message that is generated according to the results of the past week. Missing input generates reminders.

*Examples of **goals** on **eating behavior**: "I try not to put sugar in my cup of coffee or tea" and "I try to eat two different kinds of fruits every day." Examples of goals on **physical activity**: "I try to go for a short walk every evening" and "I try to stand up every 30 minutes." Examples of goals on **wellbeing**: "I reduce my stress level by making my lunch box the day before work" and "One day a week, I try to take some me-time."*

*Examples of **tips**: "Vegetables are important in your daily food intake. Therefore, fill half of your plate with vegetables," "Did you start working again and do you have a sitting job? Try to do some physical activity. Use a printer further from your desk or go to the restroom on a different floor," and "Divide tasks that are stressful to you into different smaller tasks. This creates clarity and gives you small success experiences that might make you feel better."*

A supportive **coach website** with a dashboard has been developed where coaches can see the results of the self-weighing and the daily steps of the end users. Coaches can also add additional information (tips and links to motivational activities) for a specific end user, which can be sent out by email.

The user interface of the app is designed according to the principle of conversational interfaces [22]. All content in the app is structured as a conversation between the user and the system in a chronological stream of messages (eg, a new step count or weight) (Figure 3). Messages are clickable, and clicking opens a page that provides additional information regarding the clicked message (eg, a weight graph). This approach allows the combination of both automatic input (from the weighing scale and activity tracker) and manual input (from entered mood), the display of feedback on achieved goals, and the display of reminders after a period of nonuse in a dynamic way.

Figure 3. Wireframes of the INTER-ACT app: stream of messages (left), coaching message (center), and manual input of mood (right).

The first prototype of the app was tested for functionality and feasibility by two pregnant women and a multidisciplinary team involving a professor of gynecology, a professor of midwifery, a biostatistician, a psychologist, two lifestyle coaches, and a group of app developers. They provided feedback regarding the design from medical, wellbeing, and technical perspectives. An iterative process of adaptation led to the development of the INTER-ACT app, which is being used in an ongoing RCT.

To ensure privacy and data security, the data are stored in a database hosted at a secure data center in Katholieke Universiteit Leuven. The database can only be accessed by our Application Programming Interface via a Structured Query Language connection. Security between the Application Programming Interface and end users involves a username and password system to keep the approach user friendly, but in the background, this system is supported by token-based authentication to prevent password theft. For the external system, we access authentication for data transfer via the Withings OAUTH2 system (Withings).

Stage III: Field Evaluation

The app was assessed in a qualitative field evaluation study, in which the technical functionality and user experience were explored. We recruited two pregnant and seven postnatal women (<6 months after delivery) through social media. During a home visit, the researchers installed the app, set up the Withings Go activity tracker and Withings Body+ weighing scale, and provided a short explanation of the app functions. The women used the app and devices for 3 weeks and were contacted at least once a week by telephone to address potential usability issues with the app and devices. In case of questions, the women could also contact the researchers by email and telephone. After the 3-week period, a semistructured interview was conducted during a home visit. Technical functionality issues, such as crashes, bugs, and connectivity issues with the activity tracker and weighing scale, were explored. The evaluation of user experience involved topics, such as content of knowledge- and skill-based elements; content, number, and timing of notifications; experienced accuracy of the activity tracker; and esthetics. The interviews were recorded and transcribed verbatim for analysis. The researchers' written notes of the observations made during the home visits, user feedback of the app, and reported user experiences were analyzed through an affinity diagram using Post-It notes. These insights allowed us to improve the app for a better user experience and prepare it for a full-scale field trial. Ethical approval was obtained for all studies, and informed consent was provided by all respondents (University Hospital Universitair Ziekenhuis Leuven, Belgium; B322201730956).

Results

Health Care Providers' Perspectives

Qualitative semistructured interviews and focus group discussions revealed health care provider–experienced barriers and facilitators, and perspectives on pregnancy and postpartum lifestyle coaching supported by mHealth. According to health care providers, low social background and educational levels, increased economic difficulties, ethnic minorities, different cultural or religious context, and insufficient knowledge about healthy eating were characteristics that needed attention in performing lifestyle coaching. The experienced facilitating factors were women's motivation to change lifestyle, awareness of their own responsibility, and self-control. Some health care providers were not convinced that an app would be effective in acquiring a healthier lifestyle among obese pregnant women, and they felt that it could even induce fear and anxiety. From the open questionnaires (n=43) and interviews (n=15), the following three themes for coaching emerged: (1) in-depth communication training; (2) motivational techniques; and (3) behavioral change training, with specific attention to sensitive communication for vulnerable groups, including insights on their religious and cultural contexts.

During the focus groups with midwives, they indicated a willingness to take up the role of a coach to empower women for a healthy lifestyle, but they lacked practical knowledge and skills to support vulnerable groups. They were not sure whether an app would be helpful in lifestyle coaching. However, if combined with face-to-face coaching and not used as a tool to "monitor and control" women's behavior, they indicated that an app could be useful. Data collected in the app could facilitate a coaching session and could result in a conversation about healthy lifestyle issues. However, midwives expressed that they prefer to restrict their administrative work and do not want to spend time on integration of additional technologies.

End Users' Needs

Among the 50 pregnant women who completed the survey, 30 (60%) wanted personal advice from caregivers about a healthy lifestyle. Only 8 out of 15 women (16%) indicated currently being counselled, mostly only regarding prenatal weight management (Table 4).

Additionally, 45 out of the 50 women (90%) indicated that an app would help them to maintain a healthy lifestyle. Among the 50 women, 46 (92%) were eager to monitor their calorie consumption and 28 (56%) were eager to monitor physical activity goals using an app or diary. Moreover, among the 50 women, 45 (90%) indicated that they would like to self-monitor their mental wellbeing using a Likert scale with emoticons and 39 (78%) indicated that encouraging messages might enhance their motivation. Furthermore, among the 50 women, 36 (72%) preferred the app to display and evaluate the actual weight and weight gain, including tailored feedback. All women preferred an app that could tell them what they could eat safely in pregnancy and that included food diaries, weekly shopping lists, and pictures with recommended portion sizes (Table 4).

The women indicated that the attractiveness of the app might be enhanced by the addition of features regarding fetal development, an agenda for prenatal appointments, a checklist with hospital necessities, information on health risks for the mother and child, the ability to upload pictures and ultrasounds, a contraction counter, and a kick counter. Finally, women reported that they want their partners to be involved in the use of the app.

Field Evaluation

The qualitative user evaluation study showed a high user acceptance of the system and reported an increased consciousness regarding physical activity, eating behavior, weight management, and mental wellbeing. The activity tracker, goal setting for nutrition, and regular push notifications were especially appreciated.

Multiple users requested to increase the number of notifications and suggested to spread them during the day instead of a single evening notification. Furthermore, users preferred to configure both the kind of reminder (steps, weight, mood, and goals based on the user's own behavior) and the timing.

Participants who had an app and device installed on their smartphones besides the INTER-ACT app made comparisons between the two apps (eg, comparisons were made regarding the accuracy of the activity tracker). Participants rarely felt that the Withings Go activity tracker was more accurate than their known devices (eg, Fitbit). There were no such remarks regarding the weighing scale. Participants reported missing certain functionalities that other health- and weight-related apps incorporate, such as sleep tracking, heart-rate monitoring, and advanced food tracking and calorie counting.

The esthetics of our study app were considered less modern or attractive when compared with today's standard. Despite these remarks, our participants noted important value in the INTER-ACT app when combined with face-to-face coaching.

Table 4. Results from the survey of pregnant women (N=50).

Factor	Value, n (%)
Preferred personal lifestyle advice	
Yes	30 (60)
No	20 (40)
Current lifestyle follow-up by a health care provider	
Yes	8 (16)
No	42 (84)
If yes, focus of current lifestyle follow-up by a health care provider	
Weight	4 (8)
Eating behavior	2 (4)
Physical activity	2 (4)
Desired frequency of lifestyle follow-up[a]	
Daily	0 (0)
Once a week	6 (12)
Once a month	26 (53)
On request	14 (29)
Other	3 (6)
A smartphone app might support a healthy lifestyle	
Yes	45 (90)
No	5 (10)
Preferred content of a smartphone app supporting lifestyle	
Eating behavior	46 (92)
Physical activity	28 (56)
Weight	38 (76)
Mental wellbeing	16 (32)
Preferred mHealth tools to support healthy eating	
Eating diary	35 (70)
Weekly shopping list	37 (74)
Pictures of portion sizes	31 (62)
List of allowed foods in pregnancy	50 (100)
Preferred frequency to complete an eating diary[a]	
Never	6 (12)
A few times a month	4 (8)
Once a week	3 (6)
A few times a week	10 (20)
Every day	26 (53)
Preferred method of self-monitoring of physical activity	
Pedometer	35 (70)
Registration of physical activity duration in a smartphone app	37 (74)
Preferred follow-up of mental wellbeing	
Registration on a Likert scale with emoticons	45 (90)
Receiving motivating messages	39 (78)

XSL•FO
RenderX

Factor	Value, n (%)
Preferred display of weight in the app	
Weight + weight gain	36 (72)
Only weight	11 (22)
Only weight gain	3 (6)
Preferred frequency of self-weighing	
Never	2 (4)
Once a week	32 (64)
Twice a week	4 (8)
3-6 times a week	6 (12)
Daily	6 (12)

[a]One survey was missing.

Discussion

Principal Findings

This paper reports on the development and evaluation of a mHealth app designed to help women improve their lifestyle during and between pregnancies. We found that pregnant women and health care providers valued the combination of the INTER-ACT app with face-to-face contact in supporting a healthy lifestyle. Personalized feedback from the system with different frequencies according to the focus of health behavior is highly appreciated and increases awareness about healthy behavior. Health care providers stress the importance of considering the vulnerability of risk groups within their cultural and religious contexts when introducing mHealth apps. On one hand, midwives were keen to improve knowledge and skills about sensitive communication and were interested in tools to enhance the intrinsic motivation for behavioral change. On the other hand, they reported reluctance to integrate new technologies fearing a high practical and administrative workload.

Comparison With Prior Work

Few studies have been published about app development processes for weight management in pregnant women [22,23]. Some studies focused on preconception health only [24,25]; however, to the best of our knowledge, there are no studies on app development targeting women in the interpregnancy period.

Participants in this study reported the need for mHealth as an addition to face-to face contact. This is supported by the findings in a recent RCT comparing the effectiveness of face-to-face contact, that of mHealth, and that of a combination of face-to-face contact with mHealth for 5% weight loss in an obese population. They concluded that a conventional face-to-face weight loss program can partially be replenished with an mHealth program without losing effectiveness [26].

A healthy prepregnancy BMI is an important indicator for optimal pregnancy and birth outcomes [27]. Reaching women with unhealthy lifestyles in due time is a challenge. The effects of preconception interventions for improving pregnancy outcomes in overweight and obese women are scarce [28].

Concurrently, health care providers indicate that they need more training and education about effective obesity communication and weight management practice [29,30]. Women themselves felt that tailored advice specific to their personal situation and weight monitoring would help them implement changes [31]. Both conclusions have been confirmed in this study.

Hence, we developed the INTER-ACT protocol consisting of a mHealth-supported lifestyle program [19]. The INTER-ACT app monitors women's weight and physical activity through connections with a weighing scale and activity tracker. Eating behavior and mental wellbeing were both self-reported. According to the data, algorithms provide continuous coaching through positive behavioral change techniques. The app targets women with excessive weight gain in a previous pregnancy and can be a low-cost alternative to labor-intensive face-to-face programs for the prevention of postnatal weight retention and excessive gestational weight gain in the subsequent pregnancy. Well-designed intervention trials with attention to structure, method of information delivery, and look and feel are required to further assess the feasibility and effects of such a technology for this target population.

A recent pilot mHealth-supported intervention study that included 40 postnatal women (6-16 weeks) showed that a higher intervention adherence was associated with greatly lower body weight and percentage body fat [32]. It is known that self-monitoring and increased intervention adherence are associated with increased weight loss [33,34]. Concurrently, Herring and colleagues [35] showed that peer support and interaction by social networking in the mHealth app can increase intervention adherence in urban low-income mothers. The high variability in intervention adherence in both mHealth- [32] and non–mHealth-supported lifestyle interventions [7] indicates that it is important to work on these barriers in the future through cocreation with end users.

Strengths and Limitations

A strength of this study is the mixed methods design used to explore the experiences and views of different health care providers, as well as pregnant women and mothers in the postnatal period. The iterative approach with user participation allowed us to adapt the content and functionality of the app.

Limitations are possible biases for the results because of the selection of experienced health care providers and motivated women in the pilot study. Besides, a rather short timeframe for the field evaluation of 3 weeks complicated the technical readiness of the app and thus could influence the crucial adherence and compliance of the program in the longer run. Furthermore, developing tailored feedback is complex and needs more time than was used in this approach to reach deeper levels. However, actual user evaluation showed that the INTER-ACT app increased the awareness for behavioral change.

Recommendations for upgrading the app include subsequent iterations with focus on graphical design, improving stability and performance, making notifications and reminders configurable, and achieving optimal adherence and compliance for using the app and coaching program. Furthermore, an RCT is needed to validate the app, including the coaching program, for long-term use and health-related outcomes.

Conclusion

Health care providers appreciate the INTER-ACT app in combination with face-to-face contact and emphasize the importance of paying attention to reach the most vulnerable groups, and they are keen on enhancing their sensitive communication skills. On the other hand, they are reluctant to take up additional administrative tasks and to handle technical issues that might be accompanied with the implementation of the INTER-ACT app.

Pregnant women and postnatal mothers value the combination of the INTER-ACT app with face-to-face coaching over more commercial and visually attractive apps. Technological readiness is crucial to refine the app before integration in an RCT. Future studies should evaluate the effectiveness of combinations of face-to-face programs and mHealth apps for this targeted population at risk.

Conflicts of Interest

None declared.

References

1. Devlieger R, Benhalima K, Damm P, Van Assche A, Mathieu C, Mahmood T, et al. Maternal obesity in Europe: where do we stand and how to move forward?: A scientific paper commissioned by the European Board and College of Obstetrics and Gynaecology (EBCOG). Eur J Obstet Gynecol Reprod Biol 2016 Jun;201:203-208. [doi: 10.1016/j.ejogrb.2016.04.005] [Medline: 27160501]

2. Bogaerts A, Van den Bergh B, Nuyts E, Martens E, Witters I, Devlieger R. Socio-demographic and obstetrical correlates of pre-pregnancy body mass index and gestational weight gain. Clin Obes 2012 Oct 14;2(5-6):150-159. [doi: 10.1111/cob.12004] [Medline: 25586250]

3. Catalano PM, Shankar K. Obesity and pregnancy: mechanisms of short term and long term adverse consequences for mother and child. BMJ 2017 Feb 08;356:j1 [FREE Full text] [doi: 10.1136/bmj.j1] [Medline: 28179267]

4. Bogaerts A, Van den Bergh BR, Ameye L, Witters I, Martens E, Timmerman D, et al. Interpregnancy Weight Change and Risk for Adverse Perinatal Outcome. Obstetrics & Gynecology 2013;122(5):999-1009. [doi: 10.1097/aog.0b013e3182a7f63e]

5. McBain RD, Dekker GA, Clifton VL, Mol BW, Grzeskowiak LE. Impact of inter-pregnancy BMI change on perinatal outcomes: a retrospective cohort study. Eur J Obstet Gynecol Reprod Biol 2016 Oct;205:98-104. [doi: 10.1016/j.ejogrb.2016.07.487] [Medline: 27567535]

6. Villamor E, Cnattingius S. Interpregnancy weight change and risk of adverse pregnancy outcomes: a population-based study. The Lancet 2006 Sep;368(9542):1164-1170. [doi: 10.1016/s0140-6736(06)69473-7]

7. Bogaerts AF, Devlieger R, Nuyts E, Witters I, Gyselaers W, Van den Bergh BR. Effects of lifestyle intervention in obese pregnant women on gestational weight gain and mental health: a randomized controlled trial. Int J Obes (Lond) 2013 Jun 2;37(6):814-821. [doi: 10.1038/ijo.2012.162] [Medline: 23032404]

8. Dodd JM, Turnbull D, McPhee AJ, Deussen AR, Grivell RM, Yelland LN, LIMIT Randomised Trial Group. Antenatal lifestyle advice for women who are overweight or obese: LIMIT randomised trial. BMJ 2014 Feb 10;348(feb10 3):g1285 [FREE Full text] [doi: 10.1136/bmj.g1285] [Medline: 24513442]

XSL•FO

RenderX

9. Thangaratinam S, Rogozinska E, Jolly K, Glinkowski S, Roseboom T, Tomlinson JW, et al. Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence. BMJ 2012 May 16;344(may16 4):e2088 [FREE Full text] [doi: 10.1136/bmj.e2088] [Medline: 22596383]

10. Dodd J, Deussen A, Mohamad I, Rifas-Shiman S, Yelland L, Louise J, et al. The effect of antenatal lifestyle advice for women who are overweight or obese on secondary measures of neonatal body composition: the LIMIT randomised trial. BJOG 2016 Jan 29;123(2):244-253 [FREE Full text] [doi: 10.1111/1471-0528.13796] [Medline: 26841217]

11. Poston L, Bell R, Croker H, Flynn AC, Godfrey KM, Goff L, et al. Effect of a behavioural intervention in obese pregnant women (the UPBEAT study): a multicentre, randomised controlled trial. The Lancet Diabetes & Endocrinology 2015 Oct;3(10):767-777. [doi: 10.1016/s2213-8587(15)00227-2]

12. Vinter CA, Jensen DM, Ovesen P, Beck-Nielsen H, Jørgensen JS. The LiP (Lifestyle in Pregnancy) study: a randomized controlled trial of lifestyle intervention in 360 obese pregnant women. Diabetes Care 2011 Dec 04;34(12):2502-2507 [FREE Full text] [doi: 10.2337/dc11-1150] [Medline: 21972411]

13. Poston L, Caleyachetty R, Cnattingius S, Corvalán C, Uauy R, Herring S, et al. Preconceptional and maternal obesity: epidemiology and health consequences. The Lancet Diabetes & Endocrinology 2016 Dec;4(12):1025-1036. [doi: 10.1016/s2213-8587(16)30217-0]

14. O'Donnell BE, Lewkowitz AK, Vargas JE, Zlatnik MG. Examining pregnancy-specific smartphone applications: what are patients being told? J Perinatol 2016 Oct 19;36(10):802-807. [doi: 10.1038/jp.2016.77] [Medline: 27195980]

15. Kennelly MA, Ainscough K, Lindsay K, Gibney E, Mc Carthy M, McAuliffe FM. Pregnancy, exercise and nutrition research study with smart phone app support (Pears): Study protocol of a randomized controlled trial. Contemp Clin Trials 2016 Jan;46:92-99. [doi: 10.1016/j.cct.2015.11.018] [Medline: 26625980]

16. Krishnamurti T, Davis AL, Wong-Parodi G, Fischhoff B, Sadovsky Y, Simhan HN. Development and Testing of the MyHealthyPregnancy App: A Behavioral Decision Research-Based Tool for Assessing and Communicating Pregnancy Risk. JMIR Mhealth Uhealth 2017 Apr 10;5(4):e42 [FREE Full text] [doi: 10.2196/mhealth.7036] [Medline: 28396302]

17. Willcox J, Wilkinson S, Lappas M, Ball K, Crawford D, McCarthy E, et al. A mobile health intervention promoting healthy gestational weight gain for women entering pregnancy at a high body mass index: the txt4two pilot randomised controlled trial. BJOG 2017 Oct 20;124(11):1718-1728 [FREE Full text] [doi: 10.1111/1471-0528.14552] [Medline: 28220604]

18. Redman LM, Gilmore LA, Breaux J, Thomas DM, Elkind-Hirsch K, Stewart T, et al. Effectiveness of SmartMoms, a Novel eHealth Intervention for Management of Gestational Weight Gain: Randomized Controlled Pilot Trial. JMIR Mhealth Uhealth 2017 Sep 13;5(9):e133 [FREE Full text] [doi: 10.2196/mhealth.8228] [Medline: 28903892]

19. Bogaerts A, Ameye L, Bijlholt M, Amuli K, Heynickx D, Devlieger R. INTER-ACT: prevention of pregnancy complications through an e-health driven interpregnancy lifestyle intervention - study protocol of a multicentre randomised controlled trial. BMC Pregnancy Childbirth 2017 May 26;17(1):154 [FREE Full text] [doi: 10.1186/s12884-017-1336-2] [Medline: 28549455]

20. Rasmussen KM, Abrams B, Bodnar LM, Butte NF, Catalano PM, Maria Siega-Riz A. Recommendations for Weight Gain During Pregnancy in the Context of the Obesity Epidemic. Obstetrics & Gynecology 2010;116(5):1191-1195. [doi: 10.1097/aog.0b013e3181f60da7]

21. Vanhoof JM, Vandenberghe B, Geerts D, Philippaerts P, De Mazière P, DeVito Dabbs A, PICASSO-Tx Consortium. Technology Experience of Solid Organ Transplant Patients and Their Overall Willingness to Use Interactive Health Technology. J Nurs Scholarsh 2018 Mar 28;50(2):151-162. [doi: 10.1111/jnu.12362] [Medline: 29193654]

22. Lau Y, Cheng LJ, Chi C, Tsai C, Ong KW, Ho-Lim SS, et al. Development of a Healthy Lifestyle Mobile App for Overweight Pregnant Women: Qualitative Study. JMIR Mhealth Uhealth 2018 Apr 23;6(4):e91 [FREE Full text] [doi: 10.2196/mhealth.9718] [Medline: 29685868]

23. 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 Jun 05;3(2):e44 [FREE Full text] [doi: 10.2196/mhealth.4071] [Medline: 26048313]

24. Materia FT, Smyth JM, Heron KE, Hillemeier M, Feinberg ME, Fonzi P, et al. Preconceptional health behavior change in women with overweight and obesity: prototype for SMART strong healthy women intervention. Mhealth 2018 Jul;4:24-24 [FREE Full text] [doi: 10.21037/mhealth.2018.06.06] [Medline: 30148139]

25. Nwolise CH, Carey N, Shawe J. Exploring the acceptability and feasibility of a preconception and diabetes information app for women with pregestational diabetes: A mixed-methods study protocol. Digit Health 2017 Aug 30;3:2055207617726418 [FREE Full text] [doi: 10.1177/2055207617726418] [Medline: 29942610]

26. Hurkmans E, Matthys C, Bogaerts A, Scheys L, Devloo K, Seghers J. Face-to-Face Versus Mobile Versus Blended Weight Loss Program: Randomized Clinical Trial. JMIR Mhealth Uhealth 2018 Jan 11;6(1):e14. [doi: 10.2196/mhealth.7713]

27. Goldstein RF, Abell SK, Ranasinha S, Misso M, Boyle JA, Black MH, et al. Association of Gestational Weight Gain With Maternal and Infant Outcomes: A Systematic Review and Meta-analysis. JAMA 2017 Jun 06;317(21):2207-2225 [FREE Full text] [doi: 10.1001/jama.2017.3635] [Medline: 28586887]

28. Opray N, Grivell R, Deussen A, Dodd J. Directed preconception health programs and interventions for improving pregnancy outcomes for women who are overweight or obese. Cochrane Database Syst Rev 2015 Jul 14(7):CD010932. [doi: 10.1002/14651858.CD010932.pub2] [Medline: 26171908]

XSL•FO
RenderX

29.   Heslehurst N, Russell S, McCormack S, Sedgewick G, Bell R, Rankin J. Midwives perspectives of their training and
      education requirements in maternal obesity: a qualitative study. Midwifery 2013 Jul;29(7):736-744. [doi:
      10.1016/j.midw.2012.07.007] [Medline: 22951422]
30.   Heslehurst N, Newham J, Maniatopoulos G, Fleetwood C, Robalino S, Rankin J. Implementation of pregnancy weight
      management and obesity guidelines: a meta-synthesis of healthcare professionals' barriers and facilitators using the Theoretical
      Domains Framework. Obes Rev 2014 Jun 16;15(6):462-486. [doi: 10.1111/obr.12160] [Medline: 24629076]
31.   Heslehurst N, Dinsdale S, Brandon H, Johnston C, Summerbell C, Rankin J. Lived experiences of routine antenatal dietetic
      services among women with obesity: A qualitative phenomenological study. Midwifery 2017 Jun;49:47-53. [doi:
      10.1016/j.midw.2016.11.001] [Medline: 27986354]
32.   Gilmore LA, Klempel MC, Martin CK, Myers CA, Burton JH, Sutton EF, et al. Personalized Mobile Health Intervention
      for Health and Weight Loss in Postpartum Women Receiving Women, Infants, and Children Benefit: A Randomized
      Controlled Pilot Study. J Womens Health (Larchmt) 2017 Jul;26(7):719-727 [FREE Full text] [doi: 10.1089/jwh.2016.5947]
      [Medline: 28338403]
33.   Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a
      randomized controlled trial. Obesity (Silver Spring) 2010 Feb 20;18(2):308-313 [FREE Full text] [doi: 10.1038/oby.2009.242]
      [Medline: 19696764]
34.   Brindal E, Freyne J, Saunders I, Berkovsky S, Smith G, Noakes M. Features predicting weight loss in overweight or obese
      participants in a web-based intervention: randomized trial. J Med Internet Res 2012 Dec 12;14(6):e173 [FREE Full text]
      [doi: 10.2196/jmir.2156] [Medline: 23234759]
35.   Herring SJ, Cruice JF, Bennett GG, Davey A, Foster GD. Using technology to promote postpartum weight loss in urban,
      low-income mothers: a pilot randomized controlled trial. J Nutr Educ Behav 2014 Nov;46(6):610-615 [FREE Full text]
      [doi: 10.1016/j.jneb.2014.06.002] [Medline: 25069621]

Abbreviations

BMI: body mass index
mHealth: mobile health
RCT: randomized controlled trial

Original Paper

Using an Electronic Medication Event–Monitoring System for Antiretroviral Therapy Self-Management Among African American Women Living With HIV in Rural Florida: Cohort Study

Robert Lucero[1], MPH, PhD, FAAN; Renessa Williams[1], BSN, AA; Tanisia Esalomi[2], BSN; Paula Alexander-Delpech[2], PhD; Christa Cook[3], PhD; Ragnhildur I Bjarnadottir[4], MPH, PhD

[1]Department of Family, Community, and Health Systems Science, Southern HIV & Alcohol Research Consortium, University of Florida, College of Nursing, Gainesville, FL, United States

[2]Department of Family, Community, and Health Systems Science, University of Florida, College of Nursing, Gainesville, FL, United States

[3]College of Nursing, University of Central Florida, Orlando, FL, United States

[4]Department of Family, Community, and Health Systems Science, Gainesville, FL, United States

Corresponding Author:
Robert Lucero, MPH, PhD, FAAN
Department of Family, Community, and Health Systems Science
Southern HIV & Alcohol Research Consortium
University of Florida, College of Nursing
1225 Center Drive
Gainesville, FL
United States
Phone: 1 3522736370
Email: rlucero@ufl.edu

Abstract

Background: HIV remains a significant health issue in the United States and disproportionately affects African Americans. African American women living with HIV (AAWH) experience a particularly high number of barriers when attempting to manage their HIV care, including antiretroviral therapy (ART) adherence. To enable the development and assessment of effective interventions that address these barriers to support ART adherence, there is a critical need to understand more fully the use of objective measures of ART adherence among AAWH, including electronic medication dispensers for real-time surveillance.

Objective: This study aimed to evaluate the use of the Wisepill medication event–monitoring system (MEMS) and compare the objective and subjective measures of ART adherence.

Methods: We conducted a 30-day exploratory pilot study of the MEMS among a convenience sample of community-dwelling AAWH (N=14) in rural Florida. AAWH were trained on the use of the MEMS to determine the feasibility of collecting, capturing, and manipulating the MEMS data for an objective measure of ART adherence. Self-reported sociodemographic information, including a self-reported measure of ART adherence, was also collected from AAWH.

Results: We found that the majority of participants were successful at using the electronic MEMS. Daily use of the MEMS tended to be outside of the usual time participants took their medication. Three 30-day medication event patterns were found that characterized ART adherence, specifically uniform and nonuniform medication adherence and nonuniform medication nonadherence. There were relatively few MEMS disruptions among study participants. Overall, adjusted daily ART adherence was 81.08% and subjective ART adherence was 77.78%.

Conclusions: This pilot study on the use and evaluation of the Wisepill MEMS among AAWH in rural Florida is the first such study in the United States. The findings of this study are encouraging because 10 out of 12 participants consistently used the MEMS, there were relatively few failures, and objective adjusted daily and overall subjective ART adherence were very similar. On the basis of these findings, we think researchers should consider using the Wisepill MEMS in future studies of AAWH and people living with HIV in the United States after taking into account our practical suggestions. The following practical considerations are suggested when measuring objective medication adherence: (1) before using an MEMS, be familiar with the targeted populations' characteristics; (2) choose an MEMS that aligns with the participants' day-to-day activities; (3) ensure the MEMS' features and resulting data support the research goals; (4) assess the match among the user's ability, wireless features of the MEMS, and the geographic location of the participants; and (5) consider the cost of MEMS and the research budget.

XSL•FO
RenderX

Introduction

Background

HIV remains a significant health issue in the United States and disproportionately affects African Americans [1]. African Americans represent 41% of all Americans living with the virus while comprising only 12% of the US population [2,3]. The burden of HIV is highest among African American women with HIV (AAWH) in the southern United States, including Florida, accounting for 63% of all cases in the past 10 years [4]. Among women diagnosed with HIV in 2014, 62% were African American women [3,5]. African American women also account for the largest share of deaths among women with HIV [6]. In 2010, HIV was the leading cause of death for African American women aged 25 to 44 years [7]. Moreover, the HIV mortality rate for African American women in that age group was 10.3 per 100,000 women compared with 0.7 per 100,000 among white women. This was second only to the rate among African American men. For African American women, there appears to be a complex intersection of race, class, and gender, making AAWH one of the most vulnerable groups in the United States [8,9].

The emergence of antiretroviral therapy (ART) has transformed HIV from an acute to a chronic condition, thus allowing individuals to live long lives with HIV [10]. To manage HIV, people living with HIV (PLWH) need to manage medication regimens that demand a high level of adherence. This is in addition to managing symptoms and side effects and various other challenges and barriers to self-management [10]. AAWH experience a particularly high number of barriers when attempting to manage their HIV care [9]. These include community stigma and lack of both general and disease-specific support. In addition, while financial issues, low income, lack of health insurance and other structural barriers in general affect minority populations, these factors disproportionately influence AAWH [9,10]. AAWH often go without care because of limited funds for food, clothing, housing, and other necessities, or postpone care because of lack of transportation [10]. AAWH are also more likely than their white female counterparts to experience unemployment. Those who are employed often report not being able to leave work for medical appointments. When AAWH do access care, a high proportion report not being referred to a case manager and not having enough time with their care provider [10]. Owing to these barriers, AAWH are more likely to rely on the emergency room to receive necessary care and are at higher odds of not receiving prescriptions for ART. In comparison to white women with HIV, a significantly lower proportion of AAWH receive ART or achieve viral suppression [8]. An estimated 50% of AAWH in Florida do not have a suppressed viral load [4]. The reasons for this disparity remain unclear.

To enable the development and assessment of effective interventions that address these barriers to support ART adherence, there is a critical need to understand more fully the use patterns of the objective measures of ART adherence among AAWH. A meta-analysis of the correlations of objective medication adherence via a medication–event monitoring system (MEMS) and self-reported questionnaire revealed that the mean of adherence measured by the MEMS was 74.9% (range 53.4%-92.9%) vs 84.0% by the self-reported questionnaire (range 68.35%-95%) among 11 studies and 1684 PLWH [11]. The correlation between two measures ranged from 0.24 to 0.87. The pooled correlation coefficient for the 11 studies was 0.45 (95% CI 0.34-0.56, *P*=.001), indicating a moderate relationship. There are few studies that report on the actual use patterns that underlie the objective measurement of ART adherence vis-à-vis an MEMS, and none that we know of that compare the objective and subjective ART adherence rates of AAWH [12-14]. Nonetheless, researchers have been steadily increasing the use of MEMS in research among diverse populations [15-22].

Research Questions

We sought to answer the following research questions in this pilot study:

1. What use patterns of the Wisepill MEMS emerge from the utilization of the system by AAWH?
2. Are there observable differences in an objective measure of ART adherence based on the Wisepill MEMS data and in a subjective measure of ART adherence based on self-reported data among AAWH?

Methods

Study Design

As part of a larger mixed method study, we conducted a 30-day pilot study of the Wisepill MEMS among a convenience sample of community-dwelling AAWH in rural Florida. We collected self-reported sociodemographic information and trained AAWH on the use of the MEMS to determine the feasibility of collecting, capturing, and manipulating the MEMS data. In the study reported here, we compared the observational MEMS and self-reported adherence data to address the stated research questions. The qualitative data obtained, not reported here, are being analyzed separately to address a research question related to MEMS use and HIV-related stigma.

Medication Event Monitoring via the Wisepill Dispenser

The Wisepill MEMS was chosen from among other MEMS based on the ability to organize daily medication events, as shown in Figure 1, and system design (ie, not a pill bottle) in an attempt to avoid inducing stigma. In this pilot study, we trained AAWH to use the Wisepill MEMS. A 1100 mA lithium polymer rechargeable battery provides power to the Wisepill RT2000, which holds approximately 30 large pills or 60 small pills in a seven-compartment inner container. Each time the compartment is opened, a cellular signal is sent and recorded

in real-time on a Web-based server. Each Wisepill device contains a subscriber identity module, and the transmission of data is primarily by general packet radio service to the server. Data transfer may also occur via SMS. However, general packet radio service is preferred to short message service because (1) it is less expensive and (2) the server deletes the data after it receives it. In addition to recording device openings, the Wisepill signal reports the remaining battery power for the device, airtime on the subscriber identity module, and strength of the signal. In a signaling subsystem, nonvolatile, electrical, erasable, programmable, read-only memory maintains data for later transmission if there is a power failure and connectivity is lost. The data are immediately accessible to research staff via a secure internet interface.

Figure 1. Wisepill RT2000 medication event–monitoring system.

Figure 2 depicts the flow of information captured and stored on the Wisepill secured server as well as information flow between the Wisepill client, or research site, and the Wisepill secured server. The research site programs each participant's prescription times using a unique identifier for each Wisepill user. In addition to automated downloadable reports from the Wisepill server, discreet medication events can be imported from the Wisepill server via a CSV file for individual-level analyses.

Figure 2. Wisepill medication event–monitoring system data capture and collection.

Recruitment

We recruited community-dwelling AAWH using two approaches: (1) searched two research registries (ie, the Florida Cohort Study and HealthStreet) that contain information on PLWH who may have consented to share their information for research opportunities and (2) collaborated with support groups that serve PLWH. The Florida Cohort Study began in 2014 and collects demographic, behavioral, and social factors affecting health outcomes for PLWH within the state of Florida. Any person with HIV aged 18 years or older is eligible to participate in the study. As part of the informed consent process, individuals can agree to participate in future research studies. In addition, HealthStreet is a community engagement program at the University of Florida that aims to improve the health of community members by bridging health care and health research. Community health workers assess health concerns, conditions, and research perceptions of community members, and provide referrals to community members for medical and social services, as well as opportunities to participate in health research. Finally, we attended HIV support group meetings regularly, presented our proposed study, and recruited interested individuals. The University of Florida institutional review board approved this study.

AAWH were eligible to participate in the study if they met the following criteria: (1) age between 18 and 55 years, (2) female, (3) on ART, (4) willing to use an electronic monitoring event system to monitor adherence to one of their ARTs, and (5) not an employee or a student of the University of Florida.

Participants received up to US $75 on a cash card for their time and effort upon completion of the 30-day study period.

Data Collection

Once a member of the research team determined a potential participant's eligibility, we met with the participant during an agreed upon date and time to obtain consent, conduct training on the use of the Wisepill MEMS, and collect baseline measures including sociodemographic characteristics. After completing the baseline measures and MEMS use training, the research team member provided the participant US $25 cash card for time and travel expenses. After 15 days of using the MEMS, a research team member followed up with participants by telephone to ask how the participant was doing with the use of the MEMS. We added US $25 to a participant's cash card when a member of the research team successfully contacted them by telephone. The telephone call was intended to troubleshoot potential technical barriers, and no additional data were collected from participants. At the end of the 30-day study period, a member of the research team met with each participant to collect the follow-up self-reported data. After the meeting was completed, we added a final US $25 to the participant's cash card.

Data Analysis

Baseline and follow-up self-reported data were stored in a secured Research Electronic Data Capture database at the University of Florida and exported to a Microsoft Excel file for analysis. The MEMS data are stored in the Wisepill secured server. Self-reported data were evaluated and reported using descriptive statistics appropriate for measurement level (eg, mean, median, standard deviations, frequencies, percent, and range). We checked for implausible or out-of-range values, distributional forms, and missingness.

Wisepill Use Patterns Among African American Women With HIV

We evaluated the participant's Wisepill MEMS data transmission to identify scheduled, unscheduled, and missed medication events as well as MEMS disruptions. *Disruptions* could occur because of battery failure, forwarder malfunction, or a participant's decision not to use the MEMS. *Scheduled medication events* were captured when a participant opened the MEMS during a 1-hour period that was typical of their daily dosing pattern. We used the 1-hour period based on the *7 rights* of safe medication preparation and administration, including right time [23]. A research team member programmed the scheduled medication event or events within the Wisepill research interface according to the participant's self-reported information. An *unscheduled medication event* occurred when a participant opened the MEMS outside of a scheduled medication event period. *Extra medication events* were logged when a participant opened the MEMS more than once during a scheduled medication event period. A *missed medication event* was recorded when participants neglected to open the MEMS during a scheduled medication event period. Using summary statistics (ie, range, mean, and median), we characterized the data transmission patterns of medication events (ie, scheduled, unscheduled, and extra), missed medication events, and disruptions for the entire sample.

We generated graphical dot plots of each participant's 30-day medication events, including scheduled, unscheduled, and extra medication events. After printing each dot plot, two members of the research team (RL and RW) reviewed the collection of dot plots on an open table to identify and create categories of patterns that could emerge from the 30-day medication events among study participants. We utilized a consensus process that included one other research team member (PD) to settle disagreement between RL and RW about a 30-day medication event pattern.

Objective and Subjective Antiretroviral Therapy Adherence Among African American Women With HIV

Objective Antiretroviral Therapy Adherence

Our primary outcome was an objective measure of ART adherence using the Wisepill MEMS. We measured objective ART adherence using two approaches: (1) scheduled ART adherence or the proportion of scheduled medication events during the study period (ie, total number of scheduled medication events/30 dosing events), and (2) daily ART adherence or the proportion of scheduled and unscheduled medication events during the study period (ie, total number of scheduled + unscheduled medication events/30 dosing events). In addition, given advances in wireless connectivity, it was unlikely that there would be any system failures. However, we also considered a separate calculation of objective ART adherence that would account for disruptions (ie, loss of wireless signal or technical failure) in receiving data. We censored technical failures and considered the absence of a clear technical failure a missed medication event [24].

Subjective Antiretroviral Therapy Adherence

Our secondary outcome measure was subjective ART adherence using a validated three-item self-report instrument that assesses adherence during the previous 30 days (Cronbach alpha: .86-.89) [25]. The three items include (1) How many days did you miss at least one dose of any of your HIV medicines? (response options: 0-30), (2) How good a job did you do at taking your HIV medicines in the way you were supposed to? (response options: very poor/poor/fair/good/very good/excellent), and (3) How often did you take your HIV medicines in the way you were supposed to? (response options: never/rarely/sometimes/usually/almost always/always). We calculated a proportion of the total number of days participants successfully took their medication based on the answer to question 1 (eg, two missed doses: 28/30=98%). We assigned adherence in 20% increments to the response categories in questions 2 and 3 (eg, very poor=0%, poor=20%, and fair=40%) [26]. We report the mean and median of overall adherence to each question and an overall summary of ART adherence by taking an average of the three questions answered and an aggregate across participants.

Results

Participant's Characteristics

As the eligibility criteria were applied directly to the existing consent-to-share registries (ie, the Florida Cohort Study and HealthStreet) and the criteria were shared at support group meetings, no one was excluded from the study because they were not eligible. We had a refusal rate of 76% across the three recruitment strategies. A total of 14 AAWH participated in this pilot study. All participants, except for one, completed the 30-day follow-up meeting with a member of the research team. Table 1 contains summary statistics of the study participants'

characteristics who completed the study. The average age of the study participants was 49 years with a range from 23 to 63 years. Nearly 46% (6/13) of the participants were not a high school graduate or did not have a general educational diploma. Only 3 of the participants were married or living with a long-term partner. As we were interested in personal technology use among our study participants, a set of internet use questions were included among baseline measures. It appears that at least 10 participants owned and used their own computer, laptop, or tablet, and 7 participants owned and used a mobile phone for internet activity. Notably, at least eight participants relied on public access computing to access the internet.

Table 1. Summary of study participants' characteristics (N=13).

Characteristics	Values
Age (years), mean (SD), range	48.9 (11.5), 23-63
Highest grade/year of school completed, n (%)	
Some high school (grades 9-12)	5 (39)
High school graduate or general education diploma	5 (39)
Some college or technical/trade school	1 (8)
College or trade school graduate	2 (15)
Marital status, n (%)	
Married	3 (23)
Divorced	2 (15)
Widowed	1 (8)
Separated	2 (15)
Never married/single	5 (39)
Living with a long-term partner	0 (0)
Primary use of the internet, n (%)	
Own computer/laptop/tablet at home	
Never	3 (23)
Rarely	1 (8)
About once a week	1 (8)
A few times a week	4 (31)
Daily	4 (31)
Computer at public locations (eg, library)	
Never	5 (39)
Rarely	5 (39)
About once a week	1 (8)
A few times a week	2 (15)
Daily	0 (0)
Mobile phone	
Never	5 (39)
Rarely	5 (39)
About once a week	1 (8)
A few times a week	2 (15)
Daily	0 (0)

Medication Event Patterns

There was 510 expected medication event days for the study sample, which includes multiple scheduled medication events per day for 3 participants. We observed 144 scheduled, 216 unscheduled, one extra, and 366 missed medication events among the 14 participants, not including disruptions. We also detected disruptions with two of the participant's use of the Wisepill MEMS. These disruptions were because of battery failures, which resulted in 41 disruption days.

We identified three patterns of 30-day medication events among the graphical dot plots of 12 participants and 14 scheduled medication events. The 2 participants who had a disrupted experience with the Wisepill MEMS were not included in this analysis. Figures 3-5 represent uniform and nonuniform medication adherence and nonuniform medication nonadherence, respectively. The grey horizontal bar indicates the time in which the participant reported usually taking their daily ART. Each dot is a scheduled, unscheduled, or extra medication event. In all, 5 participants had uniform medication adherence or took their ART mostly (ie, ≥24 events) during a uniform range of time (see Figure 3). Of these 5 participants, 1 had two uniform medication adherence patterns. A total of 5 participants had nonuniform medication adherence or mostly took (ie, ≥24 events) their ART but not in a uniform range of time (see Figure 4). Finally, 3 participants had nonuniform medication nonadherence or did not take their ART during a uniform range of time and missed a substantial number of ART events (ie, ≥10 events; see Figure 5). None of the participants had uniform medication adherence that was consistent with the time they reported usually taking their daily ART.

Figure 3. Uniform medication adherence.
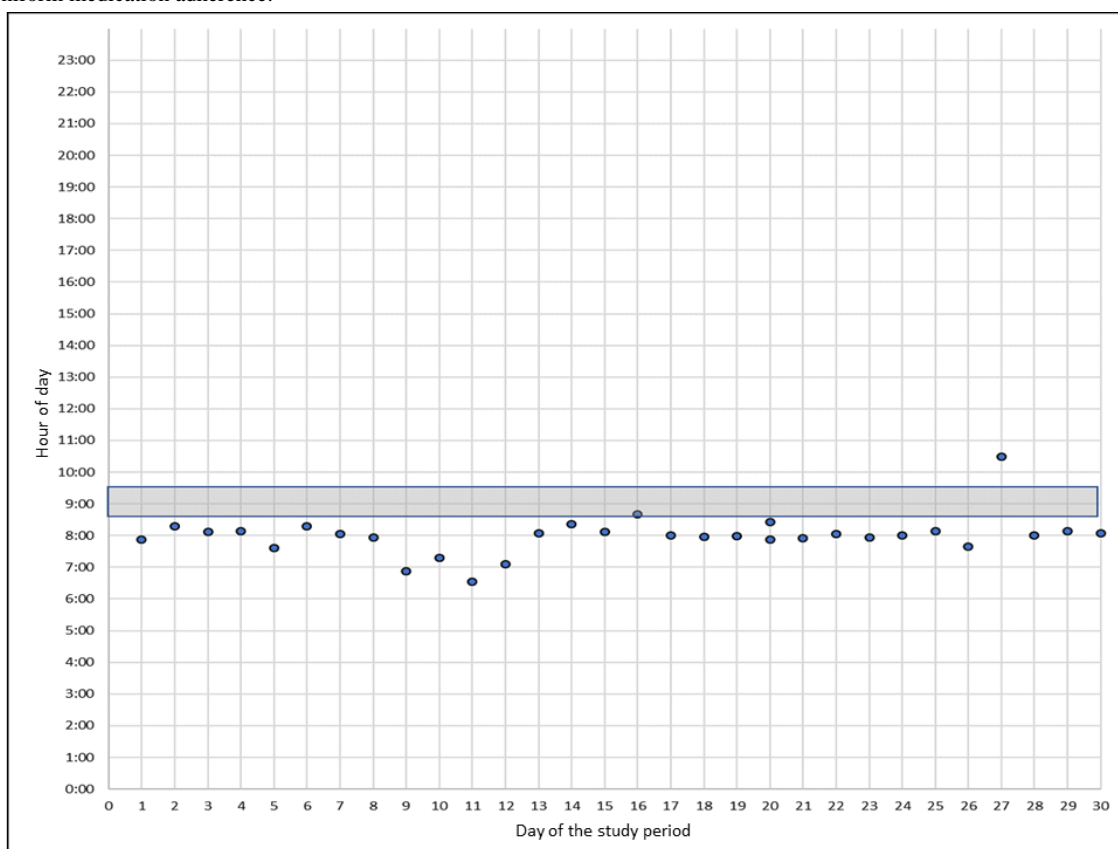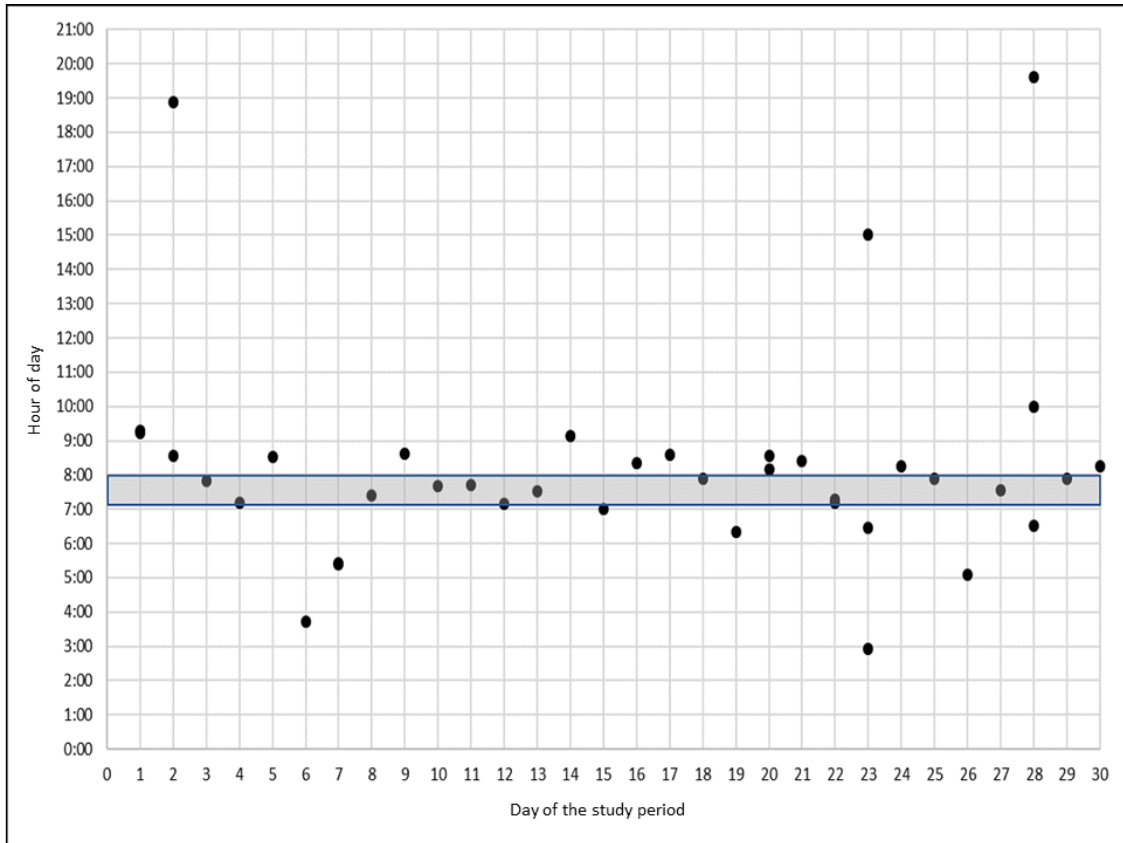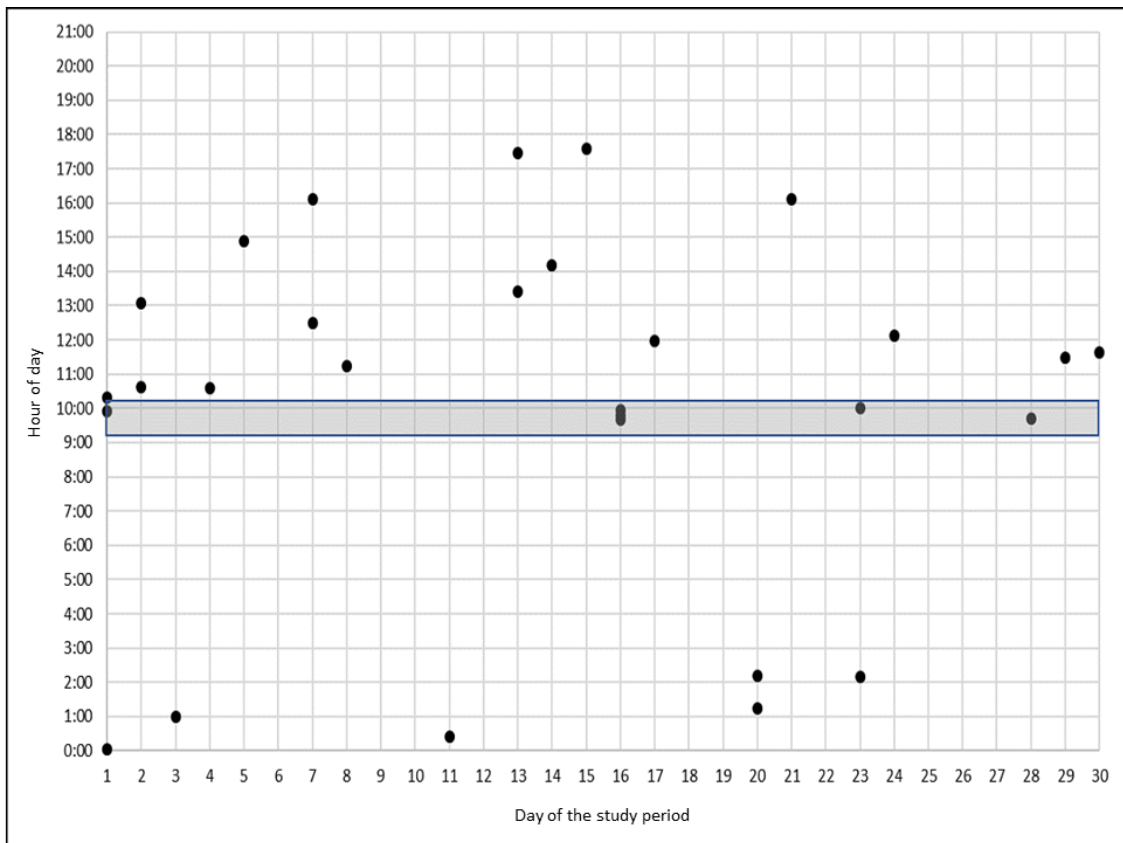
Figure 4. Nonuniform medication adherence.

Figure 5. Nonuniform medication nonadherence.

Antiretroviral Therapy Adherence

Table 2 contains a summary of the objective and subjective ART adherence measures we constructed from the observations captured using the Wisepill MEMS and the self-reported data of AAWH. There was a wide variation in the use of the MEMS, and extreme differences in the scheduled and daily ART adherence. On average, study participants used the MEMS with around 28% (3/10) of scheduled medication events during the study period. Some participants either chose not to or were unable to adhere to their usual medication schedule while others were able to with 67% (2/3) of events. However, based on adjusted scheduled and unscheduled daily medication events, participants on average used the MEMS at 81% (8/10) of events and some up to 100% for 30 days.

In comparison, the average number of days that participants self-reported not taking their ART was 2.05, or on average 93.17% adherence during the 30 days. Participants reported that they generally did a good job taking their ART 70.8% of the time. On the other hand, participants reported that sometimes (ie, 64.62%) they consistently took their ART as prescribed. On the basis of these self-reported behaviors, participants in general were 77.78% adherent to the ART.

Table 2. Objective and subjective antiretroviral therapy adherence among a pilot sample of African American women with HIV.

Adherence measures	Values
Objective measures of adherence using the Wisepill MEMS[a] for 30 days (N=12)	
Scheduled ART[b] adherence, mean (SD), range	
Unadjusted	28.24 (33.33), 0-66.67
Adjusted	28.24 (33.33), 0-66.67
Daily ART adherence, mean (SD), range	
Unadjusted	70.59 (83.33), 10.00-100.00
Adjusted	81.08 (86.67), 26.67-100.00
Three-item self-report measure of adherence[c](N=13)	
Overall subjective ART adherence, mean (SD), range	77.78 (80.00), 56.67-100.00
Individual-item subjective ART adherence	
1. In the last 30 days, on how many days did you miss at least one dose of any of your antiretroviral medication?[d], n	
Did not miss	9
1 day	2
2 days	0
3 days	2
Overall adherence for item 1 (0-100), mean (median)[e]	97.95 (100)
2. In the last 30 days, how good a job did you do at taking your antiretroviral medication in the way that you were supposed to?, n	
Very poor	0
Poor	0
Fair	4
Good	2
Very good	3
Excellent	4
Overall adherence for item 2 (0-100), mean (median)[e]	70.77 (80)
3. In the last 30 days, how often did you take your antiretroviral medication in the way that you were supposed to?, n	
Never	1
Rarely	1
Sometimes	1
Usually	3
Almost always	5
Always	2
Overall adherence for item 3 (0-100), mean (median)[e]	64.62 (80)

[a]MEMS: medication event–monitoring system.

[b]ART: antiretroviral therapy.

[c]The total number of responses for each question are based on all participants except for one who did not complete the end-of-study data collection meeting.

[d]The range of responses for this question is based on the minimum and maximum missed doses reported by participants.

[e]We calculated a proportion of the total number of days participants successfully took their medication in question 1 (eg, two missed doses: 28/30=98.3%). We assigned adherence in 20% increments to the response categories in questions 2 and 3 (eg, very poor=0%, poor=20%, and fair=40%).

Discussion

Principal Findings

We found among a sample of AAWH in rural Florida who were mostly high school graduates, unmarried, and owned some consumer technology that a majority were successful at using an electronic MEMS. Even though participants had a typical time to take their ART, their daily use of the MEMS tended to be outside of this usual time. None of the participants used the MEMS 100% of the time for scheduled medication events compared with 4 participants who successfully used the device 100% for daily medication events. This variation in medication administration was also identified among the studies including in the meta-analysis of ART adherence by Shi et al [11]. The barriers that have been reported to influence ART adherence may also modify the ability of AAWH to maintain a typical time of day to take their medication.

We identified three 30-day medication event patterns that characterized ART adherence among participants, namely, uniform and nonuniform medication adherence and nonuniform medication nonadherence. This finding could be important for the development of targeted self-management interventions based on the three use patterns [27-29]. In other words, the identification of an underlying medication event pattern can inform goal setting, action planning, or problem solving based on some observed reality.

Although few participants self-reported less than 100% ART adherence, a much smaller proportion thought they did a good job repeatedly taking ART as prescribed. These findings are nearly the same as reported by Shi et al [11]. On average, the self-reported ART adherence in this study was similar to the adjusted objective ART adherence. Unlike other studies, on average, our participants had greater objective ART adherence than self-reported ART adherence. This may be, in part, because of the difference in creating a quantitative measure from categorical data.

Limitations

This study provides important foundational insights into the use of an electronic MEMS to determine an objective measure of ART adherence among African American women who are disproportionately affected by HIV. However, there were limitations to our pilot study, and we acknowledge the small sample size. This was the first study conducted by the authors using the Wisepill MEMS. This may have unduly led to disruptions or technical failures that might not have otherwise occurred with experienced investigators or users. Second, we recruited a convenience sample of AAWH to use the MEMS. The participants may have been motivated to use the MEMS because of social desirability, which can occur when an individual's behavior is favorable to others. However, given the variation in the scheduled and daily ART adherence and self-reported ART adherence summary measures, there appears to be minimal influence on the participant's behavior by intrinsic (eg, desire to appear more adherent to ART than usual) or extrinsic (ie, baseline and 15- and 30-day US $25 compensation) motivators.

Comparison With Previous Work

Previous studies that have focused on the feasibility of using the Wisepill MEMS to monitor ART adherence occurred mostly in settings outside of the United States [14]. In Uganda, researchers found that the Wisepill MEMS produced similar results as that of MEMS pill bottle cap, and male (n=2) and female (n=8) study participants described the device as easy to use and convenient [24]. Another group of researchers in China found that, although using the Wisepill MEMS for real-time medication monitoring was technically feasible with men (n=2) and women (n=8), there were concerns regarding the acceptability of the device to patients [13]. Only half of all participants reported positive experiences. Researchers in Tanzania showed that real-time medication monitoring using the Wisepill MEMS was both a feasible and acceptable way to measure ART adherence among men (n=2) and women (n=3) [12].

To the best of our knowledge, ours is the first study to report empirical data on system use of the Wisepill MEMS in the United States even though the device is being used increasingly by researchers to construct an objective measure of ART adherence [15-22]. Similar to studies conducted in other countries, our experience was not a perfect one. Not all of our participants appeared to find the Wisepill MEMS easy to use. This may have been partly because of the battery failures we identified and also because none of our participants used the MEMS 100% of the times during scheduled events. Therefore, pretesting devices is necessary before releasing the MEMS to study participants. The fact that the majority of our participants were not able to use the MEMS uniformly could suggest something about the overall usability of the device's design.

A recent laboratory study assessed the accuracy of 10 commercially available MEMS, including Wisepill [30]. The researchers measured the accuracy of three devices and defined this parameter as scheduled events that fell within 120 seconds of the date and time recorded on paper for three devices. Of the 10 MEMS, 7 accurately registered ≥96% of the scheduled events across the three devices while the Wisepill device did so with accuracy at 100%, 92%, and 84%. We concur with McGrady's [30] conclusion that the *best* MEMS depends on the research study and population sample. We also suggest the following practical considerations when measuring objective medication adherence: (1) before using the MEMS, be familiar with the targeted populations' characteristics; (2) choose an MEMS that aligns with the participants' day-to-day activities; (3) ensure the MEMS' features and resulting data support the research goals; (4) assess the match among the user's ability, wireless features of the MEMS, and the geographic location of the participants; and (5) consider the cost of MEMS and the research budget. Although we did not necessarily follow these simple considerations before using the Wisepill MEMS, our study results are encouraging because 10 out of 12 participants had consistent use of the MEMS, there were relatively few disruptions in the device use, and objective adjusted daily and overall subjective ART adherence were very similar in our study.

Conclusions

This pilot study on the use and evaluation of the Wisepill MEMS among AAWH in rural Florida is the first such study in the United States. We found that AAWH were generally successful at using the MEMS for 30 days. Overall adjusted daily ART adherence was 81.08% and subjective ART adherence was 77.78%. The use of the MEMS among study participants resulted in three clear patterns of behavior: uniform and nonuniform medication adherence and nonuniform medication nonadherence. In summary, we think researchers should consider using the Wisepill MEMS in future studies of AAWH and PLWH in the United States after considering our practical suggestions.

Conflicts of Interest

None declared.

References

1. Centers for Disease Control and Prevention. 2018. HIV Surveillance Reports: Diagnoses of HIV Infection in the United States and Dependent Areas, 2018 URL: https://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2018-08-19]
2. Centers for Disease Control and Prevention. 2016. Living with HIV URL: https://www.cdc.gov/hiv/basics/livingwithhiv/ [accessed 2016-11-23]
3. Centers for Disease Control and Prevention. 2016. HIV and African Americans URL: https://www.cdc.gov/hiv/group/racialethnic/africanamericans/index.html [accessed 2018-08-19]
4. Florida Department of Health. 2018. AIDS Surveillance: Epidemiologic Profile URL: http://www.floridahealth.gov/diseases-and-conditions/aids/surveillance/epi-profiles/index.html [accessed 2018-08-27]
5. Centers for Disease Control and Prevention. 2018. HIV Among Women URL: http://www.cdc.gov/hiv/group/gender/women/index.html [accessed 2018-08-27]
6. National Center for Health Statistics. Centers for Disease Control and Prevention (CDC). 2017. Health, United States, 2016: With Chartbook on Long-Term Trends in Health URL: https://www.cdc.gov/nchs/data/hus/hus16.pdf [accessed 2020-02-10]
7. National Center for Health Statistics. Centers for Disease Control and Prevention (CDC). 2013. Health, United States, 2012: With Special Feature on Emergency Care URL: https://www.cdc.gov/nchs/data/hus/hus12.pdf [accessed 2020-02-10]
8. Cohen MH, Cook JA, Grey D, Young M, Hanau LH, Tien P, et al. Medically eligible women who do not use HAART: the importance of abuse, drug use, and race. Am J Public Health 2004 Jul;94(7):1147-1151. [doi: 10.2105/ajph.94.7.1147] [Medline: 15226135]
9. Toth M, Messer LC, Quinlivan EB. Barriers to HIV care for women of color living in the Southeastern US are associated with physical symptoms, social environment, and self-determination. AIDS Patient Care STDS 2013 Nov;27(11):613-620 [FREE Full text] [doi: 10.1089/apc.2013.0030] [Medline: 24138485]
10. Aziz M, Smith KY. Challenges and successes in linking HIV-infected women to care in the United States. Clin Infect Dis 2011 Jan 15;52(Suppl 2):S231-S237. [doi: 10.1093/cid/ciq047] [Medline: 21342912]
11. Shi L, Liu J, Fonseca V, Walker P, Kalsekar A, Pawaskar M. Correlation between adherence rates measured by MEMS and self-reported questionnaires: a meta-analysis. Health Qual Life Outcomes 2010 Sep 13;8:99 [FREE Full text] [doi: 10.1186/1477-7525-8-99] [Medline: 20836888]
12. de Sumari-de Boer IM, van den Boogaard J, Ngowi KM, Semvua HH, Kiwango KW, Aarnoutse RE, et al. Feasibility of real time medication monitoring among HIV infected and TB patients in a resource-limited setting. AIDS Behav 2016 May;20(5):1097-1107. [doi: 10.1007/s10461-015-1254-0] [Medline: 26604004]
13. Desilva MB, Gifford AL, Keyi X, Li Z, Feng C, Brooks M, et al. Feasibility and acceptability of a real-time adherence device among HIV-positive IDU patients in China. AIDS Res Treat 2013;2013:957862 [FREE Full text] [doi: 10.1155/2013/957862] [Medline: 23956851]
14. Musiimenta A, Atukunda EC, Tumuhimbise W, Pisarski EE, Tam M, Wyatt MA, et al. Acceptability and feasibility of real-time antiretroviral therapy adherence interventions in rural Uganda: mixed-method pilot randomized controlled trial. JMIR Mhealth Uhealth 2018 May 17;6(5):e122 [FREE Full text] [doi: 10.2196/mhealth.9031] [Medline: 29773527]

XSL•FO
RenderX

15.   Blashill AJ, Safren SA, Wilhelm S, Jampel J, Taylor SW, O'Cleirigh C, et al. Cognitive behavioral therapy for body image and self-care (CBT-BISC) in sexual minority men living with HIV: A randomized controlled trial. Health Psychol 2017 Oct;36(10):937-946 [FREE Full text] [doi: 10.1037/hea0000505] [Medline: 28541068]

16.   Castel AD, Qasmieh S, Greenberg D, Ellenberger N, Howell TH, Griffith C, et al. Digital gaming to improve adherence among adolescents and young adults living with HIV: mixed-methods study to test feasibility and acceptability. JMIR Serious Games 2018 Oct 15;6(4):e10213 [FREE Full text] [doi: 10.2196/10213] [Medline: 30322838]

17.   Craker L, Tarantino N, Whiteley L, Brown L. Measuring antiretroviral adherence among young people living with HIV: observations from a real-time monitoring device versus self-report. AIDS Behav 2019 Aug;23(8):2138-2145. [doi: 10.1007/s10461-019-02448-1] [Medline: 30888573]

18.   Koss CA, Hosek SG, Bacchetti P, Anderson PL, Liu AY, Horng H, et al. Comparison of measures of adherence to human immunodeficiency virus preexposure prophylaxis among adolescent and young men who have sex with men in the United States. Clin Infect Dis 2018 Jan 6;66(2):213-219 [FREE Full text] [doi: 10.1093/cid/cix755] [Medline: 29020194]

19.   Mayer KH, Safren SA, Elsesser SA, Psaros C, Tinsley JP, Marzinke M, et al. Optimizing pre-exposure antiretroviral prophylaxis adherence in men who have sex with men: results of a pilot randomized controlled trial of 'Life-Steps for PrEP'. AIDS Behav 2017 May;21(5):1350-1360 [FREE Full text] [doi: 10.1007/s10461-016-1606-4] [Medline: 27848089]

20.   Mimiaga MJ, Kuhns LM, Biello KB, Olson J, Hoehnle S, Santostefano CM, et al. Positive STEPS - a randomized controlled efficacy trial of an adaptive intervention for strengthening adherence to antiretroviral HIV treatment among youth: study protocol. BMC Public Health 2018 Jul 13;18(1):867 [FREE Full text] [doi: 10.1186/s12889-018-5815-9] [Medline: 30001703]

21.   Pellowski JA, Kalichman SC, White D, Amaral CM, Hoyt G, Kalichman MO. Real-time medication adherence monitoring intervention: test of concept in people living with HIV infection. J Assoc Nurses AIDS Care 2014;25(6):646-651 [FREE Full text] [doi: 10.1016/j.jana.2014.06.002] [Medline: 25043931]

22.   Whiteley L, Brown L, Lally M, Heck N, van den Berg JJ. A mobile gaming intervention to increase adherence to antiretroviral treatment for youth living with HIV: development guided by the information, motivation, and behavioral skills model. JMIR Mhealth Uhealth 2018 Apr 23;6(4):e96 [FREE Full text] [doi: 10.2196/mhealth.8155] [Medline: 29685863]

23.   Smeulers M, Verweij L, Maaskant JM, de Boer M, Krediet CT, van Dijkum EJ, et al. Quality indicators for safe medication preparation and administration: a systematic review. PLoS One 2015;10(4):e0122695 [FREE Full text] [doi: 10.1371/journal.pone.0122695] [Medline: 25884623]

24.   Haberer JE, Kahane J, Kigozi I, Emenyonu N, Hunt P, Martin J, et al. Real-time adherence monitoring for HIV antiretroviral therapy. AIDS Behav 2010 Dec;14(6):1340-1346 [FREE Full text] [doi: 10.1007/s10461-010-9799-4] [Medline: 20809380]

25.   Wilson IB, Lee Y, Michaud J, Fowler FJ, Rogers WH. Validation of a new three-item self-report measure for medication adherence. AIDS Behav 2016 Nov;20(11):2700-2708 [FREE Full text] [doi: 10.1007/s10461-016-1406-x] [Medline: 27098408]

26.   Musinguzi N, Muganzi CD, Boum Y, Ronald A, Marzinke MA, Hendrix CW, Partners PrEP Ancillary Adherence Study Team. Comparison of subjective and objective adherence measures for preexposure prophylaxis against HIV infection among serodiscordant couples in East Africa. AIDS 2016 Apr 24;30(7):1121-1129. [doi: 10.1097/QAD.0000000000001024] [Medline: 26785125]

27.   Cooper V, Clatworthy J, Whetham J, Consortium E. mHealth interventions to support self-management in HIV: a systematic review. Open AIDS J 2017;11:119-132 [FREE Full text] [doi: 10.2174/1874613601711010119] [Medline: 29290888]

28.   Gifford AL, Groessl EJ. Chronic disease self-management and adherence to HIV medications. J Acquir Immune Defic Syndr 2002 Dec 15;31(Suppl 3):S163-S166. [doi: 10.1097/00126334-200212153-00016] [Medline: 12562043]

29.   Lucero RJ, Frimpong JA, Fehlberg EA, Bjarnadottir RI, Weaver MT, Cook C, et al. The relationship between individual characteristics and interest in using a mobile phone app for HIV self-management: observational cohort study of people living with HIV. JMIR Mhealth Uhealth 2017 Jul 27;5(7):e100 [FREE Full text] [doi: 10.2196/mhealth.7853] [Medline: 28751298]

30.   McGrady ME, Holbein CE, Smith AW, Morrison CF, Hommel KA, Modi AC, et al. An independent evaluation of the accuracy and usability of electronic adherence monitoring devices. Ann Intern Med 2018 Sep 18;169(6):419-422 [FREE Full text] [doi: 10.7326/M17-3306] [Medline: 29800004]

Abbreviations

AAWH: African American women with HIV
ART: antiretroviral therapy
MEMS: medication event–monitoring system
NIH: National Institutes of Health
PLWH: people living with HIV

XSL•FO

RenderX

XSL•FO
RenderX

Original Paper

A Patient-Centered Information System (myED) for Emergency Care Journeys: Design, Development, and Initial Adoption

Monika Westphal[1], BA, MSc; Galit Bracha Yom-Tov[1], BSc, MSc, PhD; Avi Parush[1], BA, MA, PhD; Nitzan Carmeli[1], BSc, MSc; Alina Shaulov[1], BA; Chen Shapira[1], MD; Anat Rafaeli[1], BA, MA, PhD

Technion - Israel Institute of Technology, Haifa, Israel

Corresponding Author:
Galit Bracha Yom-Tov, BSc, MSc, PhD
Technion - Israel Institute of Technology
Technion City
Haifa, 3200003
Israel
Phone: 972 48294510
Email: gality@technion.ac.il

Abstract

Background: Medical care is highly complex in that it addresses patient-centered health goals that require the coordination of multiple care providers. Emergency department (ED) patients currently lack a sense of predictability about ED procedures. This increases frustration and aggression. Herein, we describe a system for providing real-time information to ED patients regarding the procedures in their ED medical journey.

Objective: This study aimed to develop a system that provides patients with dynamically updated information about the specific procedures and expected waiting times in their personal ED journey, and to report initial evaluations of this system.

Methods: To develop the *myED* system, we extracted information from hospital databases and translated it using process mining and user interface design into a language that is accessible and comprehensible to patients. We evaluated the system using a mixed methods approach that combined observations, interviews, and online records.

Results: Interviews with patients, accompanying family members, and health care providers (HCPs) confirmed patients' needs for information about their personal ED journey. The system developed enables patients to access this information on their personal mobile phones through a responsive website. In the third month after deployment, 492 of 1614 (30.48%) patients used *myED*. Patients' understanding of their ED journey improved significantly ($F_{8,299}$=2.519; P=.01), and patients showed positive reactions to the system. We identified future challenges, including achieving quick engagement without delaying medical care. Salient reasons for poor system adoption were patients' medical state and technological illiteracy. HCPs confirmed the potential of *myED* and identified means that could improve patient experience and staff cooperation.

Conclusions: Our iterative work with ED patients, HCPs, and a multidisciplinary team of developers yielded a system that provides personal information to patients about their ED journey in a secure, effective, and user-friendly way. *MyED* communicates this information through mobile technology. This improves health care by addressing patients' psychological needs for information and understanding, which are often overlooked. We continue to test and refine the system and expect to find positive effects of *myED* on patients' ED experience and hospital operations.

Introduction

Background

Emergency department (ED) care involves multiple assessments, tests, and treatments and engages multiple service providers, stakeholders, and resources [1]. The unpredictability and diversity of the medical state of ED patients poses operational and managerial challenges for sharing information with patients about their hospital ED journey. The lack of such information leads to helplessness and aggression in patients [2,3].

XSL•FO
RenderX

Technology is drastically changing health care delivery [4,5]. It facilitates physician support and patient monitoring, notably through electronic medical records [6,7] and dashboards [8]. Internet-based websites and patient forums increase communication between community clinics, patients, and health care providers (HCPs) [9-13], thus offering patients a wider scope of health and treatment information. Mobile apps are used to support patient self-monitoring, particularly for primary care (eg, medication reminders [14-17]). Novel technologies have begun to provide real-time, patient-centered information [18] for a series of medical care procedures, referred to as the *Patient Journey* [19-21]. The term is derived from the widely accepted concept of *Customer Journey* in Marketing literature, in which it refers to the activities and events included in service delivery from the customer's perspective [22,23]. In this spirit, we are promoting a platform for informing patients about their hospital ED journey, to improve their understanding of the multiple procedures their medical situation requires.

Recently, Vorakulpipat et al [24] developed a system that shows patient status in real time, including waiting times, treatment locations, and treating teams. The system was developed for outpatient clinics and, therefore, also includes billing information and the number of people ahead in line. To reduce patient uncertainty, it presents an updated snapshot of the situation at any given moment. Similarly, Google developed a patent for the automated patient management system [25]; this enables patients to track their own status by viewing patient information on hospital servers at a kiosk or on their own mobile devices. In contrast to these two systems, *myED* reveals completed, current, and anticipated ED procedures, in addition to updated patient status. This is predicted to increase patients' understanding of their personal ED journey.

Objectives

Currently, patients in the ED depend completely on HCPs for information about their medical situation. However, HCPs do not have a systematic protocol for sharing such information with patients and are often working under time constraints. Hence, the communication of information to patients is frequently stalled and, typically, very brief. We searched for a means of communicating information about ED processes to patients and of reducing patient confusion *without* adding to the HCPs' workload or delaying medical procedures.

MyED provides patient-centered information through a responsive website to facilitate access and support of all mobile devices. Accordingly, *myED* breaks down the complexity of ED care for each patient, building on the processing of existing electronic medical records. The system is designed to improve patients' understanding of their personal ED journeys by means of a useful and user-friendly design, which ensures security and privacy [26,27]. Patients access a highly secure platform through a text message they receive to their personal mobile phone on ED admission. The system delivers information regarding individual patients' ED *procedures* (eg, assessments, tests, and treatments) and associated waiting times.

Our design and evaluation of *myED* integrates key elements of the technology acceptance model (TAM; Davis et al [28]) and the information system (IS) success model [29,30]. We used the two key variables *perceived usefulness* and *perceived ease of system use* as guides in our evaluation process because they are known to be valid predictors of attitudes toward system use [28], actual use, and user satisfaction [29,30]. Figure 1 shows screenshots of *myED,* which present an actual patient's journey.

The primary objective of this line of research is to validate effective information communication to patients through their mobile phones and to increase patient satisfaction through an enhanced understanding of their personal ED journey. The goal of this paper was to report on the design, development, and initial evaluation of *myED*.

Figure 1. Screenshots of *myED*, available to an actual patient on September 28, 2018, 15:56.
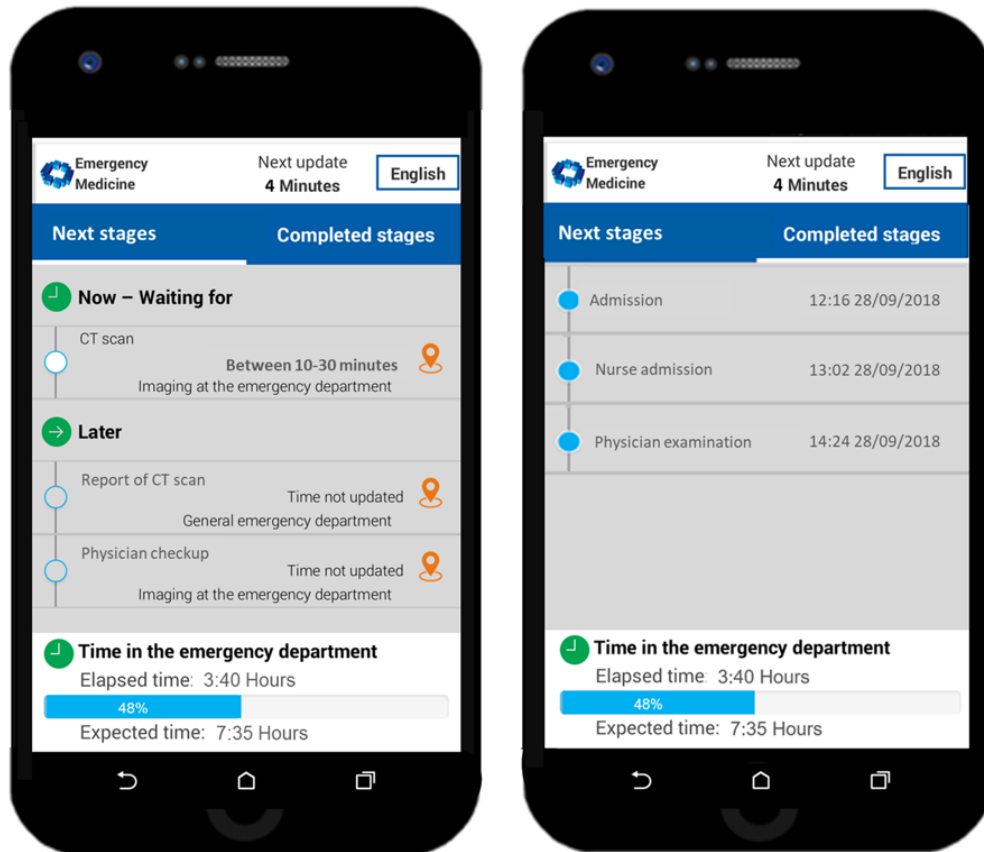

Methods

Overview

The research was conducted in collaboration with an ED of a medium-sized (477-bed) tertiary hospital. The study was reviewed and approved by the Institutional Review Board and conducted in accordance with the Declaration of Helsinki. Phase

I included assessing patient needs, using process mining [31,32] to dynamically create patient information, designing an initial user interface (UI), conducting laboratory evaluations of this design, and redesigning the UI. Phase II comprised deploying and testing the system, identifying barriers to adoption, and refining the design accordingly. The methods and results are reported separately for the two phases, as illustrated in Figure 2.

Figure 2. Overview of the methods used to design, develop, and evaluate *myED*. UI: user interface.
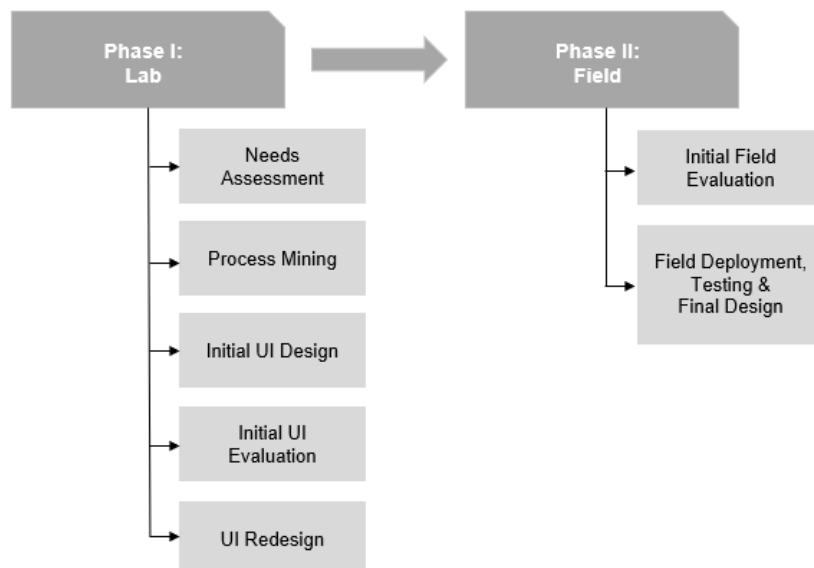
Phase I

Needs Assessment

We conducted semistructured interviews with 2 ED patients, 4 family members, and 5 HCPs to assess patients' needs for information about their ED care. This sample size follows the data saturation criteria for qualitative research (see Sandelowski [33]). The following questions roughly guided interviews with patients and family members: What information did you receive about what is going on with your ED care? Have you tried to find out what is going on? Who did you talk to in the ED? Do you understand why you are waiting and whom you are waiting for? Do you know how much time you will have to wait? What information may be helpful to you at this time? HCPs were asked about *their* perspectives on these patient-related questions.

Process Mining: Mapping Patient Journeys and Predicting Waiting Times

We used *process mining* [31,32] tools (ie, *process discovery* and *queue mining*) to mine patient-related information stored in the medical databases of the hospital ED. We accessed all available information of patients in the ED for 39 months (2014-2017).

First, we mapped all possible patient ED journeys. Using process discovery tools [34], developed in the *Technion Service Enterprise Engineering (SEE)* Lab [35], we decoded the medical procedures in each archived patient's medical record. We then aggregated these data across all patients and visualized this aggregation as a process chart that shows all possible ED journeys (see Figure 3). This information enabled building real-time techniques that detect individual patient journeys and dynamic updating of information during the ED stay.

Figure 3. Process chart of all possible emergency department patient journeys.
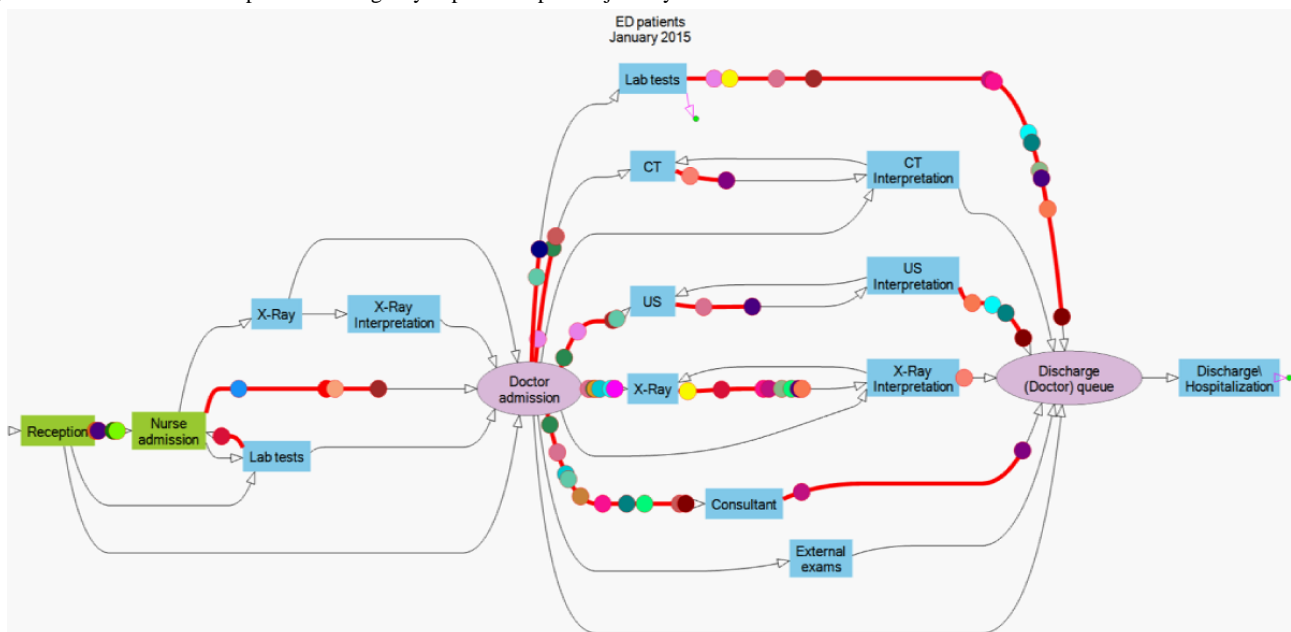
Second, we developed a means of estimating individual patient waiting times for each specific ED procedure. Following Ang et al [36] and Carmeli et al [37], we incorporated queuing theory–based results as features in machine learning methods. We started by estimating the workload in each procedure of ED care. For example, we calculated the number of people queued for a computed tomography (CT) scan when a specific patient entered this queue, and the service rate of the CT scan (ie, the number of patients who undergo a CT scan per hour). We trained a machine learning model (eg, random forest [38]) to predict waiting times for each patient regarding each procedure. The learning model comprised the following types of variables: (1) *time* variables: hour of day, weekday; (2) *patient static* variables: triage level, arrival type (eg, ambulance, walk-in), age, gender; (3) *patient dynamic* variables: completed and anticipated procedures in the patient ED journey; and (4) *dynamic workload* variables: queue length, service rate, the time waited by the last patient to receive treatment, and the total number of patients in the ED.

Initial User Interface Design

Parallel to mining hospital information, we translated some incomprehensible language of medical information into lay terms. This was consequent to the review of a large sample of medical records that identified confusing or unclear medical terms (eg, hemoglobin level and white blood cell count). In consultation with hospital staff, we identified appropriate substitute lay terms. We further verified the clarity of the terms with patients during the first few days of system deployment and did not find any problems regarding the comprehension of the text.

We developed an initial UI for communicating the relevant patient-centered information; our aims were usefulness and ease of use. We included three UI views of ED journeys: *Completed steps* (procedures a patient already completed); *Now* (procedures for which a patient is currently waiting); and *Future steps* (anticipated procedures). The *Now* view shows an estimate of the waiting time for the current procedure, and all views show an estimate of the patient's total length of stay (LoS). Time
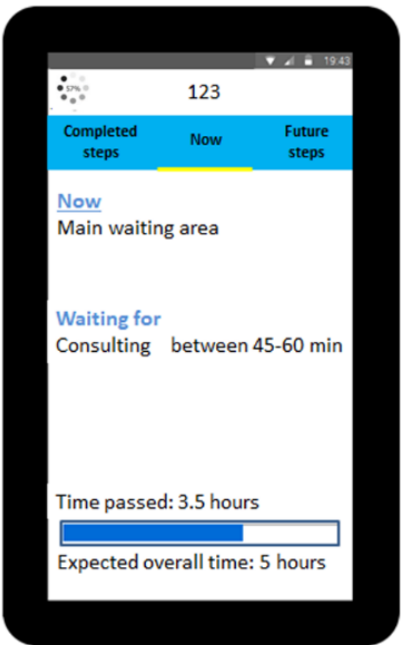
estimates are updated every 5 min, based on changes in ED load and patient prognosis.

Initial User Interface Evaluation and Redesign

We ran a study to evaluate the understanding of the information communicated by our three UI views. We recruited 255 participants on PanelView [39], an online platform for creating surveys and recruiting participants to take the surveys. We asked them to imagine arriving at an ED and being informed about

their ED journey through a novel system that they access through their own mobile phone. The three static UI views were embedded in a three-part storyline. Participants were led through these three screens as depictions of the scripted ED journey and then responded to questions about what they saw. Figure 4 illustrates the *Now* part of the storyline; each participant saw similar views for *Completed steps* and *Future Steps*. On the basis of the initial UI evaluation, we revised the UI design, as described in the Phase I Results section.

Figure 4. Sample storyline, screen, and questions used to test the initial user interface design.

Phase II

Initial Field Evaluation

We deployed *myED* in the pediatric section of the hospital's general ED. This enabled a pilot in a smaller, more controlled environment. The 6-month pilot identified issues and potential obstacles arising from the mining process.

Field Deployment, Testing, and the Final Design

After initial UI evaluation and redesign, we deployed *myED* in the ambulant adult section of the same ED. The ED includes all disciplines except maternity and otolaryngology. The adult section is divided into three subsections: ambulant (triage score 3-5), lying-in (triage score 2-3), and trauma (triage score 1).

As part of the routine ED admission process, patients were asked to provide their phone numbers. They were then informed about *myED* and sent a text message with a link to the log-in screen. After agreeing to the terms, including their consent to be part of a study on improving patient ED experience, they could enter the website anytime with their details. We designed *myED* to impart a high level of security and protection of privacy, thus mirroring the hospital's ED medical records and to extract only information about the patients' medical procedures. The

system's architecture is based on a demilitarized zone (DMZ) server that is separate from the hospital databases; *myED* has access to the DMZ server only, adding an additional layer of security. Nonsensitive patient information is extracted and displayed on the patient's mobile phone; no confidential information is displayed on the screen, and patients are not identified in the *myED* records.

As part of the system evaluation, we assessed the reactions of patients and HCPs and made final small changes to the system design.

First, 5 students who served as research assistants (RAs) shadowed and interviewed a sample of 482 patients for 2 hours a day during the first four weeks after deployment (July-August 2018) to understand the perceived usefulness and attitude toward *myED* use. RAs approached ED patients one-by-one (excluding those who seemed to be in great pain) and asked if they had entered *myED* and if they were willing to share their feedback. If a patient had not yet entered *myED*, the RA asked if he or she was interested in doing so. If relevant, the RAs helped with the entry process; if not, the RA asked if the patient was willing to share why he or she did not want to use *myED*. Specifically, we measured four types of attitudes toward *myED* use: (1) self-initiated entry to *myED*, (2) *myED* entry initiation once

approached, (3) inability to enter *myED*, and (4) disinterest in entering *myED*. We also noted perceived usefulness from *myED* users (eg, "I really like this system. Finally, someone cares about the patients!") and nonusers ("I have my file in my hand, I don't need your system!").

Second, we analyzed *myED* records for actual system use during 3 months of system deployment (August-October 2018). We excluded patients who did not receive a text message (888/4767, 18.63%: eg, their phones were not with them or were without battery power or internet connection). This decreased the baseline relevant population size to 3879. We then computed the number of people who used *myED* during August-October 2018 (1131/3879, 29.16%) and identified the point of their ED journey when they first logged in. We also reported the system adoption rates that we reached in the third month of deployment (October 2018): 1614 people received a text message (81.39% of 1983 who arrived), of whom 492 (30.48%) used *myED*.

Third, 5 RAs (students) surveyed a sample of 349 people about their understanding of the personal ED journey, both before and after system deployment. RAs approached ED patients one-by-one (excluding those who seemed to be in great pain) and invited them to participate in a survey regarding the ED service. In June 2018, 60.2% (210/349) of people responded (system nonusers, the *control group*), and in August-October 2018, 39.8% (139/349) responded (system users, the *intervention group*). Short surveys assessed patient understanding ("I understand the sequence of procedures of my treatment"; "I understand the various stages of my treatment," on a scale from 1 [not at all] to 7 [very much]), the time they had already spent in the ED until they filled out the survey, and patient demographic characteristics (age; gender; economic status, defined as the number of people divided by the number of rooms at home; religion: Jewish, Muslim, Christian, Druze, other).

Fourth, we interviewed 5 HCPs regarding their own attitudes and their perceptions of patients' attitudes of the usefulness and ease of use of *myED*. We followed a semistructured interview protocol (eg, Have you seen *myED*? Do your patients use it? How do you feel about *myED*? How do you think patients feel about *myED*? Does *myED* influence your work? Does anything bother you about *myED*? Any ideas on what could improve *myED*?).

Results

Phase I

Needs Assessment

Interviews strongly underlined patients' need for information about their medical procedures (what, when, where), and about waiting times (when, how long). All 11 respondents mentioned all these issues. Hence, *myED* was developed to address the following needs: information about (1) procedures in the ED journey, (2) estimated waiting times, (3) the location for each procedure (because some procedures occur outside the ED, eg, in outpatient clinics), and (4) total ED LoS.

Process Mining: Mapping Patient Journeys and Predicting Waiting Times

MyED generates individual, constantly updated information, fed to the mobile phone of patients (see Figure 1). The system translates information stored in hospital ED medical records to patient-friendly information and updates itself every 5 min, using the following analyses:

Process Discovery: Identifying Procedures in a Patient Journey

We aggregated detailed medical examinations that are conducted at the same time and location into operational procedures. For example, hemoglobin level and white blood cell count (which are recorded separately in the ED medical records) were aggregated into *lab tests* because patients experience them together and view them as a single procedure. Aggregating information across all patients produced all possible ED journeys, as depicted in Figure 3 (for a dynamic view, see [40]).

Each patient's journey is depicted as a distinct path in this graph, comprising a certain order of medical procedures. There is almost no predetermined order, and patients can simultaneously wait for two or more ED procedures, with no clear indication of which should be performed first. Patient journeys also vary in complexity. For example, Figure 5 shows a more and a less complex journey (Patient A and Patient B, respectively). The full complexity of a patient's journey evolves continuously during the patient's ED stay. Therefore, *myED* constantly adjusts information communicated to patients, based on updates to ED medical records.

We also identified procedures that typically occur sequentially; for example, ultrasound (US) examinations are always followed by US interpretation. We use such identified sequences to predict elements in the ED journey before they appear in the ED medical records. This is the foundation for building the information *myED* presents to patients as *anticipated procedures*.

Queue Mining: Estimating Waiting Times

ED medical records include completion times for ED procedures and are the source of the waiting times presented in the *Completed steps* view of *myED*. Queue mining methods enable *myED* to present patients with waiting times for anticipated procedures. Following Carmeli et al [37], we identified a probabilistic range of waiting times for each procedure. We used the 0.15-quantile as the lower and the 0.85-quantile as the upper bound of the reported range, which ensured that waiting times of no more than 15% of the patients exceeded our prediction.

Figure 5. Two patients' emergency department journeys identified using process mining. CT: computed tomography, US: ultrasound.
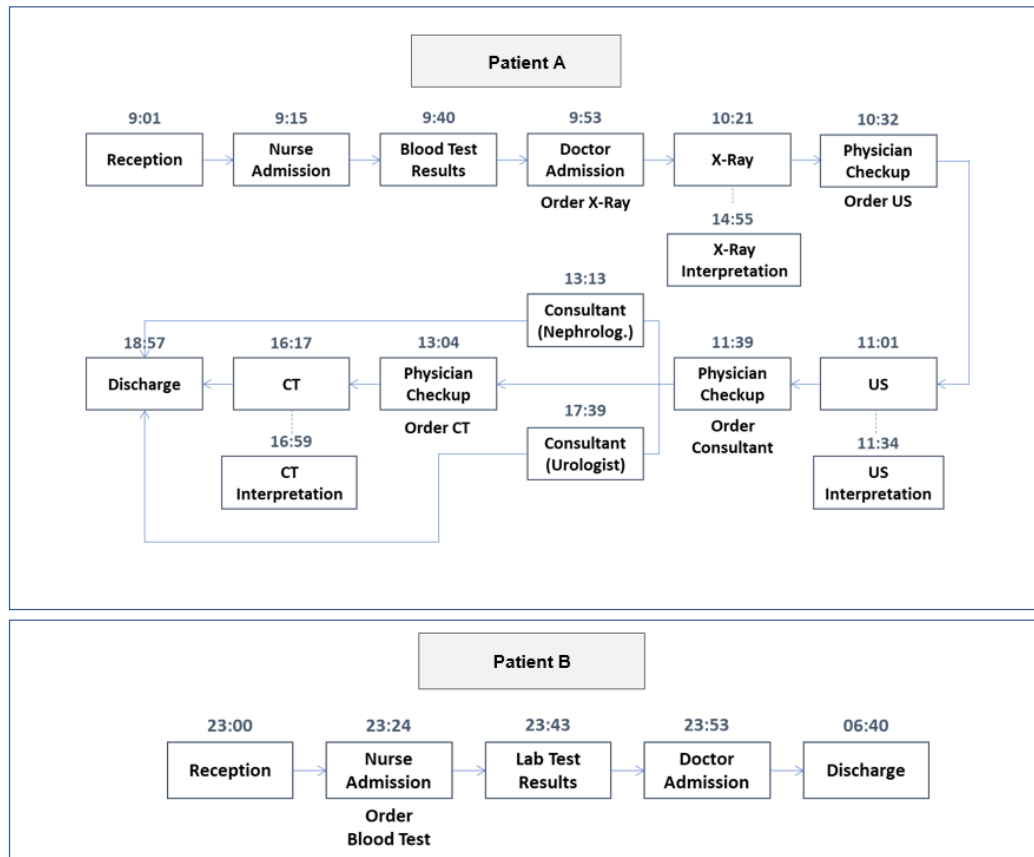
Initial User Interface Evaluation and Redesign

In total, 255 participants (age: mean 53 years, SD 8.6 years; 140/255, 54.9% female) evaluated the three UI-view design in the online study. The results showed that a high proportion of participants (202/255, 79.2%) indicated understanding the presented information. However, responses to specific questions about the anticipated procedures showed that 29.8% (76/255) were confused by the separation between *Now* and *Future Steps* views. Hence, we redesigned the UI into two views: *Next stages* and *Completed stages*, as depicted in Figure 1. *Next stages* includes both upcoming procedures (*Now – Waiting for*) and subsequent procedures (*Later*). Hence, the *Now* and *Future steps* were collapsed into one view.

We found that people prefer that waiting times are presented as a range (*between X and Y min*). Most participants (151/255, 59.2%) reported that this presentation seems more reliable and trustworthy than *at least X min* (60/255, 23.5%) and *about X min* (44/255, 17.3%). Figure 1 illustrates this. The patient depicted needs to wait 10-30 min for a CT scan. The same view shows that later he or she will wait for a report of the CT scan and meet a physician (Figure 1, left side). The *Completed stages* view (Figure 1, right side) shows that the patient was admitted (at 12:16 PM) and already saw a nurse (at 13:02 PM) and a physician (at 14:24 PM).

Phase II

Initial Field Evaluation

The 6-month pilot in the pediatric ED identified several issues and potential obstacles arising from the mining process. First, hospital databases are not always updated by the medical staff in real time. This creates gaps between the actual patient journey and the information available in *myED*. For example, we found that for 74.74% (1098/1469) of the patients, laboratory tests were reported in the databases only *after* completion. Hence, many patients could not see any information in *myED* regarding these laboratory tests while they were still waiting for them. Similarly, some procedures cannot be accessed by *myED* because they are recorded in inaccessible databases. For example, *myED* cannot access outpatient clinic databases. According to the data we gathered during April 2014-March 2015, this occurs for 0.26% (50/19,279) of the patients. Missing updates of delays also create a statistical challenge as they create missing data in the model that predicts waiting times and LoS. The field evaluation revealed inaccuracies in predictions, which were handled by retraining our learning models. For example, our algorithm underestimated 47% (46/98) of the waiting times for US examinations before the retraining (July 2018), and only 22% (17/76) afterward (August 2018).

Second, hospital guidelines change over time, creating changes in patient journeys, and, if not updated, inaccuracies in *myED*. For example, in the historical data used to create the system, all x-ray tests were analyzed by a radiology specialist. However, during *myED* deployment, new guidelines allowed regular physicians to analyze some simple x-ray tests. This created

mismatched information because *myED* informed patients they were waiting for an x-ray interpretation, instead of a physician. We quickly adapted *myED* to such guideline changes. This example emphasizes the need to continually update the system to reflect policy and protocol changes and to evaluate system accuracy periodically.

Third, the unpredictability of some ED procedures means that the actual waiting time for a specific ED procedure may exceed the waiting time predicted by *myED*. For example, a meeting with a specialist may be delayed substantially because of an emergency in his or her home unit. This creates a dilemma in deciding what information to show patients when the actual waiting time exceeds the upper bond (85th-quantile) reported. We considered three options: (1) show a generic estimate such as *up to 15 min* until a procedure is completed; (2) stop showing information, noting *time estimates are not available*; or (3) tell patients to check with HCPs regarding the delay. Consultation with hospital staff identified the first option as the best.

Field Deployment, Testing, and the Final Design

During the first month of deployment, 482 patients were actively approached (age: mean 50 years, SD 18.9 years; range 20-90 years; 265/482, 55.0% female) to use the system. Of them, 19.9% (96/482) were already using it, and 40.1% (193/482) then agreed to use it.

Of 349 respondents, 49 (11.7%) provided incomplete demographic information and were, therefore, excluded from the analysis that assessed patients' understanding of their personal ED journey. The intervention group (*myED users*) comprised 139 respondents, and the control group 210 respondents. Patient age was similar between the groups (age: mean 46 years, SD 16.1 years, range 18-83 years; age: mean 46 years, SD 17.9 years, range 18-93 years, respectively). Sex distribution was also similar (51.9% and 51.8%, respectively). The patients in the intervention group had a significantly better understanding of their ED journeys than did the patients in the control group after controlling for their age, gender, economic status, and religion ($F_{8,299}=2.519$, $P=.01$).

Of the patients actively approached, 39.2% (189/482) provided open-ended responses about their attitudes toward system use. All following responses of patients and medical staff were coded as [X.Y, gender, age], with 'X' representing the interviewee number and 'Y' the date in 2018 in ddmm format. Positive responses (118/189, 62.4%) included short praises (eg, "good!" "nice!" and "great!"), and general delight:

> *I used it all day long! It's good.* [311.0908, female, 28 years]

> *This is something new, good idea.* [112.2607, male, 47 years]

> *A very effective system, thank you.* [326.1908, male, 40 years]

> *The system is excellent.* [297.1608, female, 35 years]

All 5 HCPs interviewed also perceived *myED* as both useful and easy to use and agreed with its potential benefits for both patients and medical staff:

> *The staff should accept [the system] as an inseparable part of handling patients here. There were some patients who came to us and didn't know how to activate the system, but the procedure is very simple; it is quite friendly and easy to use.*

> *It is very effective, it's a shame that it wasn't there before. It really helps us, people really check and come to me and say they see their tests.* [1.0708, male, 40 years]

> *I hope it will help both patients and us, as it will enable them to understand their treatments.'*

> *It will definitely do no harm, it can only be useful.* [5.0708, female, 48 years]

Barriers to System Adoption and Corresponding Modifications

In the first week of deployment in the ambulant adult ED, *myED* log-in rates varied greatly (ie, between 7/54, 13% on day 3 and 27/64, 42% on day 1). We noticed several causes of poor system adoption.

Initially, admission staff were required to offer *myED* to patients and to ask for consent before manually registering them. One of 5 HCPs expressed concerns about delaying medical procedures:

> *I worry that the system will add time until patient triage, and that we will lose critical time-to-triage, which must be done within 15 minutes.* [5.0708, female, 48 years]

We saw that HCPs were not offering *myED*, especially when the ED was loaded, to not delay time-to-triage. Another cause of poor adoption was the identity format required to log into *myED*, which was designed according to the format used in ED medical records. The design was different from the routinely used identity number. This confused patients and stalled the use of *myED*. We handled both issues with a *login page redesign*. Specifically, we introduced automatic registration, thus integrating patient consent into the log-in process and modified the format of patient identification. This yielded log-in rates of 26.26% (266/1013) within 1 month of system deployment.

Second, not all patients knew about *myED*. We introduced *local advertising of myED* within the ED; flyers about the system increased log-in rates by 13.44% to 29.79% (373/1252) in the second month of system deployment.

Third, people arriving at the ED were often preoccupied and stressed and frequently missed the text message that was automatically sent for logging into the system:

> *I have such a migraine I can't look at anything.* [128.2607, female, 55 years]

> *I don't have the patience for this now.* [129.0808, female, 38 years]

> *I can't listen to what you say, I don't feel well.* [90.0608, male, 35 years]

> *I'll look at it at home after all this is over. I don't feel well right now.* [31.2407, male, 70 years]

Fourth, *myED* reports of 266 users (August 2019) showed the mean time-to-first-entry as high as 60 (SD 74) min after arrival. We, therefore, introduced a new *reminder text message*, sent 30 min after arrival to anyone who had not yet logged in; 30 min allow most patients to complete the initial triage, nurse admission, and first physician examination. This shortened time-to-first-entry in the following 2 months by 18%, to 49 (SD 62) min (865 users, September-October 2019).

Figure 6 summarizes the effects of these design modifications, which further increased log-in rates by 2.32% to 30.48%

(492/1614) in the third month—an impressive adoption rate in such a short time [41].

Further barriers to *myED* log-in attempts included issues with people's phones, which precluded their receiving the text message (888/4767, 18.63%). As depicted in Figure 7, 81.37% (3879/4767) of patients in the ED received the text message, of whom 37.69% (1462/3879) attempted to log in. Log-in failures (331/1462, 22.64%) included technical issues such as disabled cookies.

Figure 6. Influence of design modifications on myED login rates.
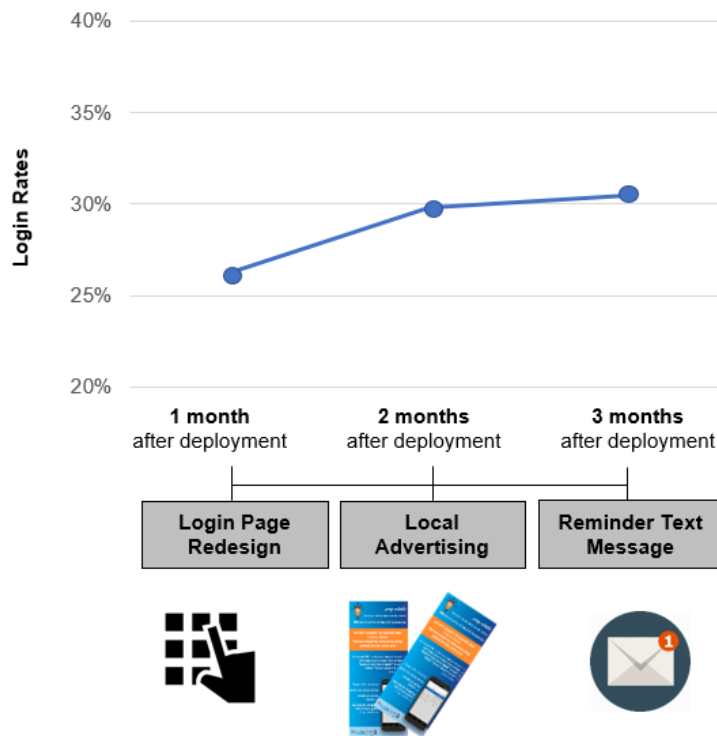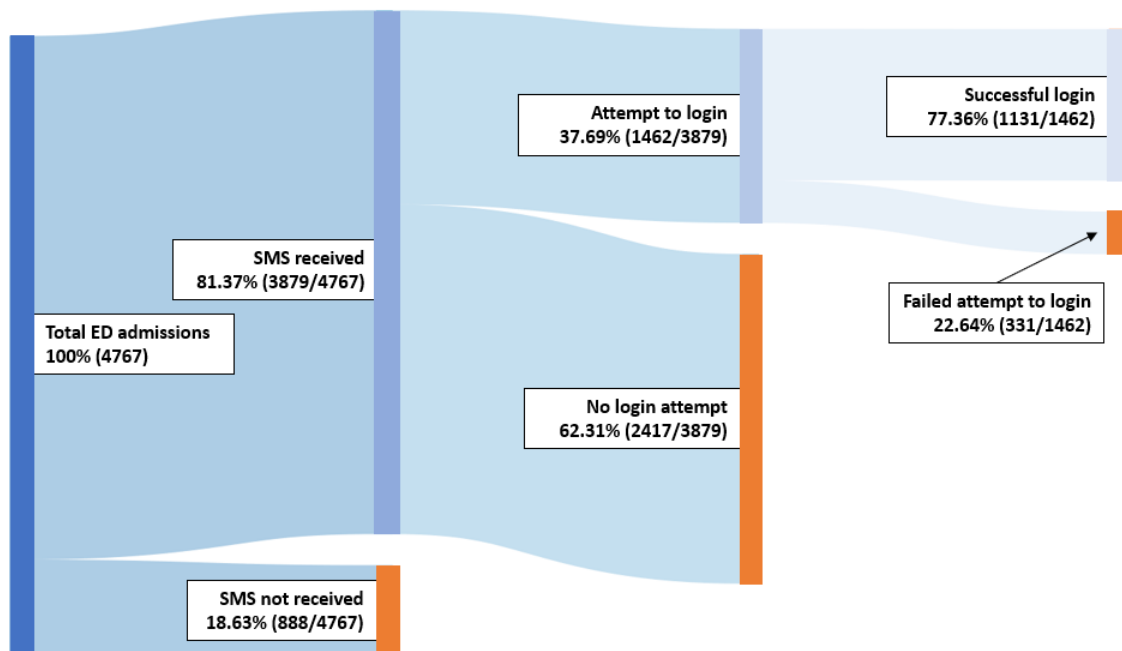

Figure 7. System adoption in first 3 months of deployment (08-10/2018). ED: emergency department.

Of 189 patients who provided open-ended responses about their attitudes toward system use, 37.6% (71/189) relayed negative or mixed (both negative and positive) attitudes. In addition, 14% (10/71) alluded to technology literacy or dependence on others as an issue:

> *I am still in the Stone Age with regard to anything digital.* [94.2407, female, 80 years]

> *I talk to, and text only with my children.* [71.0107, male, 70 years]

> *I trust my wife will take care of everything.* [12.1807, male, 40 years]

Most challenging to system acceptance were doubts about the value of *myED* (49/71, 69%):

> *I have my file in my hand, I don't need your system.* [20.1807, female, 20 years]

> *Why do I need this if the doctor will tell me the results of the blood test?!* [223.0708, female, 65 years]

These reactions seem to be related to the long and frustrating time spent in the ED:

> *What for? I know what to expect...I know what I am waiting for and for how long - basically, all night.* [48.2607, male, 60 years]

> *It serves no purpose if I have to wait for 4 hours.* [96.2407, female, 79 years]

> *This is worthless! I am waiting for the cardiologist and I don't care that this is what it says on my mobile's screen! What I care about is that I've been waiting here already for 5 hours!* [316.1208, female, 60 years]

However, longer LoS did not hinder system adoption (August-October 2018). On the contrary, *myED* users had longer mean LoS than nonusers (4.5 and 3.9 hours, respectively).

Ten percent (7/71) of the comments we received were related to the accuracy of *myED* information:

> *Worthless! I am half an hour past my neurological consultation and it says here I am waiting for it!* [13.1807, male, 20 years]

> *I'm not waiting for the procedure it says!* [268.1208, male, 50 years]

Two of 5 HCPs initially expressed concerns about *myED* and their workload:

> *I don't feel it will reduce the workload. They [patients] still come to ask what is going on, and what they are waiting for.*

> *Another thing that we feared could happen is that there will be more nagging once they receive information that results arrived.* [5.0708, female, 48 years]

> *Once test results arrive and the patient knows it, it will only put more pressure. They [patients] can start knocking on doors, and this will put pressure on [other] patients and us.* [3.0708, male, 39 years]

These concerns suggest that merely informing patients about procedures could potentially increase HCPs' workload, if not done appropriately. However, when asked whether these concerns materialized, HCPs responded that they did not. Rather, patients do not seem to ask more, but simply ask different questions:

> *I feel that patients' questions have changed. Now they already know that the test results have arrived.* [2.0708, female, 51 years]

Suggested Improvements for the System

Eleven percent (21/189) of the patients who responded to open-ended questions suggested ideas for improvement. For example, they asked why the system does not provide more detailed personal or medical information:

> *After logging in there should be some identification that this is indeed my account, like my name or identity number, on the screen.* [211.0608, male, 80 years]

> *I would have liked to see results of tests, like I can see in the HMO website.* [219.0608, male, 50 years]

However, owing to security concerns, identifying patients and providing medical information would require an even higher level of security in the system, which would substantially complicate registration. We were concerned that this would hamper people's willingness to adopt a system that they can use only for a few hours. Showing test results can also be risky, as 2 of 5 HCPs noted, for example:

> *I was afraid that the system would send the actual results of the tests...I know from other HMOs [health maintenance organizations], that if people suddenly see that the result of a blood test appears in red, they are stressed. It can be very stressful [for us] in the ED.* [3.0708, male, 39 years]

Abnormal test results can add to patient stress, whereas normal test results may cause patients to leave without being seen—both are undesirable consequences.

Patients also asked for the names of the HCPs they were waiting for:

> *The system does not give me a lot of information, I suggest to add the doctor's name in the relevant step.* [109.2607, male, 50 years]

Providing this information in *myED* is impossible because it does not appear in ED medical records.

Discussion

Principal Findings and Comparisons With Previous Work

This paper describes the development and initial implementation of *myED*, a system that addresses the need of ED patients for information about their medical journey. *MyED* is a personalized, frequently updated information system, accessible by patients anywhere and anytime during the ED visit, on their personal mobile phones. Vorakulpipat et al [24], who developed

a system comparable with *myED*, focused on outpatient clinics. Although their system provides additional information, such as the number of people ahead in line and the name of the treatment team, only a patient's current status is revealed. In contrast, *myED* reveals the entire patient ED journey. The underlying assumption is that comprehensible, continuously updated information about personal ED journeys will improve patients' understanding of the process and reduce frustration and anger [42]. Our results attest to increased patient understanding and overall positive responses to the system. The adoption rate of *myED* at the end of the first 3 months was satisfactory (492/1614, 30.48%), thus confirming the viability of the system.

Our design and evaluation of *myED* integrates key elements of the TAM [28] and IS success model [29,30]. Specifically, we report positive perceptions by patients and HCPs of ease of use and usefulness, positive and negative attitudes toward use, and a reasonable proportion of actual use of *myED*. The mixed methods approach enabled presenting multifaceted data of *myED* users and nonusers, as well as of HCPs regarding these cognitive and behavioral aspects of *myED* use.

Meeting Challenges

Our study offers ways to tackle three distinctive challenges that we encountered: (1) extracting real-time ED medical records and transforming them into comprehensible and accurate information, (2) providing information to patients without disrupting the ED workflow [43,44], and (3) getting patients to use their personal mobile phones to obtain information during ED visits, which are short-term and have a limited user-engagement period (a few hours, during the day of their visit).

To address the first challenge, we modeled ED patient journeys and presented them to patients, using an innovative, unique combination of operations research tools (ie, process discovery and queue mining [31,32]) and user-centered design methods. The interdisciplinary effort enabled translating medical and process-related information in ED medical records into real-time information regarding personal procedures. Specifically, *myED* translates existing but fragmented information into clearly structured information regarding completed, current, and anticipated medical procedures, including estimated waiting times.

Meeting the second challenge, *my*ED works with minimal disruption to the ED workflow. First, the system relies on available information, extracted directly from ED medical records that HCPs routinely update. Second, patients do *not* see actual test results or other concrete medical information. Such information requires professional interpretation and can cause patient anxiety if not communicated appropriately. Third, the *myED* design enables automatic log-in and thus ensures that the registration process does not increase time-to-triage. Fourth, the *myED* design affords patients independence by requiring no HCP involvement. The intuitive *myED* design is easy to use and employs lay terms (not medical jargon), reducing patient confusion. Finally, *myED* provides location information, which has thus far been provided only by HCPs, if at all. This improves patient orientation and can thereby reduce delays.

Regarding the third challenge, *myED* was well accepted and adopted by patients despite the short duration of user-system engagement. Mobile technology increasingly provides patients with health care information; we showed it can also be useful in the ED. As *myED* users are mostly one-time visitors, the increase in system adoption reflects the success of *myED* design modifications.

Limitations

A number of limitations are relevant to the conduct of this study. First, the number of persons interviewed for the needs assessment was small (11 in total). In addition, the cohort of users of *myED* may reflect a selection bias. One reason is that RAs who encouraged the use of the system were instructed not to approach patients in great pain. Moreover, we only deployed and tested *myED* in the pediatric and ambulant adult ED of one specific hospital. Patients in the lying-in unit, or in the same units in smaller or larger hospitals, may have needs that were not identified in this research.

Future Work

The accuracy of *myED* depends on regular updates of medical information by HCPs. Otherwise, discrepancies may occur between the information presented in *myED* and the information patients obtain from other sources. Such discrepancies can reduce patient trust and—rather than offering relief—exacerbate patient confusion. Ways to avoid such discrepancies must be sought, without disrupting the ED workflow and HCP workload. For example, updating information during patient meetings on portable computers could improve real-time information, without adding ED workload. Second, the quick and short-term user-system engagement of *myED* is a continuous challenge. An effective system would avail rapid acceptance and adoption beyond what we accomplished. Low system adoption among patients who feel sick or are in pain, and among those who are technologically illiterate, remains a challenge. In the effort to improve system adoption, the *onboarding process*, (ie, receiving and opening the link in the text message, logging in, relogging) should be further simplified. Third, a new challenge we identified during *myED* implementation is the need to secure the stability of the fragile relationship between ED staff and patients. HCPs recognized the benefits of *myED* but expressed concerns that it could change the power relations between staff and patients if patients see information *before* the staff [45]. To avoid such concerns, we must strive for mutual HCP-patient empowerment (see Parush [46]).

Research has shown that providing information about waiting time affects people's behavior [47]. Our design includes waiting times in range format (eg, *between 30 and 45 min*) to decrease the risk that patients would be unavailable once they reach the head of the line. Future research is needed to verify this prediction. Inaccuracy of waiting time estimates has been shown to decrease trust and increase frustration [48-50]. Hence, future research should continue to explore the optimal means for presenting waiting times to ED patients. Finally, another direction for future research is the investigation of patient-related measures beyond patient understanding. These include patient satisfaction of the ED visit and experienced frustration, as well

as other important operational measures, such as a patient leaving without completing the planned procedures [51].

In summary, our research shows that *myED* is a novel and revolutionary approach for improving a patient's understanding of his or her personal ED journey. This new use of ED medical records can improve patients' experience of ED visits.

Conflicts of Interest

None declared.

References

1. Armony M, Israelit S, Mandelbaum A, Marmor YN, Tseytlin Y, Yom-Tov GB. On patient flow in hospitals: a data-based queueing-science perspective. Stoch Syst 2015;5(1):146-194. [doi: 10.1214/14-ssy153]

2. Anderson P. Perspective: complexity theory and organization science. Organ Sci 1999;10(3):216-232. [doi: 10.1287/orsc.10.3.216]

3. Efrat-Treister D, Moriah H, Rafaeli A. The effect of waiting on aggressive tendencies toward emergency department staff: Providing information can help but may also backfire. Plos One 2020;15(1):e0227729 [FREE Full text] [doi: 10.1371/journal.pone.0227729]

4. Srinivasa K, Sowmya B, Shikhar A, Utkarsha R, Singh A. Data analytics assisted internet of things towards building intelligent healthcare monitoring systems. J Organ End User Comput 2018;30(4):83-103. [doi: 10.4018/joeuc.2018100106]

5. Britnell M, Bakalar R, Sehata A. KPMG International. Digital Health: Heaven or Hell? How Technology Can Drive or Derail the Quest for Efficient, High Quality Healthcare URL: https://assets.kpmg.com/content/dam/kpmg/pdf/2016/03/digital-health-heaven-hell.pdf [accessed 2020-01-16]

6. Darr A, Harrison MI, Shakked L, Shalom N. Physicians' and nurses' reactions to electronic medical records. Managerial and occupational implications. J Health Organ Manag 2003;17(5):349-359. [doi: 10.1108/14777260310505129] [Medline: 14628488]

7. Wiljer D, Urowitz S, Apatu E, DeLenardo C, Eysenbach G, Harth T, Canadian Committee for Patient Accessible Health Records. Patient accessible electronic health records: exploring recommendations for successful implementation strategies. J Med Internet Res 2008 Oct 31;10(4):e34 [FREE Full text] [doi: 10.2196/jmir.1061] [Medline: 18974036]

8. Yoo J, Jung KY, Kim T, Lee T, Hwang SY, Yoon H, et al. A real-time autonomous dashboard for the emergency department: 5-year case study. JMIR Mhealth Uhealth 2018 Nov 22;6(11):e10666 [FREE Full text] [doi: 10.2196/10666] [Medline: 30467100]

9. Olthuis G, Prins C, Smits MJ, van de Pas H, Bierens J, Baart A. Matters of concern: a qualitative study of emergency care from the perspective of patients. Ann Emerg Med 2014 Mar;63(3):311-9.e2. [doi: 10.1016/j.annemergmed.2013.08.018] [Medline: 24054787]

10. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Participation in online patient support groups endorses patients' empowerment. Patient Educ Couns 2009 Jan;74(1):61-69. [doi: 10.1016/j.pec.2008.07.044] [Medline: 18778909]

11. Merolli M, Gray K, Martin-Sanchez F. Health outcomes and related effects of using social media in chronic disease management: a literature review and analysis of affordances. J Biomed Inform 2013 Dec;46(6):957-969. [doi: 10.1016/j.jbi.2013.04.010] [Medline: 23702104]

12. Vreugdenhil MM, Ranke S, de Man Y, Haan MM, Kool RB. Patient and health care provider experiences with a recently introduced patient portal in an academic hospital in the Netherlands: mixed methods study. J Med Internet Res 2019 Aug 20;21(8):13743. [doi: 10.2196/13743] [Medline: 31432782]

13. van den Bulck SA, Hermens R, Slegers K, Vandenberghe B, Goderis G, Vankrunkelsven P. Designing a patient portal for patient-centered care: cross-sectional survey. J Med Internet Res 2018 Oct 1;20(10):e269 [FREE Full text] [doi: 10.2196/jmir.9497] [Medline: 30287416]

14. Lau AY, Piper K, Bokor D, Martin P, Lau VS, Coiera E. Challenges during implementation of a patient-facing mobile app for surgical rehabilitation: feasibility study. JMIR Hum Factors 2017 Dec 7;4(4):e31 [FREE Full text] [doi: 10.2196/humanfactors.8096] [Medline: 29217504]

15. Quanbeck A, Gustafson DH, Marsch LA, Chih M, Kornfield R, McTavish F, et al. Implementing a mobile health system to integrate the treatment of addiction into primary care: a hybrid implementation-effectiveness study. J Med Internet Res 2018 Jan 30;20(1):e37 [FREE Full text] [doi: 10.2196/jmir.8928] [Medline: 29382624]

16. Bernhard G, Mahler C, Seidling HM, Stützle M, Ose D, Baudendistel I, et al. Developing a shared patient-centered, web-based medication platform for Type 2 Diabetes patients and their health care providers: qualitative study on user requirements. J Med Internet Res 2018 Mar 27;20(3):e105 [FREE Full text] [doi: 10.2196/jmir.8666] [Medline: 29588269]

17. Rohde A, Duensing A, Dawczynski C, Godemann J, Lorkowski S, Brombach C. An app to improve eating habits of adolescents and young adults (Challenge to Go): systematic development of a theory-based and target group-adapted mobile app intervention. JMIR Mhealth Uhealth 2019 Aug 12;7(8):e11575 [FREE Full text] [doi: 10.2196/11575] [Medline: 30903746]

18. Kim S, Mikesell L, Fadem S, Aakhus M. Designing a Personalized Support Tool for Patients facing Bone Marrow Transplant. In: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2019 Presented at: PervasiveHealth'19; May 20 - 23, 2019; Trento, Italy p. 41-50. [doi: 10.1145/3329189.3329206]

19. Layton A, Moss F, Morgan G. Mapping out the patient's journey: experiences of developing pathways of care. Qual Health Care 1998 Dec(7 Suppl):S30-S36. [Medline: 10339033]

20. Trebble TM, Hansi N, Hydes T, Smith MA, Baker M. Process mapping the patient journey: an introduction. Br Med J 2010 Aug 13;341:c4078. [doi: 10.1136/bmj.c4078] [Medline: 20709715]

21. Westbrook JI, Coiera EW, Sophie Gosling A, Braithwaite J. Critical incidents and journey mapping as techniques to evaluate the impact of online evidence retrieval systems on health care delivery and patient outcomes. Int J Med Inform 2007;76(2-3):234-245. [doi: 10.1016/j.ijmedinf.2006.03.006] [Medline: 16798071]

22. Zomerdijk LG, Voss CA. Service design for experience-centric services. J Serv Res 2010;13(1):67-82. [doi: 10.1177/1094670509351960]

23. Lemon KN, Verhoef PC. Understanding customer experience throughout the customer journey. J Mark 2016;80(6):69-96. [doi: 10.1509/jm.15.0420]

24. Vorakulpipat C, Rattanalerdnusorn E, Sirapaisan S, Savangsuk V, Kasisopha N. A mobile-based patient-centric passive system for guiding patients through the hospital workflow: design and development. JMIR Mhealth Uhealth 2019 Jul 22;7(7):e14779 [FREE Full text] [doi: 10.2196/14779] [Medline: 31333195]

25. Blom MG. Google Patents. Washington, DC: US Patent and Trademark Office; 2011. Automated Patient Management System URL: https://patents.google.com/patent/US8000979B2/en [accessed 2020-01-16]

26. Dillon A, Morris MG. User acceptance of new information technology: theories and models. Ann Rev Inf Sci Technol 1996;14(4):3-32 [FREE Full text]

27. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. Manag Inf Syst Q 2003;27(3):425-478. [doi: 10.2307/30036540]

28. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. Manag Inf Syst Q 1989;13(3):319-339. [doi: 10.2307/249008]

29. DeLone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. J Manag Inf Syst 2003;19(4):9-30. [doi: 10.1080/07421222.2003.11045748]

30. DeLone WH, McLean ER. Information systems success: the quest for the dependent variable. Inf Syst Res 1992;3(1):60-95. [doi: 10.1287/isre.3.1.60]

31. van der Aalst WM. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Heidelberg: Springer-Verlag Berlin Heidelberg; 2011.

32. Senderovich A, Weidlich M, Gal A, Mandelbaum A. Queue mining for delay prediction in multi-class service processes. Inf Syst 2015;53:278-295. [doi: 10.1016/j.is.2015.03.010]

33. Sandelowski M. Sample size in qualitative research. Res Nurs Health 1995 Apr;18(2):179-183. [doi: 10.1002/nur.4770180211] [Medline: 7899572]

34. Mandelbaum A, Trofimov V, Nadjharov E, Yuviler N, Senderovich A, Gavako I. Technion | Israel Institute of Technology. Haifa; 2014. SEEGraph Designer and SEEGraph Viewer URL: https://web.iem.technion.ac.il/images/see-lab/5/see_graph.pdf [accessed 2020-01-16]

35. Technion | Israel Institute of Technology. Haifa; 2019. Service Enterprise Engineering (SEE) Lab URL: https://web.iem.technion.ac.il/he/service-enterprise-engineering-see-lab/general-information.html [accessed 2020-01-16]

36. Ang E, Kwasnick S, Bayati M, Plambeck EL, Aratow M. Accurate emergency department wait time prediction. Manuf Serv Oper Manag 2016;18(1):141-156. [doi: 10.1287/msom.2015.0560]

37. Carmeli N, Yom-Tov G, Boxma O. Eurandom Preprint Series. 2018 Jan 30. State-Dependent Estimation of Delay Distributions in Fork-Join Networks URL: https://link.springer.com/content/pdf/10.1023/A:1010933404324.pdf [accessed 2020-01-16]

38. Breiman L. Random forests. Mach Learn 2001;45(1):5-32 [FREE Full text] [doi: 10.1023/A:1010933404324]

39.   PanelView. 2019. Who we are URL: https://www.panelview.co.il/About.aspx [accessed 2020-01-16]

40.   SEE Lab - Technion. YouTube. Haifa: SEE Lab Technion; 2017 Nov 2. ED Patient Flow URL: https://www.youtube.com/watch?v=HP_au996Ffw&feature=youtu.be [accessed 2020-01-16]

41.   Fraccaro P, Vigo M, Balatsoukas P, Buchan IE, Peek N, van der Veer SN. Patient portal adoption rates: a systematic literature review and meta-analysis. Stud Health Technol Inform 2017;245:79-83. [Medline: 29295056]

42.   Sutton R, Khan R. Prediction, understanding and control as antidotes to organizational stress. In: Lorsch JW, editor. Handbook of Organizational Behavior. Englewood Cliffs, New Jersey: Prentice Hall; 1987:272-285.

43.   Färber A, Zigan N, Dolata M, Stalder P, Koppitz A, Schwabe G. The Digital Transformation of Physician–Patient Consultations: Identifying Problems and Approaches to Improve Adherence. In: Proceedings of the 52nd Hawaii International Conference on System Sciences.: HICSS; 2019 Presented at: HICSS'19; January 8-11, 2019; Maui, Hawaii p. 4125-4134. [doi: 10.24251/hicss.2019.500]

44.   Sanner TA, Øvrelid E. Informing hospital workflow coordination. Comput Supported Coop Work 2019 May 28 [FREE Full text] [doi: 10.1007/s10606-019-09362-z]

45.   Forrester JK, Rafaeli A, Dimotakis NE, Graham KA, Shaulov A, Shepard AK, et al. Employee reactions to uncertainty: an exploration of individual, role-based, & situational factors. Acad Manag 2019;2019(1):15640. [doi: 10.5465/ambpp.2019.15640symposium]

46.   Parush A. Designing health information for mutual empowerment in the joint journey of patients and healthcare professionals. Stud Health Technol Inform 2019;257:341-345. [Medline: 30741220]

47.   Dong J, Yom-Tov E, Yom-Tov GB. The impact of delay announcements on hospital network coordination and waiting times. Manag Sci 2019;65(5):1969-1994 [FREE Full text] [doi: 10.1287/mnsc.2018.3048]

48.   Taylor S. Waiting for service: the relationship between delays and evaluations of service. J Mark 1994;58(2):56-69. [doi: 10.2307/1252269]

49.   Ibrahim R, Armony M, Bassamboo A. Does the past predict the future? The case of delay announcements in service systems. Manag Sci 2017;63(6):1762-1780 [FREE Full text] [doi: 10.1287/mnsc.2016.2425]

50.   Maister D. The psychology of waiting lines. In: Czepiel JA, Solomon MR, Surprenant CF, editors. The Service Encounter: Managing Employee/Customer Interaction in Service Business. Lexington, Massachusetts: Lexington Books; 1985:113-123.

51.   Batt RJ, Terwiesch C. Waiting patiently: an empirical study of queue abandonment in an emergency department. Manag Sci 2015;61(1):39-59 [FREE Full text] [doi: 10.1287/mnsc.2014.2058]

Abbreviations

CT: computed tomography
DMZ: demilitarized zone
ED: emergency department
HCP: health care provider
IS: information system
LoS: length of stay
RA: research assistant
TAM: technology acceptance model
UI: user interface
US: ultrasound

XSL·FO
RenderX

XSL•FO

RenderX

<u>Original Paper</u>

Effectiveness of the Fun for Wellness Web-Based Behavioral Intervention to Promote Physical Activity in Adults With Obesity (or Overweight): Randomized Controlled Trial

Nicholas D Myers[1], PhD; Adam McMahon[2], MSc; Isaac Prilleltensky[2], PhD; Seungmin Lee[1], MSc; Samantha Dietz[2], PhD; Ora Prilleltensky[2], PhD; Karin A Pfeiffer[1], PhD; André G Bateman[1], MSc; Ahnalee M Brincks[1], PhD

[1]Michigan State University, East Lansing, MI, United States
[2]University of Miami, Coral Gables, FL, United States

Corresponding Author:
Nicholas D Myers, PhD
Michigan State University
201 IM Sports Circle Building
308 W Circle Drive
East Lansing, MI, 48824
United States
Phone: 1 5174325380
Email: myersni1@msu.edu

Abstract

Background: Insufficient physical activity in the adult population is a global pandemic. Fun for Wellness (FFW) is a self-efficacy theory- and Web-based behavioral intervention developed to promote growth in well-being and physical activity by providing capability-enhancing opportunities to participants.

Objective: This study aimed to evaluate the effectiveness of FFW to increase physical activity in adults with obesity in the United States in a relatively uncontrolled setting.

Methods: This was a large-scale, prospective, double-blind, parallel-group randomized controlled trial. Participants were recruited through an online panel recruitment company. Adults with overweight were also eligible to participate, consistent with many physical activity–promoting interventions for adults with obesity. Also consistent with much of the relevant literature the intended population as simply adults with obesity. Eligible participants were randomly assigned to the intervention (ie, FFW) or the usual care (ie, UC) group via software code that was written to accomplish equal allocations to the FFW and UC groups. Data collection was Web based, fully automated, and occurred at three time points: baseline, 30 days after baseline (T2), and 60 days after baseline (T3). Participants (N=461) who were assigned to the FFW group (n_FFW=219) were provided with 30 days of 24-hour access to the Web-based intervention. A path model was fit to the data consistent with the FFW conceptual model for the promotion of physical activity.

Results: There was evidence for a positive direct effect of FFW on transport-related physical activity self-efficacy (beta=.22, *P*=.02; *d*=0.23), domestic-related physical activity self-efficacy (beta=.22, *P*=.03; *d*=0.22), and self-efficacy to regulate physical activity (beta=.16, *P*=.01; *d*=0.25) at T2. Furthermore, there was evidence for a positive indirect effect of FFW on physical activity at T3 through self-efficacy to regulate physical activity at T2 (beta=.42, 95% CI 0.06 to 1.14). Finally, there was evidence for a null direct effect of FFW on physical activity (beta=1.04, *P*=.47; *d*=0.07) at T3.

Conclusions: This study provides some initial evidence for both the effectiveness (eg, a positive indirect effect of FFW on physical activity through self-efficacy to regulate physical activity) and the ineffectiveness (eg, a null direct effect of FFW on physical activity) of the FFW Web-based behavioral intervention to increase physical activity in adults with obesity in the United States. More broadly, FFW is a scalable Web-based behavioral intervention that may effectively, although indirectly, promote physical activity in adults with obesity and therefore may be useful in responding to the global pandemic of insufficient physical activity in this at-risk population. Self-efficacy to regulate physical activity appears to be a mechanism by which FFW may indirectly promote physical activity in adults with obesity.

Trial Registration: ClinicalTrials.gov NCT03194854; https://clinicaltrials.gov/ct2/show/NCT03194854.

XSL•FO
RenderX

Introduction

Background

The objective of this study was to evaluate the effectiveness of the Fun for Wellness (FFW) intervention to increase physical activity in adults with obesity in the United States in a relatively uncontrolled (ie, real world) setting. The study described in this paper was conceptualized as an effectiveness trial (ie, participants were recruited via a national health care panel recruitment company) that built upon a 2015 FFW efficacy trial completed in a relatively controlled setting (ie, participants were recruited at a major research university in the United States) [1-3]. This study is important from a general scientific perspective because the potential utility of interventions should be evaluated under both more controlled (eg, scientifically ideal: an efficacy trial) and less controlled (eg, real-world ideal: an effectiveness trial) conditions [4,5]. Before describing the FFW intervention, we begin with a summary of the 2015 FFW efficacy trial and then introduce key components in this study: target population (ie, adults with obesity), proposed outcome (ie, physical activity) and mediator (ie, self-efficacy), and the theoretical framework (ie, self-efficacy theory).

2015 Fun for Wellness Efficacy Trial

A randomized controlled trial completed in 2015 provided the initial test of the efficacy of the FFW intervention to promote well-being [1-3]. The FFW intervention was conceptualized as exerting both a positive direct effect and a positive indirect effect through self-efficacy on well-being. Data collection occurred within a relatively controlled environment (ie, adult employees at a major research university in the United States). Results provided some initial evidence for the efficacy of FFW to promote well-being self-efficacy [3]; interpersonal, community, psychological, and economic subjective well-being [1]; and interpersonal and physical well-being actions [2]. The effectiveness trial described in this paper sought to follow up on the initial evidence provided in the 2015 FFW efficacy trial.

Adults With Obesity

Approximately 2 billion adults are overweight per the World Health Organization (WHO) [6]. Moreover, approximately one-third of adults who are overweight can more precisely be classified as adults with obesity and the size of this subgroup has tripled over the past few decades [6]. In the United States, more than 40% of women and 35% of men are obese [7]. From a public health perspective, this trend toward an increasing number of adults with obesity is problematic because obesity is a risk factor for major noncommunicable chronic diseases such as cardiovascular disease, type II diabetes, musculoskeletal disorders, and some cancers [8]. To reduce the prevalence of adults with obesity, the WHO recommends that individuals increase energy intake from high-quality food sources (eg, raw vegetables), limit energy intake from low-quality food sources (eg, highly processed foods high in fat), and engage in a recommended amount of physical activity for health [6]. Examples of a recommended amount of physical activity for health in adults include at least 150 min per week of moderate-intensity physical activity or at least 75 min per week of vigorous-intensity physical activity, or an equivalent combination of the two recommendations listed above [9,10]. However, there is evidence that a very small percentage (eg, <5%) of adults with obesity meet the public health guidelines for physical activity [11]. Fortunately, there is also evidence that cognitive behavioral interventions can successfully promote physical activity in adults with obesity [12,13] and in the more general adult population [9].

Physical Activity

Physical activity has been defined as bodily movement produced by skeletal muscles that requires energy expenditure [14]. Insufficient physical activity in the adult population is a global pandemic [15,16]. Successfully addressing this pandemic will require ongoing and wide implementation of a variety of intervention strategies (eg, community-wide, informational, behavioral, social, policy, and built environment) at multiple levels of society (eg, individual, neighborhood, municipality, and country) across the globe [17,18]. At the individual level, there is evidence that behavioral interventions designed to promote physical activity by focusing on personal psychological attributes (eg, self-efficacy) can be effective [19-21]. Delivering a physical activity intervention online has been shown to be an effective mode of delivery [22,23] that also may allow for efficient scaling up of an intervention [18]. Thus, a readily scalable, Web-based behavioral intervention that effectively promotes physical activity in adults with obesity may be useful in responding to a global pandemic (ie, physical inactivity) in an at-risk population (ie, adults with obesity).

Self-Efficacy Theory

The social cognitive theory [24] has provided the theoretical framework for many effective cognitive behavioral physical activity–promoting interventions for adults with obesity [12,13]. Self-efficacy theory [25] resides within social cognitive theory and views an individual as a proactive agent in the regulation of their emotions, cognitions, and behaviors. Self-efficacy beliefs play a primary role in the self-efficacy theory and are defined as domain-specific judgments held by an individual about their ability to successfully execute differing levels of performance given certain situational demands. Self-efficacy beliefs rely upon the cognitive processing of several potential sources of efficacy information: enactive mastery experiences, vicarious experiences, verbal persuasion, and physiological and emotional states. Furthermore, two proposed omnibus outcomes of self-efficacy beliefs are an individual's thought patterns (eg, goal setting, worry, and attributions) and behaviors (eg, challenges undertaken, effort expended on challenges undertaken, and persistence in the face of difficulties that arise during challenges undertaken). A necessary condition for valid

testing of self-efficacy theory is concordance between the domain-specific self-efficacy beliefs and the proposed outcome of interest. There is a rich literature on the potential importance of targeting self-efficacy as a potentially modifiable mediating variable in physical activity–promoting interventions [19-21].
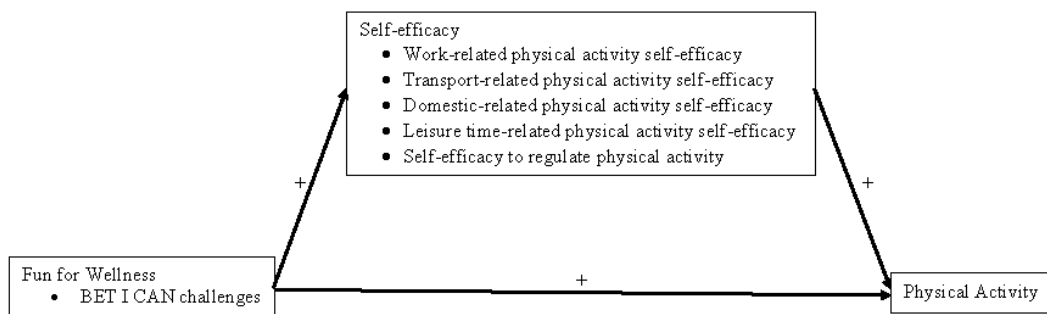
The self-efficacy theory posits that a self-efficacy–level construct may play a central role in the initiation of a behavior (eg, engaging in a recommended amount of weekly physical activity), whereas self-efficacy to regulate a behavior construct may play a central role in the maintenance of a behavior (eg, engaging in a recommended amount of weekly physical activity over time) [25]. A self-efficacy–level construct can be defined as an individual's beliefs in his or her ability to accomplish levels of a task (eg, engage in at least 150 min of moderate-intensity physical activity in the next week). Self-efficacy to regulate a behavior construct can be defined as an individual's beliefs to overcome possible barriers to accomplishing a task that he or she already knows how to do (eg, engage in at least 150 min of moderate-intensity physical activity in the next week if you are under personal stress). The importance of both a self-efficacy level construct and self-efficacy to regulate a behavior construct has been demonstrated in exercise contexts [26,27]. However, there still exists a pressing need to systematically test self-efficacy

theory–based interventions to promote physical activity in real-world settings [5,9,21].

Fun for Wellness

FFW is a self-efficacy theory–based, online (ie, Web-based and not an app) behavioral intervention developed to promote growth in well-being and physical activity by providing capability-enhancing opportunities to participants [28]. The full conceptual model for the FFW intervention is broader than this study and specifies that FFW exerts both a positive direct effect and a positive indirect effect through self-efficacy (ie, well-being self-efficacy, well-being action self-efficacy, physical activity self-efficacy, self-efficacy to regulate physical activity) on well-being (ie, subjective well-being, well-being actions, and physical activity). The narrower focus of this study was on the FFW conceptual model for the promotion of physical activity (see Figure 1). Consistent with the self-efficacy theory [24,25], the behaviors, emotions, thoughts, interactions, context, awareness, and next steps (BET I CAN) challenges provided in the FFW intervention (described in the next section) are specified as positive sources of self-efficacy information that exert a positive direct effect on self-efficacy beliefs, which are then specified to exert a positive direct effect on physical activity (ie, a behavior) [28]. Thus, self-efficacy is specified as a mediating variable in the FFW conceptual model for the promotion of physical activity.

Figure 1. The Fun for Wellness conceptual model for the promotion of physical activity. BET I CAN: behaviors, emotions, thoughts, interactions, context, awareness, and next steps.

Behaviors, Emotions, Thoughts, Interactions, Context, Awareness, and Next Steps Challenges

The self-efficacy theory provided the theoretical framework that guided the creation of capability-enhancing learning opportunities (ie, the BET I CAN challenges) with which FFW participants engage [1]. The capability-enhancing learning opportunities provided to participants exist in the form of 152 interactive and scenario-based challenges organized in the on-line environment by the BET I CAN acronym. The *B*ehavior-focused challenges are intended to increase a participant's capabilities to set a goal and to create positive habits. The *E*motion-focused challenges are intended to increase a participant's capabilities to cope with negative emotions and to cultivate positive emotions. The *T*hought-focused challenges are intended to increase a participant's capabilities to challenge negative assumptions and to create a new narrative for their life. The *I*nteraction-focused challenges are intended to increase a participant's capabilities to communicate and connect with

others. The *C*ontext-focused challenges are intended to increase a participant's capabilities to read cues and to change cues in the environment. The *A*wareness-focused challenges are intended to increase a participant's capabilities to know themselves and to know the issue. The *N*ext steps–focused challenges are intended to increase a participant's capabilities to make a plan and to stick with it. The scientific literature for each type of BET I CAN challenge has been reviewed elsewhere [28].

The capability-enhancing learning opportunity within each of the 152 BET I CAN challenges provides each FFW participant with exposure to one or more of Bandura's potential sources of self-efficacy information [3]. More specifically, each BET I CAN challenge requires a participant to do one of the following activities: (1) play an interactive game, (2) watch vignettes performed by professional actors, (3) listen and read minilectures narrated by a coach, and (4) engage in self-reflection exercises and chat rooms. An opportunity for an enactive mastery experience is provided when a participant plays an interactive

BET I CAN game. An opportunity for a vicarious experience is provided when a participant watches a BET I CAN vignette performed by professional actors. An opportunity to be verbally persuaded is provided when a participant listens to a BET I CAN minilecture narrated by a coach. An opportunity for assessing relevant physiological and emotional states is provided when a participant is asked to engage in a BET I CAN self-reflection exercise. The scientific literature supporting each of these proposed sources of self-efficacy information in physical activity contexts has been reviewed elsewhere [21,25,29].

Self-Efficacy Beliefs

Both a self-efficacy–level construct (ie, physical activity self-efficacy) and self-efficacy to regulate a behavior construct (ie, self-efficacy to regulate physical activity) are included in the FFW conceptual model for the promotion of physical activity [28]. Physical activity self-efficacy has been defined in the FFW context as the degree to which an individual perceives that they have the capability to engage in a recommended amount of weekly physical activity for health. Self-efficacy to regulate physical activity has been defined in the FFW context as the degree to which an individual perceives that they have the capability to overcome possible barriers to engagement in a recommended amount of weekly physical activity for health.

Both the physical activity self-efficacy construct and the self-efficacy to regulate physical activity construct were recently added to the FFW conceptual model based on two key results and limitations from the 2015 FFW efficacy trial [28]. First, although results from the 2015 FFW efficacy trial provided some initial evidence for the efficacy of FFW to promote physical well-being actions [2], measurement of physical well-being actions consisted of only 2 items (ie, how often do you engage in moderate physical activity such as brisk walking for about 30 min at least five times a week and eat mostly a plant-based diet of foods such as fruits, vegetables, nuts, and seeds). This study seeks to address this limitation by more thoroughly measuring physical activity across four general domains of life: leisure related, domestic related, work related, and transport related [30,31]. Second, although results from the 2015 FFW efficacy trial provided some initial evidence for the efficacy of FFW to promote self-efficacy [3], measurement of self-efficacy focused on well-being self-efficacy (ie, the degree to which an individual perceives that they have the capability to attain a positive status in key domains of their life) and, thus, was not very concordant with physical activity. This study seeks to address this limitation by more thoroughly measuring the self-efficacy beliefs for physical activity (ie, leisure-related, domestic-related, work-related, and transport-related physical activity self-efficacy and the self-efficacy to regulate physical activity).

Hypotheses

In all, three construct-level a priori hypotheses were investigated in this study based on the conceptual model depicted in Figure 1. Hypothesis 1 was that the FFW intervention would exert a positive direct effect on self-efficacy. Hypothesis 2 was that self-efficacy would exert a positive direct effect on physical activity. Hypothesis 3 was that the FFW intervention would

exert a positive direct effect on physical activity. An additional construct-level exploratory hypothesis (ie, hypothesis 4) was also investigated based on the conceptual model depicted in Figure 1: the FFW intervention would exert a positive indirect effect on physical activity through self-efficacy. Dimension-specific hypotheses were not made because of a lack of previous research on the effectiveness of the FFW intervention to promote physical activity.

Methods

The Well-Being and Physical Activity Study

The data described in this paper were collected within a more broadly focused trial, the Well-Being and Physical Activity Study (ClinicalTrials.gov, identifier: NCT03194854). Within this section, we provide an overview of the relevant methods used in the Well-Being and Physical Activity Study to provide a context for the specific focus of this paper [32]. The readers are referred to the relevant protocol paper [28] for a fuller description of the protocol for the Well-Being and Physical Activity Study. A populated Consolidated Standards of Reporting Trials-EHEALTH checklist is provided in Multimedia Appendix 1.

Ethics Approval

All procedures in this study involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The institutional review board (IRB) at the University of Miami provided necessary permission to conduct this study on July 11, 2017, IRB number 20170541. The University of Miami and Michigan State University (STUDY00000979) established an Institutional Authorization Agreement on June 26, 2018, that provided permission for the University of Miami to serve as the designated IRB for this study.

Study Design

The study design was a large-scale, prospective, double-blind (ie, investigators and outcome assessor were masked), parallel-group randomized controlled trial. Recruiting, screening, random assignment, and collection of data were conducted online from August 2018 through November 2018. Data collection was Web based, fully automated, and occurred at three time points: baseline (T1), 30 days after baseline (T2), and 60 days after baseline (T3). The timeline for this study was similar to timelines used in other physical activity interventions in adults with obesity [12,13].

Recruitment and Eligibility

A sample size of approximately 900 participants was targeted for enrollment in the study. Participants were recruited through the general population panel of the SurveyHealth recruitment company. Partnering with a panel recruitment company is consistent with recruitment in preliminary research on FFW [33,34] and with a movement toward larger and smarter physical activity promotion interventions [18]. Eligibility criteria were (a) the ability to access the Web-based intervention, (b) living in the United States, (c) aged 18 to 64 years, (d) BMI of 25.00

kg/m^2 or more, and (e) absence of simultaneous enrollment in another intervention program promoting either well-being or physical activity. The BMI criterion included both the overweight (ie, 25.00-29.99 kg/m^2) category and the obese category (ie, ≥30.00 kg/m^2) consistent with many physical activity–promoting interventions for adults with obesity [12,35].

Informed Consent

Informed consent was obtained from each participant included in the study. More specifically, immediately after being determined to be eligible for this study, each eligible individual was directed to a Web-based, IRB-approved informed consent form. Each individual who clicked *Consent to Participate* was enrolled as a participant in the study. Each individual who clicked *Decline to Consent* was denied access to the study.

Random Assignment

Random assignment of each eligible participant occurred after a unique and secure login credential was created, informed consent was obtained, a medical disclaimer was agreed to, and the T1 survey battery was completed. Eligible participants were randomly assigned to the intervention (ie, FFW) or the usual care (ie, UC) group via software code that was written to accomplish equal allocations to the FFW and UC groups. Participants assigned to the FFW group were given immediate access to the intervention. Participants assigned to the UC group were put on a waitlist for access to the intervention. Both the FFW group and the UC group were provided with modest financial incentives to provide data consistent with a general approach taken in many theory-based physical activity–promoting interventions [9]. The authors of this study are unaware of any previous research that would support casting unique doubt on the results of this study (as compared with other theory-based physical activity–promoting interventions that used modest financial incentives in a study) attributable to the particular financial incentives approach taken in this study.

Usual Care

Participants assigned to the UC group were asked to conduct their lives as usual. The login credential for each UC participant provided access to a secure website to complete the survey battery at T1, T2, and T3. UC participants had the opportunity to earn up to US $30 worth of Amazon electronic gift cards. Specifically, UC participants could earn US $5 for completing the T1 survey battery, US $10 for completing the T2 survey battery, and US $15 for completing the T3 survey battery. UC participants were given 1 month of 24-hour access to the FFW intervention after data collection for this study was closed.

Fun for Wellness

Participants assigned to the FFW group were asked to engage with the FFW intervention. The login credential for each FFW participant provided 30 days (ie, from T1 to T2) of 24-hour access to the 152 BET I CAN challenges and access to a secure website to complete the survey battery at T1, T2, and T3. FFW participants had the opportunity to earn a total of US $45 worth of Amazon electronic gift cards. Specifically, FFW participants could earn US $5 for completing the T1 survey battery, US $10 for completing both the T2 survey battery and at least 15 BET

I CAN postintroductory challenges, an additional US $15 for completing at least 30 BET I CAN post-introductory challenges, and US $15 for completing the T3 survey battery.

Each of the first four BET I CAN challenges required the participant to do one of the aforementioned activities while focusing on introductory material (orientation to the website, examples of a recommended amount of physical activity for health, etc) to provide an important context for capability-enhancing opportunities [25]. Participants were required to complete these introductory challenges to gain access to the remaining 148 postintroductory BET I CAN challenges. Participants self-selected which postintroductory BET I CAN challenges to complete. Challenges completed by each participant were tracked by computer software to provide data (ie, participation points) for the FFW engagement scoring system [1]. Earning at least 21 participation points was the operational definition for being engaged with the FFW intervention [28].

Survey Battery

Instruments designed to measure demographic information, self-efficacy, and physical activity were included in the survey battery. Proposed demographic and biological correlates of physical activity were collected via self-reporting at T1 and included participant age, gender, race/ethnicity, highest level of education completed, marital status, employment status, and household annual income [19]. This set of demographic and biological variables is collectively referred to as the demographic covariates from this point forward.

Physical Activity

Physical activity was measured at T1 through T3 with the long form of the international physical activity questionnaire (IPAQ [30,31]). The long form of the IPAQ is intended for individuals aged 15 to 69 years and purports to measure physical activity in four domains—work related, transport related, domestic related, and leisure time related—according to the frequency and duration of the physical activity performed in each domain during a week. The physical activity domains measured in the IPAQ are separated according to their intensity, which is defined as a distinction between walking, moderate physical activities, and vigorous physical activities. Moderate physical activity is defined as activities that take moderate physical effort and make you breathe somewhat harder than normal. Vigorous physical activity is defined as activities that take hard physical effort and make you breathe much harder than normal.

A total physical activity score—which is the sum of total walking time, total time in moderate physical activity, and total time in vigorous physical activity—was created based on the IPAQ data processing guidelines [36]. Total walking time is the sum of walking time in the work-related, transport-related, and leisure-related domains. Total time in moderate physical activities is the sum of moderate physical activity in the work-related, transport-related, domestic-related, and leisure-related domains. Total time in vigorous physical activities is the sum of vigorous physical activity in the work-related, domestic-related, and leisure-related domains. Outlying cases (ie, averaging 16 hours or more of physical

activity per day) were excluded from analysis based on IPAQ data processing guidelines for excluding outliers [36].

Self-Efficacy

Overall, five domains of self-efficacy were measured at T1 through T3. Each of the four physical activity self-efficacy–level domains was measured with a slightly modified version of the well-established 8-item exercise self-efficacy (EXSE; [26]) scale. The EXSE scale assesses an individual's beliefs in their ability to continue exercising on a 3-times-per-week basis at moderate intensities for more than 40 min per session in the future. The EXSE scale was tailored for the FFW context to assess the degree to which an individual perceives that they have the capability to engage in a recommended amount of weekly physical activity for health. Work-related physical activity self-efficacy was measured with a 12-item scale that was designed to be concordant with how work-related physical activity is measured in the IPAQ (ie, at both a vigorous and moderate intensity). Transport-related physical activity self-efficacy was measured with a 6-item scale that was designed to be concordant with how transport-related physical activity is measured in the IPAQ (ie, at a moderate intensity). Domestic-related physical activity self-efficacy was measured with a 6-item scale that was designed to be concordant with how domestic-related physical activity is measured in the IPAQ (ie, at a moderate intensity). Leisure-related physical activity self-efficacy was measured with a 12-item scale that was designed to be concordant with how leisure-related physical activity is measured in the IPAQ (ie, at both a vigorous and a moderate intensity). Vigorous-intensity items began with the stem "how confident are you in your current ability to engage in *work-* or *leisure*-related physical activity at a vigorous level of intensity" and then referenced six increasing periods (eg, for at least 10, 15, 30, 45, 60, or 75 min in the next week). Moderate-intensity items began with the stem "how confident are you in your current ability to engage in *work-* or *transport-* or *domestic-* or *leisure*-related physical activity at a moderate level of intensity" and then referenced six increasing time periods (eg, for at least 10, 30, 60, 90, 120, or 150 min in the next week). Responses to each item were organized within a 5-category rating scale structure, where 0=no, 1=low, 2=moderate, 3=high, and 4=complete confidence based on previous research on effective self-efficacy rating scale structures [37]. An average observed score for each of the four physical activity self-efficacy level domains was created based on relevant guidelines [26,38].

Self-efficacy to regulate physical activity was measured at T1 through T3 with a slightly modified version of the well-established 13-item barriers self-efficacy (BARSE) scale [27]. The BARSE scale assesses an individual's perceived capabilities to exercise 3 times per week for 40 min over the next 2 months in the face of commonly identified barriers to participation. The BARSE scale was tailored for the FFW context to assess the extent to which an individual believes that he or she has the ability to overcome possible barriers to engagement in a recommended amount of weekly physical activity for health. Responses to each item were organized within a 5-category rating scale structure, where 0=no, 1=low, 2=moderate, 3=high, and 4=complete confidence. An average

observed score for self-efficacy to regulate physical activity was created based on relevant guidelines [27,38].

Data Analytic Approach

Statistical models were fit in Mplus 8.3 with maximum-likelihood (ML) estimation with robust SEs [39]. Type I error rate was set equal to 0.05. Missing data were addressed with full information ML estimation using the observed information matrix under the assumption of missing at random [40]. Reliability was assessed using Cronbach alpha [41,42]. Indexes of effect size considered for direct effects were Cohen d [43] and percentage of variance explained. Commonly used heuristics were used to assist in the interpretation of an absolute value of Cohen d: 0.20 (small), 0.50 (medium), and 0.80 (large). For each indirect effect, a bias-corrected bootstrapped estimate of the 95% confidence interval was obtained with the number of draws set equal to 2000 [44]. An index of effect size was not considered for indirect effects because an effect size index for complex mediation models has not yet been firmly established [45].

Path Model

A single saturated (degrees of freedom=0) path model was fit consistent with the conceptual model depicted in Figure 1 under an intention-to-treat approach [46]. Each of the five domains of self-efficacy at T2 were regressed on FFW (ie, 0=UC, 1=FFW), physical activity at T1, and the demographic covariates. Physical activity at T3 was regressed on FFW, the five domains of self-efficacy at T2, physical activity at T1, and the demographic covariates. The expression *adjusted mean difference* is used from this point forward to acknowledge the statistical adjustment made by including covariates in the model.

There were four sets of focal parameters in the path model. The first set of focal parameters was the direct effect of FFW on each of the five domains of self-efficacy at T2 (ie, $beta_1$). Each of these five parameters was interpreted as the adjusted mean difference on a particular domain of self-efficacy at T2 for the FFW group as compared with the UC group. The second set of focal parameters was the direct effect of the five domains of self-efficacy at T2 on physical activity at T3 (ie, $beta_2$). Each of these five parameters was interpreted as the path coefficient from a particular domain of self-efficacy at T2 to physical activity at T3. The third set of focal parameters was a single parameter: the direct effect of FFW on physical activity at T3 (ie, $beta_3$). This parameter was interpreted as the adjusted mean difference on physical activity at T3 for the FFW group as compared with the UC group. The fourth set of focal parameters was the indirect effect of FFW on physical activity at T3 through each of the five domains of self-efficacy at T2 (ie, $beta_4$, where $beta_4 = beta_1 * beta_2'$). Each of these five parameters was interpreted as the product of path coefficients from FFW to physical activity at T3 through a particular domain of self-efficacy at T2. Each set of focal parameters tested the numerically corresponding hypothesis (eg, $beta_1$ tested hypothesis 1).

Necessary Sample Size

Necessary sample size was determined for a minimum fixed level of power (ie, 0.80) for rejecting the null hypothesis that each of the five focal parameters regarding a direct effect of FFW (ie, $beta_2$ and $beta_3$) was equal to 0.00 using Monte Carlo methods as implemented in Mplus 8.3 [47]. The population parameter value for each of the five relevant focal parameters was set equal to a value that corresponded to a small-to-moderate effect size (ie, $d=.35$) consistent with relevant results from previous research [2,3]. Type I error was set equal to 0.05. The number of replications was set to 10,000, and the necessary sample size was equal to 285.

Results

Participant Characteristics

Figure 2 depicts participant flow from eligibility screening to randomization to retention over the three measurement occasions. A total of 821 consenting participants were randomly assigned to FFW (n=410) or UC (n=411). Forensic analysis by a computer scientist performed before data analysis identified 154 cases as fraudulent, and these cases were excluded from analysis. The researchers initiated the forensic analysis after consulting with the designated IRB, legal counsel, and the office of research compliance and quality assurance about the computer scientist's report of suspicious activity on the website (eg, participants logging in very close temporal proximity and sending identical emails to the computer scientist in broken English). The forensic analysis revealed that all of these 154 accounts were made by 1 user or group through 2 virtual private server (VPS) services. The analysis was reported as a reportable new information (RNI#00003760) incident to the designated IRB in December 2018. Unlike the 154 fraudulent cases (ie, 154/821, 18.8%), no groupings of the 667 nonfraudulent cases (667/821, 81.2%) appeared to have been made by 1 user or group through VPS services.

Figure 2. Participant flow from screening to randomization to retention over the three measurement occasions for the physical activity–related data.
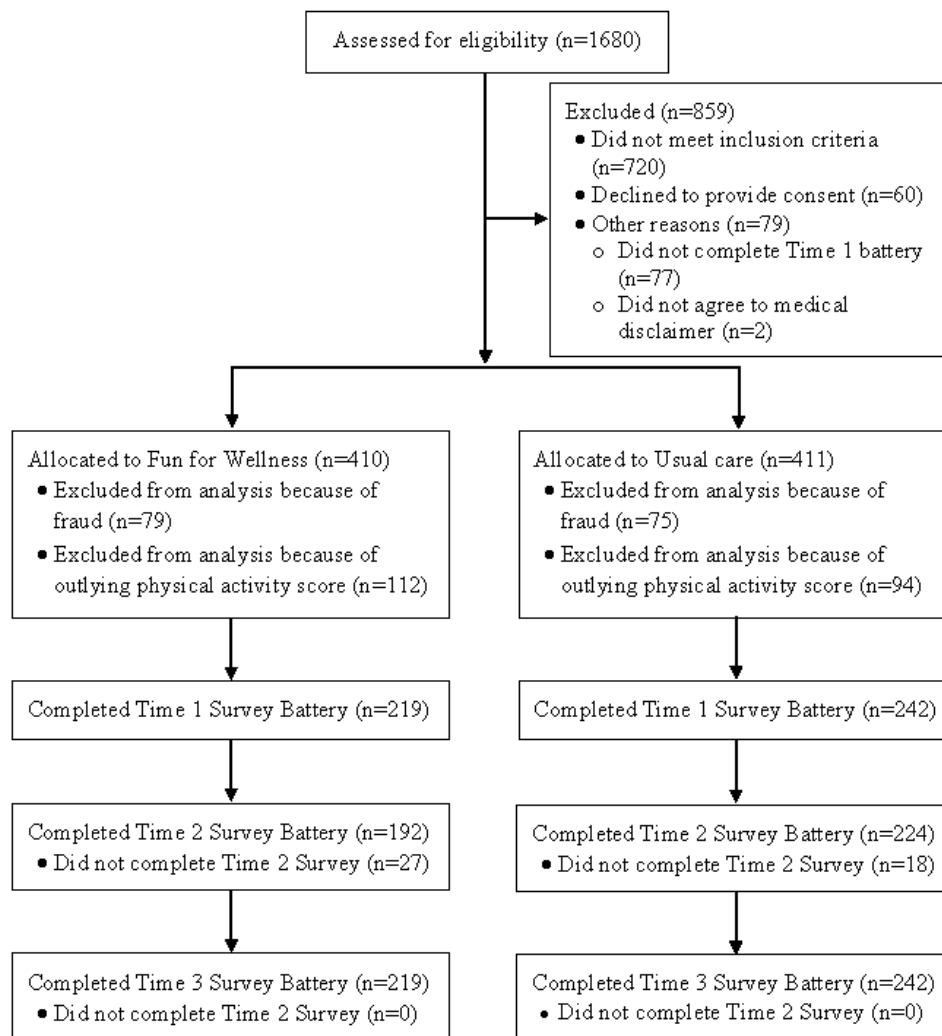
An additional 206 cases were outlying cases on the physical activity score and were excluded from the analysis, leaving 461 analyzed cases (ie, participants), FFW (n=219) and UC (n=242). A majority of the participants identified as female (302/461, 65.5%), white, non-Hispanic (342/461, 74.2%), having completed at least a 4-year college degree (307/461, 66.5%), married (314/461, 68.2%), being a full-time employee (309/461, 67.0%), being at least 40 years old (254/461, 55.1%), and as

residing in a household with an annual income of at least US $75,000 (238/461, 51.6%). The difference in the proportion of missing data observed at T2 for the FFW group (ie, 0.12) as compared with the UC group (ie, 0.07) was not statistically significant (P=.08.)

Table 1 provides a comparison of demographic characteristics, BMI values, self-efficacy domain scores, and physical activity scores at T1 for participants by randomization group. There were no statistically significant differences in the proportions (for binary variables tested via logistic regression) or means (for continuous variables tested via linear regression) of

demographic characteristics, the mean BMI value, the mean self-efficacy domain scores, or the mean physical activity scores at T1 by randomization group. The minimum value of BMI observed across the sample was 25.06 kg/m^2. The median values of physical activity in hours per week (ie, 10.61 and 9.18) were similar to IPAQ-based values in some other relevant research [48-50]. No important harms or unintended effects were observed in either group. Cronbach alpha ranged from .86 (physical activity) to .97 (work-related physical activity self-efficacy). A majority (201/219, 91.7%) of the participants who were assigned to the FFW group were engaged with the FFW intervention.

Table 1. Descriptive statistics for demographic characteristics, self-efficacy domain scores, and physical activity scores at baseline for participants by randomization group (N=461).

Variable[a]	Usual care (n=242)	Fun for Wellness (n=219)
Female, n (%)	157 (64.9)	144 (66.0)
Black, n (%)	41 (16.9)	31 (14.2)
Hispanic, n (%)	17 (7.0)	13 (5.9)
Vocational or technical school, n (%)	17 (7.0)	15 (6.9)
Some college, n (%)	37 (15.3)	39 (18.0)
Undergraduate degree, n (%)	111 (46.0)	85 (38.7)
Graduate or professional degree, n (%)	54 (22.4)	56 (25.7)
Married, n (%)	165 (68.2)	149 (68.1)
Part-time employment, n (%)	28 (11.6)	19 (8.7)
Full-time employment, n (%)	158 (65.3)	151 (69.0)
Retired, n (%)	21 (8.7)	19 (8.8)
Age (years), mean (SD)	41.97 (11.03)	41.77 (10.78)
Income in thousand dollars, mean (SD)	76.38 (47.73)	77.77 (48.20)
BMI (kg/m^2), mean (SD)	30.92 (5.83)	30.21 (5.31)
Work-related physical activity self-efficacy (alpha=.97), mean (SD)	1.19 (1.15)	1.17 (1.14)
Transport-related physical activity self-efficacy (alpha=.95), mean (SD)	1.20 (1.11)	1.28 (1.16)
Domestic-related physical activity self-efficacy (alpha=.95), mean (SD)	1.43 (1.22)	1.52 (1.25)
Leisure time–related physical activity self-efficacy (alpha=.97), mean (SD)	1.37 (1.17)	1.39 (1.25)
Self-efficacy to regulate physical activity (alpha=.90), mean (SD)	2.06 (0.74)	2.05 (0.70)
Physical activity in hours per week (alpha=.86), median (IQR)	10.61 (19.17)	9.18 (17.65)

[a]The reference group (eg, male) for each demographic variable (eg, gender) and subgroups comprising less than 5% of observations are not reported for spatial reasons. Missing data ranged from 0% to 3.5% across all the variables in this table.

Path Model

The percentage of variance accounted for ranged from 16.8% (work-related physical activity self-efficacy) to 25.3% (self-efficacy to regulate physical activity) across the five domains of self-efficacy at T2 and equaled 37.4% for physical activity at T3. The correlations among the residuals of the four self-efficacy–level constructs ranged from 0.74 (work-related physical activity self-efficacy with leisure-related physical activity self-efficacy) to 0.76 (transport-related physical activity self-efficacy with domestic-related physical activity self-efficacy). The correlations between the residuals of the four self-efficacy–level constructs with self-efficacy to regulate

physical activity ranged from 0.04 (transport-related physical activity self-efficacy with self-efficacy to regulate physical activity) to 0.13 (work-related physical activity self-efficacy with self-efficacy to regulate physical activity). The unstandardized estimates of the covariates are available in Table 2, but these estimates are not discussed because of spatial limitations. Table 3 provides the unstandardized estimate of each focal parameter from the path model by hypothesis. Figure 3 provides key focal unstandardized parameter estimates for hypothesis 1 through hypothesis 3. Estimates for hypothesis 4 are not directly provided in Figure 3 because they are not parameter estimates per SE but rather a function of existing parameter estimates. However, they are listed at the bottom of

Table 3. The paragraphs below briefly interpret these estimates   with regard to the corresponding hypothesis tested.

Table 2. Unstandardized estimate of the covariates from the path model.

Predictor	Outcome					
	Work-related physical activity self-efficacy at time 2, beta (SE)	Transport-related physical activity self-efficacy at time 2, beta (SE)	Domestic-related physical activity self-efficacy at time 2, beta (SE)	Leisure-related physical activity self-efficacy at time 2, beta (SE)	Self-efficacy to regulate physical activity at time 2, beta (SE)	Physical activity at time 3, beta (SE)
Physical activity at time 1	.01 (0.00)[a]	.01 (0.00)[b]	.01 (0.00)[a]	.01 (0.00)[b]	.01 (0.00)[c]	.46 (0.07)[b]
Female	−.17 (0.10)	−.13 (0.09)	−.02 (0.10)	−.08 (0.10)	−.10 (0.07)	1.27 (1.24)
Black	−.23 (0.12)	−.15 (0.13)	−.04 (0.14)	−.11 (0.13)	.16 (0.09)	−.32 (1.73)
Hispanic	.00 (0.21)	−.15 (0.22)	−.30 (0.22)	−.35 (0.18)[c]	.12 (0.12)	−.65 (2.61)
Vocational or technical school	.03 (0.28)	.06 (0.28)	.16 (0.31)	.29 (0.30)	.22 (0.23)	5.86 (4.44)
Some college	−.27 (0.22)	−.15 (0.25)	.11 (0.27)	.02 (0.25)	−.03 (0.21)	2.05 (3.71)
Undergraduate degree	−.12 (0.22)	.12 (0.24)	.20 (0.26)	.30 (0.25)	.13 (0.21)	.39 (3.00)
Graduate or professional degree	−.52 (0.23)[c]	−.22 (0.25)	−.12 (0.27)	−.14 (0.26)	.20 (0.21)	3.36 (3.24)
Married	−.07 (0.12)	−.15 (0.13)	−.12 (0.13)	−.09 (0.13)	−.06 (0.08)	1.62 (1.69)
Part-time employment	.11 (0.25)	−.05 (0.26)	−.32 (0.27)	−.09 (0.26)	.27 (0.17)	2.41 (3.70)
Full-time employment	−.21 (0.22)	−.58 (0.22)[a]	−.91 (0.24)[b]	−.76 (0.23)[a]	.40 (0.15)[a]	−1.81 (3.47)
Retired	−.46 (0.25)	−.79 (0.26)[a]	−.86 (0.29)[a]	−.78 (0.27)[a]	.04 (0.22)	−2.28 (4.32)
Age in years	.02 (0.01)[b]	.02 (0.01)[b]	.03 (0.01)[b]	.03 (0.01)[b]	−.02 (0.00)[b]	.18 (0.09)[c]
Income in thousand dollars	.01 (0.00)[b]	.01 (0.00)[b]	.01 (0.00)[b]	.01 (0.00)[b]	.00 (0.00)	.02 (0.02)

[a]$P<.01$, 2-tailed.

[b]$P<.001$, 2-tailed.

[c]$P<.05$, 2-tailed.

Table 3. Unstandardized estimate of each focal parameter from the path model by hypothesis.

Specific path	Beta$_1$ (SE)	95% CI	Cohen d	95% CI
Hypothesis 1: FFW[a] –> self-efficacy				
FFW –> work-related physical activity self-efficacy at time 2	.09 (0.09)	−0.09 to 0.27	0.10	−0.09 to 0.28
FFW –> transport-related physical activity self-efficacy at time 2	.22 (0.10)[b]	0.04 to 0.41	0.23	0.04 to 0.41
FFW –> domestic-related physical activity self-efficacy at time 2	.22 (0.10)[b]	0.03 to 0.41	0.22	0.03 to 0.40
FFW –> leisure-related physical activity self-efficacy at time 2	.14 (0.10)	−0.05 to 0.33	0.14	−0.04 to 0.33
FFW –> self-efficacy to regulate physical activity at time 2	.16 (0.06)[c]	0.04 to 0.29	0.25	0.07 to 0.43
Hypothesis 2: Self-efficacy –> physical activity				
Work-related physical activity self-efficacy at time 2 –> physical activity at time 3	−.17 (1.28)	−2.68 to 2.35	__[d]	—
Transport-related physical activity self-efficacy at time 2 –> physical activity at time 3	1.19 (1.28)	−1.32 to 3.6	—	—
Domestic-related physical activity self-efficacy at time 2 –> physical activity at time 3	−1.02 (1.54)	−4.03 to 1.99	—	—
Leisure-related physical activity self-efficacy at time 2 –> physical activity at time 3	3.80 (1.29)[c]	1.26 to 6.33	—	—
Self-efficacy to regulate physical activity at time 2 –> physical activity at time 3	2.55 (1.12)[b]	0.35 to 4.76	—	—
Hypothesis 3: FFW –> physical activity				
FFW –> physical activity at time 3	1.04 (1.45)	−1.80 to 3.88	0.07	−0.11 to 0.26
Hypothesis 4: FFW –> self-efficacy –> physical activity				
FFW –> work-related physical activity self-efficacy at time 2 –> physical activity at time 3	−.02 (0.12)	−0.58 to 0.27	—	—
FFW –> transport-related physical activity self-efficacy at time 2 –> physical activity at time 3	.26 (0.31)	−0.24 to 1.20	—	—
FFW –> domestic-related physical activity self-efficacy at time 2 –> physical activity at time 3	−.22 (0.34)	−1.26 to 0.3	—	—
FFW –> leisure-related physical activity self-efficacy at time 2 –> physical activity at time 3	.54 (0.41)	−0.06 to 1.76	—	—
FFW –> self-efficacy to regulate physical activity self-efficacy at time 2 –> physical activity at time 3	.42 (0.25)	0.06 to 1.14[e]	—	—

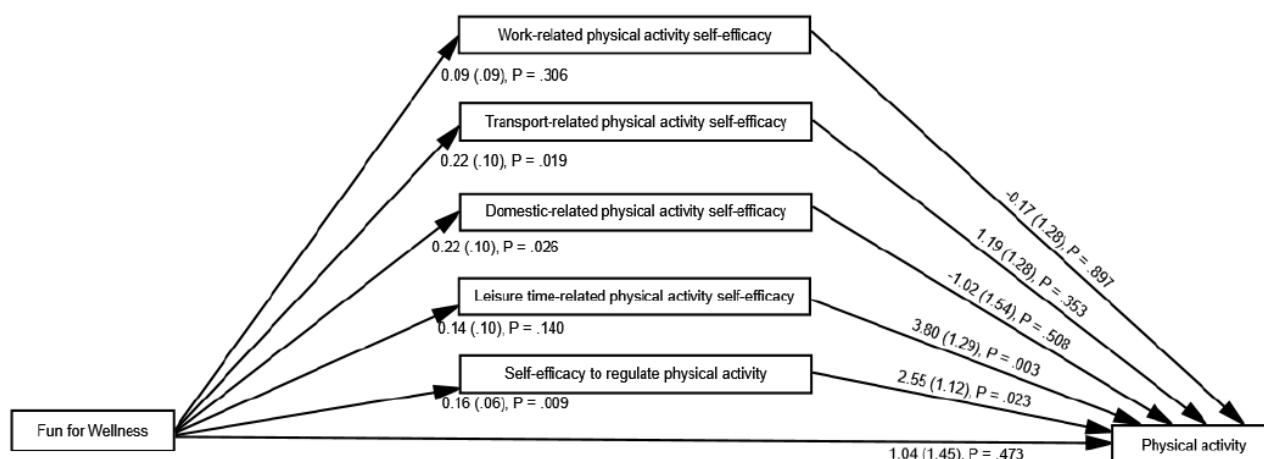[a]FFW: Fun for Wellness.

[b]$P<.05$, 2-tailed.

[c]$P<.01$, 2-tailed.

[d]Not applicable.

[e]Bias-corrected confidence interval did not include 0.

Figure 3. Key focal unstandardized parameter estimates from the path model for hypothesis 1 through hypothesis 3. The 241 nonfocal parameter estimates are not depicted to reduce clutter.

Hypothesis 1

The adjusted mean difference for the FFW group as compared with the UC group was statistically significant and approximately small in size for transport-related physical activity self-efficacy (beta=.22, P=.02; d=0.23), domestic-related physical activity self-efficacy (beta=.22, P=.03; d=0.22), and self-efficacy to regulate physical activity (beta=.16, P=.01; d=0.25) at T2. The adjusted mean difference for the FFW group as compared with the UC group was statistically nonsignificant for work-related physical activity self-efficacy (beta=.09, P=.31; d=0.10) and leisure-related physical activity self-efficacy (beta=.14, P=.14; d=0.14) at T2. Thus, only partial support was provided for hypothesis 1.

Hypothesis 2

The path coefficient to physical activity at T3 was statistically significant for leisure-related physical activity self-efficacy (beta=3.80, P=.003) and self-efficacy to regulate physical activity (beta=2.55, P=.02) at T2. The path coefficient to physical activity at T3 was statistically nonsignificant for work-related physical activity self-efficacy (beta=−.17, P=.90), transport-related physical activity self-efficacy (beta=1.19, P=.35), and domestic-related physical activity self-efficacy (beta=−1.02, P=.51) at T2. Thus, only partial support was provided for hypothesis 2.

Hypothesis 3

The adjusted mean difference on physical activity at T3 for the FFW group as compared with the UC group was statistically nonsignificant (beta=1.04, P=.47, d=0.07). Thus, no support was provided for hypothesis 3.

Hypothesis 4

The 95% CI for the product of path coefficients from FFW to physical activity at T3 through self-efficacy at T2 did not include 0.00 for self-efficacy to regulate physical activity (beta=.42, 95% CI 0.06 to 1.14). The 95% CI for the product of path coefficients from FFW to physical activity at T3 through self-efficacy at T2 included 0.00 for work-related physical activity self-efficacy (beta=−.02, 95% CI −0.58 to 0.27), transport-related physical activity self-efficacy (beta=.26, 95% CI −0.24 to 1.20), domestic-related physical activity

self-efficacy (beta=−.22, 95% CI −1.26 to 0.36), and leisure-related physical activity self-efficacy (beta=.54, 95% CI −0.06 to 1.76). Thus, only partial support was provided for hypothesis 4.

Discussion

Principal Findings

The objective of this study was to evaluate the effectiveness of the FFW Web-based behavioral intervention to increase physical activity in adults with obesity in the United States in a relatively uncontrolled setting. In general, results from this study provide both some supportive and some unsupportive initial evidence with regard to the objective of this study. Specific findings, both supportive and unsupportive, will be discussed with respect to the four construct-level hypotheses tested within the FFW conceptual model for the promotion of physical activity (see Figure 1) and to the relevant results from the 2015 FFW efficacy trial.

Partial supportive evidence was observed in this study for three of the four hypotheses tested. Supportive evidence for hypothesis 1 includes positive direct effects from the FFW intervention to transport- and domestic-related physical activity self-efficacy and self-efficacy to regulate physical activity at T2. This set of findings provides some support for the conceptualization of the BET I CAN challenges as capability-enhancing opportunities and extends the literature on the ability of FFW to promote self-efficacy beliefs [3]—a potentially modifiable mediating variable targeted by the intervention. Supportive evidence for hypothesis 2 includes positive direct effects from both leisure-related physical activity self-efficacy and self-efficacy to regulate physical activity at T2 to physical activity at T3. This pair of findings provides some support for a central contention of the self-efficacy theory—behaviors are an omnibus outcome of self-efficacy beliefs [25]—and addresses a limitation of the 2015 FFW efficacy trial: not evaluating proposed relationships between self-efficacy and physical activity [2]. Supportive evidence for hypothesis 4 includes a positive indirect effect of the FFW intervention on physical activity at T3 through self-efficacy to regulate physical activity at T2. This finding addresses a limitation of the 2015 FFW efficacy trial: not

evaluating the proposed positive indirect effect of the FFW intervention on physical activity through self-efficacy. Beyond the four hypotheses tested, this study has the potential to be important because it provides initial evidence for the effectiveness of the FFW intervention to increase physical activity (indirectly through self-efficacy to regulate physical activity) in an at-risk population [8,13]. Beyond the FFW intervention, findings from this study also contribute to a practical research need identified in the 2018 Physical Activity Guidelines Advisory Committee Scientific Report: to systematically test theory-based interventions in real-world settings [9].

At least partial unsupportive evidence was observed in this study for each of the four hypotheses tested. Unsupportive evidence for hypothesis 1 includes null direct effects from the FFW intervention to both work- and leisure-related physical activity self-efficacy at T2. Thus, it may be that the BET I CAN challenges in the FFW intervention would benefit from being further optimized for providing more meaningful exposure to relevant sources of efficacy-enhancing information with regard to these 2 domains of self-efficacy beliefs [51]. More specifically, future studies that estimate the individual effect of each BET I CAN component, and how BET I CAN components may operate synergistically with each other, may help identify active and inactive intervention components within FFW with regard to promoting self-efficacy and physical activity in adults with obesity. Unsupportive evidence for hypothesis 2 includes null direct effects from work-, transport-, and domestic-related physical activity self-efficacy at T2 to physical activity at T3. This set of null findings may be because of the relatively strong correlations among the four self-efficacy–level constructs (ie, difficult to identify unique relationships with physical activity). Unsupportive evidence for hypothesis 3 includes a null direct effect from the FFW intervention to physical activity at T3. This null finding is inconsistent with relevant results from the 2015 FFW efficacy trial [2] and may be because of differences in either model specification (ie, evaluating the direct effect of FFW on physical activity while controlling for self-efficacy beliefs in this study) or measurement of physical activity (ie, more thoroughly measuring physical activity in this study). Unsupportive evidence for hypothesis 4 includes null indirect effects from the FFW intervention to physical activity at T3 through each of the four self-efficacy–level constructs: work-, transport-, leisure-, and domestic-related physical activity self-efficacy at T2. This set of null findings may be attributable to the idea that, on average, an individual's self-efficacy beliefs regarding their capability to engage in a recommended amount of physical activity for health may be less important than an individual's self-efficacy beliefs in their capability to overcome possible barriers to their engagement in a recommended amount of weekly physical activity for health with regard to the promotion of an individual's physical activity behavior [26,27].

Conclusions

Results from this study provide some initial evidence for both the effectiveness and the ineffectiveness of the FFW Web-based behavioral intervention to increase physical activity in adults with obesity in the United States. Specifically, there is evidence that FFW may be ineffective in directly promoting physical activity in adults with obesity. Similarly, there is evidence that FFW may be ineffective in indirectly promoting physical activity through the four (ie, work-, transport-, domestic-, and leisure time–related) *self-efficacy–level constructs* (ie, the degree to which an individual perceives that they have the capability to engage in a recommended amount of weekly physical activity for health). However, there is evidence that FFW may be effective in indirectly promoting physical activity in adults with obesity by increasing an individual's *self-efficacy to regulate their physical activity* (ie, the degree to which an individual perceives that they have the capability to *overcome possible barriers to* engagement in a recommended amount of weekly physical activity for health). For this reason, we believe that the FFW Web-based behavioral intervention may have the potential to eventually become useful, in some small but important way, given the magnitude of the problem, in responding to the global pandemic of insufficient physical activity in adults with obesity by increasing an individual's self-efficacy to regulate their physical activity.

Realizing the potential for the FFW intervention to have practical implications at a local level will require future community-based studies that align with recent recommendations put forth by the Community Preventive Services Task Force [52]. More specifically, the Community Preventive Services Task Force suggests that physical activity interventions for adults with obesity should include activity monitors and promote physical activity within a more broadly focused weight management program where there is access to a health care provider. An implication from the results of this study is that a feasibility study is now underway to implement accelerometer-based assessment of physical activity within the FFW intervention in partnership with a local bariatric service center within a major health care organization in the Midwest of the United States [53]. Gaining necessary approvals for accessing medical records from participants in this ongoing feasibility study may provide important information on certain patient characteristics (eg, comorbidities) that may influence the effectiveness of the FFW intervention.

Limitations

We are aware of at least four noteworthy limitations for this study that temper the relevant conclusions that can be made. First, we recognize that our hypotheses assume additivity of FFW effects for all covariates (ie, no a priori moderators for the proposed effects of FFW). We encourage future secondary analyses that explore the prospect of heterogeneous FFW effects for subgroups of individuals (eg, comorbidities) on physical activity. Second, we note that another limitation is that all the data collected, except for engagement with the FFW intervention, were collected via self-reporting. Field-based studies that collect physical activity data from objective instrumentation [54-58] in adults with obesity are encouraged [35,52] and are underway in the FFW context [53]. This underway study is employing both self-reported and accelerometer-measured physical activity in adults with obesity, which is consistent with recommendations in previous research [59] that found the physical activity of overweight or obese individuals to be ranked higher by self-reporting than by accelerometer as compared with normal-weight individuals.

That said, it is important to note that the aforementioned published study did not provide evidence for randomized group assignment (eg, control vs experimental) as a moderator for the observed mismatch between self-reported and accelerometer-measured physical activity in overweight or obese individuals engaged in physical activity–promoting interventions. Thus, although the aforementioned study provides support for suspecting that a mismatch between self-reported and accelerometer-measured physical activity may have been observed in this study (if accelerometer-measured physical activity had been collected), it does not provide direct support for suspecting that the magnitude of the suspected mismatch may have varied as a function of randomized group assignment in this study (ie, UC group vs FFW group). The third limitation is that 360 of 820 cases (eg, 43.9%) needed to be excluded from the analyses because of either fraud (n=154) or outlying physical activity scores (n=206). Future efforts to better guard against fraud (eg, working more closely with the panel recruitment company) and possible overreporting of physical activity (eg, objective assessment of physical activity) is encouraged and may increase confidence in subsequent findings (eg, in reference to physical activity guidelines). A final limitation is that engagement data were not collected from UC participants who were given 1 month of 24-hour access to the FFW intervention (but were not provided with financial incentives to complete BET I CAN challenges) after data collection for this study was closed. Collecting these data would have provided some insight into the degree to which the very high level of engagement observed in the FFW group (ie, 201/219, 91.7%) may have been because of the inclusion of financial incentives to complete BET I CAN challenges.

Acknowledgments

Authors' Contributions

NDM, IP, and AM made substantial contributions to all facets of this work. SL made substantial contributions to the design of this work. SD made substantial contributions to the creation of intervention activities and acquisition of the data collected in this work. OP made substantial contributions to the conceptualization of this work. KAP made substantial contributions to the design of this work. AGB made substantial contributions to the design of this work. AMB made substantial contributions to the analysis of the data collected in this work. All authors read and approved the final manuscript.

Conflicts of Interest

Two coauthors, AM and IP, are partners in Wellnuts LLC. Wellnuts LLC may commercialize the FFW intervention in the future.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2303 KB - formative_v4i2e15919_app1.pdf]

References

1. Myers ND, Prilleltensky I, Prilleltensky O, McMahon A, Dietz S, Rubenstein CL. Efficacy of the fun for wellness online intervention to promote multidimensional well-being: a randomized controlled trial. Prev Sci 2017 Nov;18(8):984-994. [doi: 10.1007/s11121-017-0779-z] [Medline: 28303422]
2. Myers ND, Dietz S, Prilleltensky I, Prilleltensky O, McMahon A, Rubenstein CL, et al. Efficacy of the fun for wellness online intervention to promote well-being actions: a secondary data analysis. Games Health J 2018 Aug;7(4):225-239. [doi: 10.1089/g4h.2017.0132] [Medline: 29708773]
3. Myers ND, Prilleltensky I, Hill CR, Feltz DL. Well-being self-efficacy and complier average causal effect estimation: a substantive-methodological synergy. Psychol Sport Exerc 2017;30:135-144. [doi: 10.1016/j.psychsport.2017.02.010]
4. Singal AG, Higgins PD, Waljee AK. A primer on effectiveness and efficacy trials. Clin Transl Gastroenterol 2014 Jan 2;5:e45 [FREE Full text] [doi: 10.1038/ctg.2013.13] [Medline: 24384867]
5. Hagger MS, Weed M. DEBATE: Do interventions based on behavioral theory work in the real world? Int J Behav Nutr Phys Act 2019 Apr 25;16(1):36 [FREE Full text] [doi: 10.1186/s12966-019-0795-4] [Medline: 31023328]
6. World Health Organization. 2018 Feb 16. Obesity and Overweight URL: http://www.who.int/mediacentre/factsheets/fs311/en/ [accessed 2018-07-01]
7. Ogden CL, Carroll MD, Fryar CD, Flegal KM. Prevalence of obesity among adults and youth: United States, 2011-2014. NCHS Data Brief 2015 Nov(219):1-8 [FREE Full text] [Medline: 26633046]
8. National Heart, Lung, and Blood Institute (NHLBI). 2013. Managing Overweight and Obesity in Adults: Systematic Evidence Review From the Obesity Expert Panel, 2013 URL: https://www.nhlbi.nih.gov/sites/default/files/media/docs/obesity-evidence-review.pdf [accessed 2018-07-01]

9. Office of Disease Prevention and Health Promotion. 2018. 2018 Physical Activity Guidelines Advisory Committee Scientific Report URL: https://health.gov/paguidelines/second-edition/report/ [accessed 2019-01-18]

10. World Health Organization. Global Recommendations on Physical Activity for Health. Switzerland: World Health Organization; 2010.

11. Tudor-Locke C, Brashear MM, Johnson WD, Katzmarzyk PT. Accelerometer profiles of physical activity and inactivity in normal weight, overweight, and obese US men and women. Int J Behav Nutr Phys Act 2010 Aug 3;7:60 [FREE Full text] [doi: 10.1186/1479-5868-7-60] [Medline: 20682057]

12. Gourlan MJ, Trouilloud DO, Sarrazin PG. Interventions promoting physical activity among obese populations: a meta-analysis considering global effect, long-term maintenance, physical activity indicators and dose characteristics. Obes Rev 2011 Jul;12(7):e633-e645. [doi: 10.1111/j.1467-789X.2011.00874.x] [Medline: 21457183]

13. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Behavioral weight loss interventions to prevent obesity-related morbidity and mortality in adults: US Preventive Services Task Force recommendation statement. J Am Med Assoc 2018 Sep 18;320(11):1163-1171. [doi: 10.1001/jama.2018.13022] [Medline: 30326502]

14. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131 [FREE Full text] [Medline: 3920711]

15. Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, Lancet Physical Activity Series Working Group. The pandemic of physical inactivity: global action for public health. Lancet 2012 Jul 21;380(9838):294-305. [doi: 10.1016/S0140-6736(12)60898-8] [Medline: 22818941]

16. Sallis JF, Bull F, Guthold R, Heath GW, Inoue S, Kelly P, Lancet Physical Activity Series 2 Executive Committee. Progress in physical activity over the Olympic quadrennium. Lancet 2016 Sep 24;388(10051):1325-1336. [doi: 10.1016/S0140-6736(16)30581-5] [Medline: 27475270]

17. Heath GW, Parra DC, Sarmiento OL, Andersen LB, Owen N, Goenka S, Lancet Physical Activity Series Working Group. Evidence-based intervention in physical activity: lessons from around the world. Lancet 2012 Jul 21;380(9838):272-281 [FREE Full text] [doi: 10.1016/S0140-6736(12)60816-2] [Medline: 22818939]

18. Reis RS, Salvo D, Ogilvie D, Lambert EV, Goenka S, Brownson RC, Lancet Physical Activity Series 2 Executive Committee. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. Lancet 2016 Sep 24;388(10051):1337-1348 [FREE Full text] [doi: 10.1016/S0140-6736(16)30728-0] [Medline: 27475273]

19. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJF, Martin BW, Lancet Physical Activity Series Working Group. Correlates of physical activity: why are some people physically active and others not? Lancet 2012 Jul 21;380(9838):258-271. [doi: 10.1016/S0140-6736(12)60735-1] [Medline: 22818938]

20. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? Health Educ Res 2011 Apr;26(2):308-322. [doi: 10.1093/her/cyr005] [Medline: 21321008]

21. Beauchamp MR, Crawford KL, Jackson B. Social cognitive theory and physical activity: mechanisms of behavior change, critique, and legacy. Psychol Sport Exerc 2019;42(5):110-117. [doi: 10.1016/j.psychsport.2018.11.009]

22. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [doi: 10.1016/j.amepre.2007.02.041] [Medline: 17572313]

23. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012 Apr 30;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]

24. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001;52:1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]

25. Bandura A. Self-Efficacy: The Exercise of Control. New York: Freeman; 1997.

26. McAuley E. Self-efficacy and the maintenance of exercise participation in older adults. J Behav Med 1993 Feb;16(1):103-113. [doi: 10.1007/bf00844757] [Medline: 8433355]

27. McAuley E. The role of efficacy cognitions in the prediction of exercise behavior in middle-aged adults. J Behav Med 1992 Feb;15(1):65-88. [doi: 10.1007/bf00848378] [Medline: 1583674]

28. Myers ND, Prilleltensky I, Lee S, Dietz S, Prilleltensky O, McMahon A, et al. Effectiveness of the fun for wellness online behavioral intervention to promote well-being and physical activity: protocol for a randomized controlled trial. BMC Public Health 2019 Jun 13;19(1):737 [FREE Full text] [doi: 10.1186/s12889-019-7089-2] [Medline: 31196052]

29. Feltz DL, Short S, Sullivan P. Self - Efficacy in Sport: Research and Strategies for Working With Athletes, Teams, and Coaches. Champaign: Human Kinetics; 2008.

30. Ainsworth BE, Bassett DR, Strath SJ, Swartz AM, O'Brien WL, Thompson RW, et al. Comparison of three methods for measuring the time spent in physical activity. Med Sci Sports Exerc 2000 Sep;32(9 Suppl):S457-S464. [doi: 10.1097/00005768-200009001-00004] [Medline: 10993415]

31. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]

32. American Psychological Association. Publication Manual of the American Psychological Association. Sixth Edition. Washington, DC: American Psychological Association; 2010.

33. Prilleltensky I, Dietz S, Prilleltensky O, Myers ND, Rubenstein CL, Jin Y, et al. Assessing multidimensional well-being: development and validation of the I COPPE scale. J Commun Psychol 2015;43(2):199-226. [doi: 10.1002/jcop.21674]

34. Myers ND, Park SE, Lefevor GT, Dietz S, Prilleltensky I, Prado GJ. Measuring multidimensional subjective well-being with the I COPPE scale in a Hispanic sample. Meas Phys Educ Exerc Sci 2016;20(4):230-243. [doi: 10.1080/1091367X.2016.1226836]

35. de Vries HJ, Kooiman TJ, van Ittersum MW, van Brussel M, de Groot M. Do activity monitors increase physical activity in adults with overweight or obesity? A systematic review and meta-analysis. Obesity (Silver Spring) 2016 Oct;24(10):2078-2091 [FREE Full text] [doi: 10.1002/oby.21619] [Medline: 27670401]

36. Google Sites. 2005. International Physical Activity Questionnaire URL: https://sites.google.com/site/theipaq/home [accessed 2018-07-01]

37. Myers ND, Feltz DL, Wolfe EW. A confirmatory study of rating scale category effectiveness for the Coaching Efficacy Scale. Res Q Exerc Sport 2008 Sep;79(3):300-311. [doi: 10.1080/02701367.2008.10599493] [Medline: 18816941]

38. Bandura A. Guide for constructing self-efficacy scales. In: Pajares F, Urdan T, editors. Self-Efficacy Beliefs of Adolescents. Charlotte: Information Age Publishing; 2006:307-337.

39. Muthén LK, Muthén BO. Mplus User's Guide. Eighth Edition. Los Angeles: Muthén & Muthén; 2017.

40. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods 2002 Jun;7(2):147-177. [Medline: 12090408]

41. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297-334. [doi: 10.1007/BF02310555]

42. Raykov T, Marcoulides GA. Thanks Coefficient Alpha, We Still Need You!. Educ Psychol Meas 2019 Feb;79(1):200-210. [doi: 10.1177/0013164417725127] [Medline: 30636788]

43. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Second Edition. Hillsdale: Lawrence Erlbaum Associates; 1998.

44. MacKinnon DP. Introduction to Statistical Mediation Analysis. New York: Lawrence Erlbaum Associates; 2008.

45. Lachowicz MJ, Preacher KJ, Kelley K. A novel measure of effect size for mediation analysis. Psychol Methods 2018 Jun;23(2):244-261. [doi: 10.1037/met0000165] [Medline: 29172614]

46. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. Br Med J 1999 Sep 11;319(7211):670-674 [FREE Full text] [doi: 10.1136/bmj.319.7211.670] [Medline: 10480822]

47. Muthén LK, Muthén BO. How to use a Monte Carlo study to decide on sample size and determine power. Struct Equ Modeling 2002;9(4):599-620. [doi: 10.1207/S15328007SEM0904_8]

48. Patel MS, Asch DA, Rosin R, Small DS, Bellamy SL, Heuer J, et al. Framing financial incentives to increase physical activity among overweight and obese adults: a randomized, controlled trial. Ann Intern Med 2016 Mar 15;164(6):385-394 [FREE Full text] [doi: 10.7326/M15-1635] [Medline: 26881417]

49. Salehpour A, Hosseinpanah F, Shidfar F, Vafa M, Razaghi M, Dehghani S, et al. A 12-week double-blind randomized clinical trial of vitamin D   supplementation on body fat mass in healthy overweight and obese women. Nutr J 2012 Sep 22;11:78 [FREE Full text] [doi: 10.1186/1475-2891-11-78] [Medline: 22998754]

50. Tehard B, Saris WH, Astrup A, Martinez JA, Taylor MA, Barbe P, et al. Comparison of two physical activity questionnaires in obese subjects: the NUGENOB study. Med Sci Sports Exerc 2005 Sep;37(9):1535-1541. [doi: 10.1249/01.mss.0000177464.68521.3b] [Medline: 16177606]

51. Collins LM, Kugler K. Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST). New York: Springer; 2018.

52. The Community Guide. Physical Activity: Interventions Including Activity Monitors for Adults with Overweight or Obesity URL: https://www.thecommunityguide.org/findings/physical-activity-interventions-including-activity-monitors-adults-overweight-obesity [accessed 2018-07-01]

53. Myers ND, Lee S, Bateman AG, Prilleltensky I, Clevenger KA, Pfeiffer KA, et al. Accelerometer-based assessment of physical activity within the Fun For Wellness online behavioral intervention: protocol for a feasibility study. Pilot Feasibility Stud 2019;5:73 [FREE Full text] [doi: 10.1186/s40814-019-0455-0] [Medline: 31164990]

54. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S531-S543. [doi: 10.1249/01.mss.0000185657.86065.98] [Medline: 16294116]

55. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]

56. Lee I, Shiroma EJ. Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. Br J Sports Med 2014 Feb;48(3):197-201 [FREE Full text] [doi: 10.1136/bjsports-2013-093154] [Medline: 24297837]

57. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014 Jul;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]

58. Wijndaele K, Westgate K, Stephens SK, Blair SN, Bull FC, Chastin SF, et al. Utilization and harmonization of adult accelerometry data: review and expert consensus. Med Sci Sports Exerc 2015 Oct;47(10):2129-2139 [FREE Full text] [doi: 10.1249/MSS.000000000000661] [Medline: 25785929]

59. Tully MA, Panter J, Ogilvie D. Individual characteristics associated with mismatches between self-reported and accelerometer-measured physical activity. PLoS One 2014;9(6):e99636 [FREE Full text] [doi: 10.1371/journal.pone.0099636] [Medline: 24919185]

Abbreviations

BARSE: barriers self-efficacy
BET I CAN: behaviors, emotions, thoughts, interactions, context, awareness, and next steps
EXSE: exercise self-efficacy
FFW: Fun for Wellness
IPAQ: International Physical Activity Questionnaire
IRB: institutional review board
ML: maximum-likelihood
T1: baseline
T2: 30 days after baseline
T3: 60 days after baseline
UC: usual care
VPS: virtual private server
WHO: World Health Organization

XSL•FO
RenderX

<u>Original Paper</u>

Experiences and Needs of Multicultural Youth and Their Mentors, and Implications for Digital Mentoring Platforms: Qualitative Exploratory Study

Rebecca Lynn Radlick[1], PhD; Jelena Mirkovic[2], PhD; Sarah Przedpelska[3], MA; Elanor Halvorsen Brendmo[2], MA; Deede Gammon[2,4], PhD

[1]Norwegian Research Centre (NORCE), Bergen, Norway

[2]Center for Shared Decision-Making and Collaborative Care Research, Oslo University Hospital, Oslo, Norway

[3]Catalysts Association, Oslo, Norway

[4]Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway

Corresponding Author:
Rebecca Lynn Radlick, PhD
Norwegian Research Centre (NORCE)
Nygårdsgaten 112
Bergen, 5008
Norway
Phone: 47 56107634
Email: rera@norceresearch.no

Abstract

Background: Mentoring programs (ie, programs that connect youths with adult volunteers) have been shown to improve outcomes across the behavioral, social, and academic domains of youth development. As in other European countries, mentoring programs have few traditions in Norway, where interventions for multicultural youths are usually profession driven and publicly funded. Faced with the risk of disparities in education and health, there is a need to better understand this group's experiences and requirements relative to mentoring. This would also serve as a basis for designing and implementing digital support.

Objective: The objective of this study was to gain insight into multicultural youth mentees' and adult mentors' experiences and needs in the context of an ongoing mentoring program, how digital support (electronic mentoring) might address these needs, and how such support could be designed and implemented.

Methods: The study used a qualitative approach, with data from 28 respondents (21 mentees and 7 mentors). In total, 4 workshops with mentees as well as semistructured interviews with mentees and mentors were conducted. The sessions were audio recorded, transcribed, and analyzed thematically.

Results: In total, 3 main themes were identified from the experiences and needs reported by the mentees and mentors. These included a need for connection, help in achieving goals, and the need for security and control. Subthemes encompassed a desire to socialize with others, balancing the nature of the relationship, paying it forward, building trust, sharing insights and information with peers, goal-oriented mentees and mentors wanting to assist with goal achievement, and the fundamental need for privacy and anonymity in the digital platform.

Conclusions: The findings of this study are supported by the literature on traditional mentoring, while also offering suggestions for the design of digital solutions to supplement the in-person mentoring of multicultural youth. Suggestions include digital support for managing the mentee-mentor relationships, fostering social capital, and ways of ensuring security and control. Features of existing electronic health apps can be readily adapted to a mentoring program context, potentially boosting the reach and benefits of mentoring.

XSL•FO
RenderX

Introduction

Background

School dropout among adolescents and young adults has been increasingly reframed as a public health issue in light of the strong association between poor self-rated health in adolescence, high school dropout, and reduced labor market integration [1-3]. Immigrant youths in Norway are at a greater risk for unemployment or leaving school early and exhibit a 26-percentage point disparity in education and employment, compared with native Norwegian youths [4]. Factors contributing to these disparities within the immigrant youth population include time spent in Norway and, thus, language abilities, reason for migration (refugee vs labor migration), health status, and parental income [4].

In light of the aforementioned discrepancy in school and labor market participation across immigrant and majority youth, there is an increased willingness to try out alternative models, such as mentoring programs, to reduce these disparities. Mainly studied in the United States, mentoring is defined as "taking place between young persons (ie, mentees) and older or more experienced volunteers (ie, mentors) who are acting in a nonprofessional helping capacity to provide relationship-based support that benefits one or more areas of the mentee's development" [5]. A 2001 meta-analysis of 73 independent evaluations of mentoring programs supports the effectiveness of mentoring in improving outcomes across behavioral, social, and academic domains [6], as does a more recent meta-analysis of outcome studies [7]. However, mentoring programs, which are typically run as social entrepreneurships, have few traditions in Norway, where interventions for immigrant youths are often profession driven and publicly funded [8].

One of the first mentoring programs in Norway (called Catalysts) targets recently arrived immigrant youths aged between 16 and 25 years. The majority are recruited through introductory language classes at their schools, which are mandatory for everyone who wishes to complete high school in Norway. Participants' immigration background varies and includes unaccompanied minor refugees, family reunification, and children of labor migrants. The Catalysts mentoring program lasts for 6 months and matches immigrant youths (mentees) and volunteer adults (mentors) in mentor-mentee dyads according to their interests and needs. Program components are closely aligned with mentoring best practices [5] and build on the principles of appreciative inquiry [9,10], closely related to positive psychology [11]. The Catalysts program is primarily commissioned by municipalities and corporate social responsibility entities in businesses, in addition to funding from the Norwegian Labor and Welfare Administration (NAV).

Electronic Health and Adaptations to a Mentoring Context

The researchers who initiated collaboration with Catalysts have backgrounds in electronic health (eHealth) and sought ways of applying eHealth knowledge and apps to health promotion interventions outside of health care settings for health promotional purposes. Catalysts mentoring practitioners, for

their part, were interested in evidence-informed digital innovations to improve the reach and effectiveness of their programs. We joined forces to illuminate the following overarching question: to what extent could an evidence-informed electronic mental health platform (called ReConnect, see also the Methods section) be adapted to the needs and experiences of mentees and mentors to enhance mentoring? As described elsewhere [12,13], the original ReConnect platform was designed for individuals requiring long-term mental health care. It included 3 main components: a peer support forum, a secure messaging function between service users and their health care providers, and a toolbox of resources that could be used autonomously or in collaboration between service users and providers (with resources related to mapping strengths, mindfulness, sleep hygiene, goal-setting modules, personal network, and medications). The design and operation of ReConnect were guided by principles of recovery [14], which may or may not resonate with stakeholders of youth mentoring.

Little prior research was available, as only 3% of mentoring programs in the United States have a digital component and only 1% are exclusively digital [15]. The few implementations of digital components previously studied range from informal and supplemental to more formal or exclusive (digital interactions only) and include email, social media, and SMS as well as app-mediated connections and computer platforms [16]. Research indicates that demographic and personal circumstances, communication styles, accessibility issues, and program implementation shape the effectiveness of electronic mentoring (e-mentoring) among youth, but many questions remain about what types of digital solutions may be effective and for whom [16].

Theory of Change

To assess the relevance of the pre-existing eHealth platform to the context of mentoring and to identify necessary and desired adaptations, we encouraged Catalysts to articulate their program's *theory of change*. Along with many in the mentoring field [17], Catalysts maintains that in-person relationships between mentees and mentors are foundational and that any digitalization should aim to enhance, not replace, in-person relationships. In addition to appreciative inquiry [10], which explicitly redirects attention away from problems and vulnerabilities toward strengths and opportunities, Catalysts' theory of change encompassed elements closely aligned with social capital [18-22], particularly the *bridging* type of social capital that is found across the lines of age, social status, and ethnicity [23]. Evidence suggests that having wider support networks of peers and unrelated adults across a range of domains (that is to say, greater social capital) is associated with a variety of positive outcomes, including better youth mental health [24,25]. These perspectives might broadly be seen as complementary to the recovery-oriented perspectives and literature that guided the design of the original ReConnect platform [12]. Nevertheless, the efforts to arrive at clearer (theoretical) rationales for how to adapt ReConnect and why (or instead start from scratch) called for the perspectives of mentees and mentors. To elicit these perspectives, an open inductive approach to e-mentoring was considered most appropriate at this stage.

Research Questions

This exploratory study aimed to gain insights into the experiences and needs of immigrant youths and their mentors and how the mentoring experience might be enhanced by a digital supplement. More specifically, the following questions guided the study:

> *What are the needs of mentoring stakeholders (mentors and mentees)?*
>
> *How can these needs be addressed using a digital platform?*
>
> *How can such a platform be designed and implemented within this mentoring context?*

Insights into these issues will be used as a basis for subsequent formulations of user requirements of an e-mentoring platform, and thereafter an intervention study, to assess the effects of the platform.

Methods

Study Design

The study was conducted between winter 2018 and spring 2019 and was approved by the Data Protection Officer at Oslo University Hospital. Qualitative data were collected with different methods to illuminate the research questions, in close collaboration with stakeholders (mentees and mentors). As outlined in Table 1, we conducted interviews, focus groups, and workshops at the Center for Shared Decision Making and Collaborative Care Research (Oslo University Hospital), Norwegian Research Center (NORCE), and Catalysts localities to gain insights from stakeholders about their needs and preferences for digital support to augment their face-to-face mentoring meetings.

Participant Recruitment

Mentors and mentees were recruited from among current or previous participants in the Catalysts program in several ways: an open call for participation among active mentors and mentees or by direct contact from the program coordinator. In the latter case, a purposive sampling of mentors and mentees was used to select participants with varied backgrounds in terms of program progression and demographic characteristics [26]. Each person from this list of mentees was then contacted (by EB or SP), and interviews or workshops were scheduled. All participants received information about the purpose of the study, the voluntary nature of participating, and maintenance of confidentiality, and all participants signed informed consent forms. The interviews and focus group workshops were audio taped and transcribed, with the exception of the pair interview, which was not taped because of technical problems. There were 28 different respondents, including 21 mentees and 7 mentors (Table 1).

Table 1. Data collection method and respondent roles.

Data collection approach	Respondents' role (number of respondents)
Interview with mentees	Mentee (n=2)
Interview with mentor	Mentor (n=6)
Interview with mentor and mentee pair	Mentor (n=1) and mentee (n=1)
Focus group workshops (n=4) with mentees	Mentees (n=23; 18 unique); 5 participated in multiple workshops

Data Collection

Participants were interviewed in 8 individual interviews (2 mentees and 6 mentors) and 1 pair interview (1 mentor and 1 mentee), as illustrated in Table 1. Mentees were interviewed in person, whereas mentors were interviewed both in person and via telephone. The individual interviews were conducted by 2 of the authors (EB and RLR) and lasted between 30 min and 1 hour. A total of 4 focus group workshops were also conducted, with 18 different mentees participating; an additional 5 mentees participated in multiple workshops. Here, an author (SP) acted as a moderator and guided the activities and conversations.

Mentees were interviewed at different phases of the program, allowing us to gain insights into the youths' expectations, experiences, and needs as they progressed in the program. All mentors were interviewed toward the end of the program or after the programs had concluded. A semistructured interview guide was used, and the initial questions posed to respondents were intentionally broad, addressing expectations and motivations for participation; challenges and benefits of the program; and, specifically, the use of apps. Over time, as the requirements for the design of an e-mentoring platform began to take shape, questions became more specific, for example, "If an app were to be introduced as part of the program, what would be useful to you to have in such an app, and why?" During the discussions, participants were asked follow-up questions to expand on what was said and to reflect upon their answers if needed [26]. Respondents were also encouraged to share their concrete experiences in participating in the mentoring program, elaborating on how digital support might strengthen or undermine specific aspects of the program. Workshops allowed interactions among participants, who could mutually confirm, reinforce, and contradict each other's statements, whereas the interviews allowed deeper probing around the functionality of the mentor program and how it could be improved with a digital innovation. In the workshops and interviews, the original ReConnect platform (described earlier), which included 3 interactive components (ie, peer support forums for users and health care providers, a toolbox of resources, and secure messaging), was used as a point of departure. Features and design principles from ReConnect were mentioned when posing questions about its relevance to participants' needs and preferences in a mentoring context. Given the ReConnect platform's medical content, archaic looks, and code libraries, the study group decided to use the features and design principles in developing a new, but similar, platform. In the workshops,

participants engaged with hand-drawn wireframes based on some of the features from ReConnect, such as the forum, messaging, and toolbox. Stakeholders were asked how they would engage with the aforementioned features in a mentoring context, to assess their needs. This selection of features from the ReConnect platform potentially influenced respondents' feedback on specific possibilities for an e-mentoring platform.

Data Analysis

Data were analyzed using thematic analysis, where the researchers first familiarize themselves with the data; generate initial codes; assess, review, and name themes; and then report findings [27]. The recordings from the interviews and workshops were transcribed and imported to NVivo (QSR International Pvt Ltd) qualitative data analysis software. In the first step of the analysis, 3 authors (DG, EB, and SP) independently read the transcriptions of the interviews and workshops to get an overview of topics reported by the participants and to generate initial codes from the transcripts. The code generation and analysis were completed with the initial research questions in mind, focusing on the needs of the respondents, how an e-mentoring platform could address such requirements, and how to design and implement the platform. All manifest content, or descriptions of needs, experiences, and ideas that were considered relevant to creating, adapting, and implementing the e-mentoring platform, were extracted from the transcriptions. The initial codes were generated, compared, and discussed between the 3 authors. The fourth author (JM), who is a computer design scientist, observed several of the workshops and participated in discussions of emerging codes in light of

design implications. Overlapping or similar meaning codes were grouped together into initial themes. The themes were reviewed and discussed iteratively by 4 of the authors (DG, EB, SP, and JM), with the aim to reduce the number of themes and define meaningful and relevant names of the themes without losing relevant meaning. After defining codes and themes, the fifth author (RLR) then independently read and coded the material, comparing with already extracted codes. All authors discussed the results to reach consensus on final names of the themes and codes for subthemes. The Norwegian quotations were translated by RLR and DG, who are both native English speakers and fluent in Norwegian.

Results

Overview

A total of 3 overarching themes were identified in the respondents' descriptions of the needs and preferences for traditional mentoring. The main themes identified were related to connection, personal goals, and security and control. Most subthemes were common to both mentees and mentors, but several subthemes were more closely related to a single respondent group. Each of the overarching themes and subthemes are first presented, followed by a summary of suggestions from the informants about how digital support might address such needs. Quotations that are particularly illustrative of the theme are included in the text, whereas supplementary quotations are referenced and are available in Multimedia Appendix 1. A summary visualization of the themes and subthemes is depicted in Figure 1.

Figure 1. The main themes and subthemes.
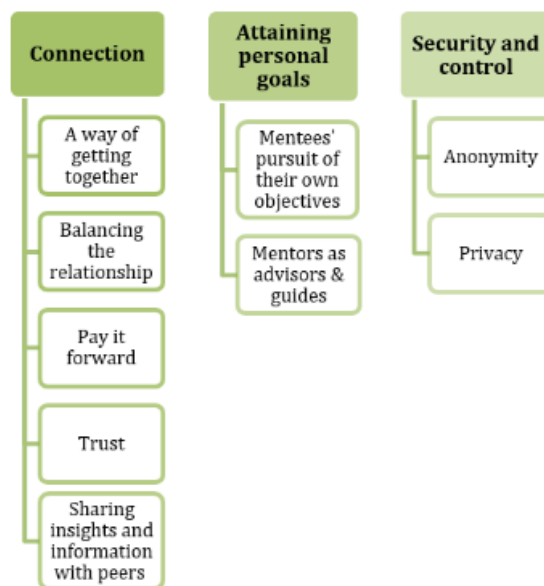
Theme 1: Connection—A Sense of Community

The first overarching theme identified in the data was the stakeholders' desire for connection or belonging to a larger community. This concept was reflected in mentor-mentee relationships, connections within the separate mentee and mentor spheres, and links to Norwegian society. A mentee described challenges with regard to the latter point:

You know Norwegians, if you don't speak to them first, they won't speak with you. [Mentee, Interview 2]

Subthemes extracted from the transcripts included a desire to get together with others, balancing the relationship, paying it forward, trust, and the need for sharing insights and information with peers.

A Way of Getting Together

The first subtheme encompassed the need for group activities and ways of socializing. A core argument for introducing a digital platform as part of the program was to enhance contact between the mentee-mentor dyads. Many of the mentees expressed a desire for frequent contact with their mentors, as illustrated by the following mentee quotation (when asked how often she would like to meet her mentor):

> If possible five days a week! Every day! [Mentee, Interview 1]

The youths emphasized that an app should not be used at the expense of human contact (meetings with mentors) but rather supplement it. Mentees already felt that the in-person contact was too infrequent (Multimedia Appendix 1) and that emotional bonding, showing empathy, and closeness could be challenging to facilitate remotely and would be preferable in person (Multimedia Appendix 1). However, the same mentee acknowledged that context was important, and for practical urgent needs, such as feedback on a job application, face-to-face contact was less essential (Multimedia Appendix 1).

In addition, multiple mentees expressed a desire to find events and arrangements locally and saw the program as a way to do so. For example, although one of the mentees wanted to be more socially active, before participating in the program, she usually just went home and stayed alone (Multimedia Appendix 1). The need to meet other people and participate in activities was also apparent in this quote from one of the youths describing his town:

> I hear those who live in [a city], they have lots of activities and stuff...And then, there's [a rural place] where I live. I see more trees than people [laughs]. And in the wintertime there's no people. Only sometimes I meet an old lady and an old man, walking their dog. [Mentee, Workshop 1]

For the adults, the social aspect of meeting other mentors with different backgrounds was mentioned as a motivation for volunteering. Some reported having expanded their own networks of contacts by meeting other mentors (Multimedia Appendix 1). Similar to the youth, mentors also emphasized the importance of in-person contact with their mentees (Multimedia Appendix 1).

Balancing the Nature of the Relationship

A second subtheme relating to the need for connection addressed the nature of the mentor-mentee relationship and balancing relationship roles. For example, several of the mentors experienced blurred boundaries between being a mentor (formal role) and being a friend (informal role). For some mentors, the blurring of the roles was quite natural (Multimedia Appendix 1) and could develop organically (Multimedia Appendix 1), whereas for others, it was a conscious choice:

> I have seen this as more of a friend relation...so I have chosen to not be so formal because...it's another person. This isn't some [social work case in the public system]. [Mentor, Interview 4]

This more informal relationship was echoed by a mentor who underscored the spontaneity of interactions with their mentee (Multimedia Appendix 1). Although many of the mentors described permeable boundaries between the formal and informal, they also expressed concerns about *intruding* on their mentees' personal spheres when using social media for communicating.

Pay it Forward

Both mentees and mentors articulated a need to share their knowledge and experiences—to *pay it forward*. The youths felt that they benefited from having a mentor, and several of them described wanting to share their experiences and knowledge with their fellow mentees, future mentees, or other youths who were not fortunate enough to have a mentor (Multimedia Appendix 1).

Mentors also mentioned the desire to pay it forward as one reason for volunteering in the program. For example, some had mentors of their own at work or had prestigious jobs, and some felt a responsibility for sharing their knowledge and experience, in helping with integration in Norway (Multimedia Appendix 1) or in trying to help as many people as possible (Multimedia Appendix 1). Other mentors felt motivated by the possibility to share their own experiences of being new to a company, city, or country or by having been in similar circumstances as the mentee. These mentors recalled what it felt like to be young and uncertain of what direction to take in life or how to get there, and they felt a strong desire to help (Multimedia Appendix 1). Mentors also recognized the effect the relationship with mentees had on their own personal development and that they learned a lot not only about other people but also about themselves (Multimedia Appendix 1). However, this expectation of *paying it forward* sometimes conflicted with reality, as experienced by a mentor whose mentee had a broad network and significant resources before joining the program:

> I learned a lot about the daily life of youths in his situation, and the situation he has been in. I would have liked to have helped him more, but in a way, that is "egotistical altruism" or whatever you want to call it. That you wish you could do more. But it is of course only a good thing that he didn't have a need for more. [Mentor, Interview 5]

Relational Trust: A Supportive and Safe Environment

The mentees highlighted trust as important, but challenging, in their relationship with mentors. This was partially attributed to the (perceived) brevity of the program (6 months). Talking about personal issues requires trust; however, trust takes time to develop. However, the requirement for mentors to sign confidentiality agreements ameliorated this issue for some mentees (Multimedia Appendix 1). There was also a discussion on whether it would be preferable to meet face-to-face first or via an app. Several mentees suggested that it would be preferable to meet in person first before engaging in any kind of digital interaction, as they only interact with people online when they actually know the person and have built trust with them (Multimedia Appendix 1). However, others opined that using an app first could be useful for getting to know each other before

the first meeting. The fact that only authorized users could have access to the platform (as discussed in the section below on security and control) could also help enhance the feeling that it is a safe environment to share personal information. Trust was also discussed by the mentors as fundamental in developing the relationship with their mentee (Multimedia Appendix 1).

Sharing Insights and Information With Peers

Both mentors and mentees expressed a need for information and insights in navigating their relationships, particularly through discussion with others in the same role. Several of the mentees wanted to talk with other youths, without mentors being present:

> Because this was the first time I had a mentor in my life. Maybe it's the same for the others. So it can be a bit challenging to navigate, at least the first time. So we share tips between us, and meet up, and similar. [Mentee, Workshop 1]

The adults were also new to mentoring and wanted more information on best practices and specific scenarios that they might encounter. A mentor discussed how she needed to give her mentee information on how to apply for study grants, as well as about citizenship requirements, but that this information was not readily available and was time-consuming to find (Multimedia Appendix 1). Another mentor mentioned that the communication between the group of mentors at meetings was quite good, but that there was no communication outside of the meetings, and that supplementary contact was desirable (Multimedia Appendix 1). Some mentors expressed uncertainty as to what they should do at the in-person meetings with their mentees and wanted additional guidelines, potentially to ensure that they were maintaining standards of best practice. However, although many of the respondents shared a desire for peer-to-peer support, some of the mentors and mentees did not find support and contact with other adults or youths, respectively, as something necessary.

Theme 2: Attaining Personal Goals

A second overarching theme emerging from the data was related to goals and expectations and was reflected in both mentor and mentee responses. The mentees were motivated and had clear ideas of what they needed support with, whereas the mentors frequently viewed themselves as guides, encouraging and helping their mentees to achieve these goals.

Mentees' Pursuit of Their Own Objectives

The youths overwhelmingly viewed the program as a way to help them in attaining their goals, and many sought support for broad needs related to education and career (Multimedia Appendix 1). Mentees were sometimes very specific about their needs for assistance with homework and school; learning better computer skills; or improving their Norwegian, both formal and colloquial:

> I would really like advice on what's best for me. For example, what I'll do in the future. What career I should choose, or how I can be integrated into Norwegian society, and better understand Norwegian culture. I'm going to live in Norway, right? ...I really

> want to learn Norwegian [slang], like [young people speak] on the streets. [Mentee, Workshop 1]

A number of the mentees mentioned strengthening or broadening their networks as another key motivation for participating in the Catalysts program, and this opportunity to meet a supportive nonrelative adult was appealing (Multimedia Appendix 1).

Mentors as Advisors and Guides

Mentors reported seeing themselves as guides, advisors, or *coaches*. Many mentors viewed one aspect of their work as supporting their mentee in achieving specific objectives, such as finding a job, as well as motivating and following up on the mentee's progress (Multimedia Appendix 1).

One mentor helped her mentee get a part-time job at a customer service center after the mentee expressed an interest in a job where she could help other people (Multimedia Appendix 1). At the same time, some mentors expressed uncertainty about their roles when faced with highly motivated mentees who had already identified career objectives that they appeared well equipped to pursue. Sometimes these mentees were able to ask for and receive help from others in their existing network. This led to some frustration from one of the mentors who clearly had a strong desire to help with specific tasks:

> I think it was a great experience [mentoring], but also a little frustrating. I felt that my mentee didn't really have much of a need for a mentor. We struggled with finding things to work with, since he had connections with other adults that he was in contact with regularly, who he could ask the same questions to...He had a better, more well-established network around him rather than just him and me... [Mentor, Interview 5]

Theme 3: Security and Control

The respondents expressed a range of needs that had to do with security and control over their personal information. A main reference point was social media (such as Facebook and Instagram), and many of the mentees experienced these platforms as being unsafe (Multimedia Appendix 1). *Anonymity* was a subtheme identified in the data, although both mentors and mentees were not in complete agreement on whether anonymity was desirable or not. The second subtheme was a concern about *privacy*. The need for security and control was discussed in light of particular e-mentoring platform features (as presented in the subsequent section): a forum, messaging, and toolbox.

Anonymity

Some mentees wanted to remain anonymous in the forum all or some of the time, whereas others were uninterested in the option; their reasoning was often context dependent. Several youths wanted very personal discussions related to mental health to be private, whereas less sensitive topics such as choosing school courses or dealing with parents were acceptable for nonanonymous discussion with mentors and other mentees. Those supporting anonymity described how they might want to share sensitive issues without concerns of being identified. These respondents felt that it could be useful to present problems

anonymously and to discuss them, both with other mentees, in case of having misunderstandings with the mentor, and with the mentor later, if it was a personal issue (Multimedia Appendix 1). Those opposed to the option of anonymity expressed concerns that it might undermine the networking and connective potentials of the forum as well as present a hindrance in building trust between participants (also discussed in the first section on *connection*; Multimedia Appendix 1). One mentee suggested that it is quite difficult to trust, build connections, and discuss personal topics if some people are anonymous, whereas others are not:

> But [being] anonymous, it is [a] little bit...maybe someone will not talk with you. They will think, ok, she is anonymous...I don't know [who she is]...and...don't want to put her information. So, with who I'm talking. It is, like, difficult to trust. [Mentee, Workshop 3]

Mentors expressed concerns for respecting their mentees' private spheres and discomfort in mixing private social media usage with mentoring. This also relates closely to distinguishing between more formal and informal roles (as discussed in the section on balancing the nature of the relationship).

Privacy

The second subtheme, *privacy*, relates to who has access to personal information. Mentors and mentees expressed needs for limiting unauthorized access and being able to trust that people are who they say they are in the platform. Ensuring that mentee data would not be leaked or that unauthorized people would not gain access was described as being critical. Similarly, the mentors articulated strong concerns about their mentees' privacy and maintaining confidentiality. Privacy was mentioned by mentees relative to family members who could sometimes be intrusive; this becomes more poignant when one considers the potential for social control. As a female mentee expressed:

> I have a brother, and he said to me, "you aren't allowed to talk with [this] boy"...my father knows...but [my brother] thinks this way because he doesn't like that I talk with boys or hang out with boys. [Mentee, Interview 2]

Cultural differences between countries of origin and Norway with regard to privacy were mentioned by another mentee who appreciated high-level security in logging on and how it prevented others from gaining access to private information (Multimedia Appendix 1).

Digital Support for Identified Needs

The youths and mentors mentioned numerous experiences and needs, as outlined above, and suggested multiple ways that digital support could help meet these needs. Throughout these discussions, various elements from the ReConnect platform [12] were presented as examples of what might be possible in the e-mentoring platform. These included a *forum*, *messaging function*, and *toolbox*. The youths also proposed additional potential features in the e-mentoring platform including GPS and an activity *calendar*.

Forum

Both mentees and mentors identified multiple purposes for *forums*. Forums were suggested as a way for mentees to pay it forward, sharing knowledge they had gained through the mentoring experience with other youths in the program, with the next group, or with youths not in the program (Multimedia Appendix 1). Initiating and discussing issues anonymously was another use for a forum, as seen by the mentees. A group mentor-mentee forum could allow connections between mentors and mentees with similar interests or mentees who needed advice on a particular topic or career, allowing better maximization of mentor resources, as suggested by both youths and mentors (Multimedia Appendix 1).

A mentor-only forum was mentioned by the mentors as potentially providing a place where they could initiate conversations with other mentors, becoming more secure in their role, and increasing feelings of connection to the mentor group as a whole. Being able to ask for advice and best practices in a forum was mentioned by many of the mentors (Multimedia Appendix 1). However, one mentor felt that concrete and simple issues could be discussed in an electronic forum but questioned the extent to which complex issues could be discussed this way without losing important context (Multimedia Appendix 1).

In the discussions on security and control, the desire for *anonymity* in the e-mentoring platform focused on the potential to allow mentors access to posts in the *forum* and whether forums should be separate for mentees and mentors. The mentors also expressed concerns related to privacy and anonymity of their mentees in the forum. One of the mentors was worried that even in cases where one could post anonymously, the information might be able to be associated with a specific mentee anyway (Multimedia Appendix 1). Anonymity could, thus, lead to a false sense of security and create uncertainty among the users.

Messaging

Mentees mentioned potential benefits of sending and receiving documents and similar items (such as a curriculum vitae [CV]) in private messages to mentors. One such benefit was enhancing accountability for doing what was agreed upon, as this gave mentors a way to follow-up (Multimedia Appendix 1). However, there was some debate among the mentees about what information they would want to share with their mentors (as discussed in the theme of security and control). Some of the mentees indicated that they were more comfortable with discussing their own strengths (a component of the nondigitized program) and receiving feedback from mentors online rather than face to face. It was suggested that a digital component could provide a place to initiate discussions of sensitive topics with mentors, building trust by opening the doors for more personal in-person communication. For example, one mentee said she wanted to ask her mentor for advice about sensitive situations, such as how to deal with a fight with a friend and how to resolve problems with her teacher who expressed anti-immigrant sentiments (Multimedia Appendix 1).

Mentors felt that a chat or forum function could be useful in supplementing the face-to-face program by providing a specific

arena where they could follow up on tasks and objectives that they and their mentees had agreed upon. One mentor communicated on a weekly basis with his mentee via social media and telephone and felt it benefited the relationship and made it easier to follow up on various goal-oriented tasks (such as writing a job application; Multimedia Appendix 1). This type of follow up could be done via messaging. This mentor noted how his mentee appreciated the opportunity to practice Norwegian language with the mentor on the phone or Facebook and suggested that a messaging function would be similarly useful (Multimedia Appendix 1).

As several mentors expressed concerns about *intruding* in their mentees' private spaces or social media personas, a dedicated channel for communication (program platform) might be preferable to using social media traditionally used with peers, as suggested by one mentee (Multimedia Appendix 1). This was seen as particularly relevant in cases where mentors experienced challenges in differentiating between mentor and friend roles (Multimedia Appendix 1).

Many of the mentors and mentees were optimistic about the potential for communicating via a messaging function, whereas others did not view this aspect as positively. Mentors expressed concerns about being available at all times for mentee messages and the potential burden this might represent. This specifically concerned messages that might indicate that the mentee was depressed, feeling lonely, or missing their family (Multimedia Appendix 1). Several other mentors felt that they already had sufficient tools for communication (eg, SMS, WhatsApp, and phone) and did not require a new platform for communication with their mentees (Multimedia Appendix 1).

An important issue, which was evident in the discussions with mentors and mentees on the forum and messaging function, was the need for security and control. A suggestion for ensuring privacy (security of data) was to use a stringent log-on (at the same cryptographic level as banking), which almost everyone in Norway has, for access. This would ensure that only people who were part of the program could log on and that family members would not be able to read any personal information (Multimedia Appendix 1).

Toolbox

The *toolbox* concept in the original ReConnect platform referred to a wide range of optional interactive support tools and resources. The contents of a toolbox tailored to mentoring activated a variety of ideas among the respondents, ranging from one-way information snippets about the mentoring program to more interactive and individualized support related to the program components.

One activity in the Catalysts program is identifying the youths' personal strengths, but sometimes it was challenging for mentees to identify and discuss these. It was suggested that a toolbox could support this process by including concrete topics for discussion or allowing mentees to select from predefined categories (Multimedia Appendix 1). Other suggestions for potential components for a toolbox included guidance for getting a job, writing a CV, strengthening Norwegian skills, and similar, thus facilitating goal achievement and connection to Norway.

The ability to follow progression in the program generally or progress toward personal goals was also a desirable function (Multimedia Appendix 1). For some of the mentees, mentors were seen as possessing a special kind of knowledge that they wanted to gain access to, both generally and more specifically related to school, the job market, or language. Mentors, in turn, expressed the desire to share their knowledge about their education and career *paths* with mentees or simply assist them in getting work. The toolbox was discussed explicitly as a *super solution* in this context (Multimedia Appendix 1).

Despite receiving information and guidance during the monthly meetings, mentors frequently mentioned needing additional tools during the course of the program. One mentor revealed that he had not used some of the basic components of the program that they learned in training because his mentee did not find them relevant. However, this mentor and others felt that a toolbox could help them concretely understand what tools could be used in practice and, possibly, send them reminders to do so (Multimedia Appendix 1).

By having the platform on a mobile phone, mentors and mentees could also make specific plans for each meeting, addressing the need for information and uncertainty experienced by both parties and providing a knowledge basis with which to *pay it forward*:

> Yes, and the toolbox, I was thinking...when the mentor and mentee meet. The first meeting is a bit awkward. So I thought that if the other mentees could write about what they would do with their mentors, then the others could see. So the mentees can learn from...other mentees. [Mentee, Workshop 2]

Several mentors viewed the toolbox as a potential *information bank* with practical examples, information on best practices, text or links to various resources, and suggestions on what mentors could do if a challenging situation arose (Multimedia Appendix 1). Such examples could help the mentors to feel more comfortable in their roles and better equipped to deal with uncertainties (Multimedia Appendix 1) as well as better able to help mentees achieve their goals.

Additional Components

Through the course of the discussions, additional components for the e-mentoring platform were suggested, such as *calendar* and *GPS* or *position tracking*. Introducing a location-specific calendar could provide an oversight over both program milestones, reminders of program meetings, and activities and events happening in the local community, helping to activate participants and increase the youths' sense of connection (Multimedia Appendix 1). A shared calendar could also allow mentees to connect with others participating in the program, and it was suggested that mentees could bring along their friends, allowing a *pay it forward* aspect for peers outside of the program. Another proposal was to allow the calendar to be connected to external calendars, so that mentors could check their mentees' availability (Multimedia Appendix 1). However, some youths found it unnecessary to have a calendar on the digital platform and felt that their own personal calendars were sufficient. Several of the mentees also expressed a desire for tools, such as a GPS, that would help them geographically

navigate the local community, especially for meetings in the mentor program or for events (Multimedia Appendix 1). The idea of having a shared *calendar* function was also discussed in light of the need for security and control, with mixed opinions. Some of the mentees wanted the calendar to be private, whereas others felt that they should be open with their mentors (Multimedia Appendix 1).

Discussion

Principal Findings

The original pragmatic question posed in this study was about the relevance of an eHealth platform (ReConnect) in the context of a mentoring program (Catalysts). After identifying the key elements of Catalysts' theory of change as a basis for assessing this question (see the Introduction section), we posed open questions to immigrant youths and mentors about their experiences and needs related to mentoring and potential areas for digital support. The findings are first discussed in light of the literature, irrespective of digital support, followed by implications for digital design and implementation in a mentoring context.

Mentees and Mentors

Broadly, the experiences and needs expressed by the multicultural youths in this study coincide well with studies that have examined similar groups relative to the concepts of health and social capital. For example, a Canadian study [28] found that refugee youths defined *health* in terms of a sense of belonging, an ability to cope, and self-determination, dimensions closely aligned with the themes identified in this study.

Although the resultant themes were distinct, they, nevertheless, overlapped in certain areas. For example, *security and control* was primarily discussed in the context of digital support issues (protection of personal data), but it tied in with the broader issues of trust and belonging—trust in who the youths communicated with and being part of a community that was safe and supportive. Similarly, the mentees' need for connection was also reflected in the *goals* theme in that youths sought broader connections to the Norwegian community through their mentors as a means of attaining their goals. These findings are reflective of what others have found [29].

Mentors' experiences and needs largely reflected their desire to contribute to society by supporting immigrant youths in their development and integration. Mentors saw their role, in part, as sharing their knowledge of how society and its institutions such as school and work functioned. Mentors also expressed a desire for connection and to understand other cultures, as well as enhance their own networks via contact with fellow volunteers. Their reported motivations for becoming a mentor and the insecurities experienced in their roles (eg, how to blend formal [mentor] and informal [friendship] roles) are similar to what others have found [30-34].

Implications for Electronic Mentoring Design

Mentees and mentors offered a wide range of ideas and preferences for adapting the main components from the original ReConnect platform (eg, forum, messaging, toolbox) and suggested new components for the e-mentoring platform (such as calendar and GPS). Perhaps the clearest implication of the findings for e-mentoring design, expressed by all informants, is that it must not replace in-person contact between the youths and mentors. Although mentoring that is exclusively based on digital contact may be a *better than nothing* solution for some, informants in this study instead focused on the ways digital support could enhance in-person mentee-mentor encounters and relationships. This echoes the view that the core benefit of mentoring is the *relationship* between the mentee and mentor and concerns that e-mentoring might undermine the dyads' ability to forge quality relationships [17]. Support for broader networks among mentees and mentors, both individually and collectively, was also emphasized, again as a way of fostering in-person relationships. These findings reflect Catalysts' own theory of change, highlighting specific areas where e-mentoring could enhance their traditional mentoring in terms of managing the relationships, supporting social capital, and ensuring security.

Managing the Relationships

Reflective of the primacy placed on the mentee-mentor relationship, informants' suggestions for digital resources largely had to do with fostering in-person relationships and activities toward specific goals (relating to main themes 1 and 2). Information that aimed to boost mentor skills and sense of efficacy (eg, case descriptions and various scenarios for responding) as well as mentees' skills and self-efficacy in eliciting mentor support for their given needs were suggested [35]. Specific tools (forum and toolbox) to support mentee-mentor collaboration around concrete tasks were also important in this regard. For example, many of the suggestions for digital support had to do with tools that could boost specific skills such as Norwegian training, writing a CV, and goal-based planning, along with ways the mentors and mentees could become engaged in developing skills. This ties in with what others have suggested [36]—that specific activities toward concrete goals (*relationship as the context for an intervention*) are preferable to friendship models of mentoring where the main objective is to forge a close bond (with the relationship as the intervention). Although the forum and toolbox were considered valuable for relational support, sending private messages was viewed as less important by most of the stakeholders, as existing tools (eg, SMS and WhatsApp) were largely viewed as sufficient.

Social Capital Support

Digital options for facilitating connection and goal orientation (main themes 1 and 2) are further discussed here in the light of social capital [37,38]. Although some discuss digital options for boosting social capital with regard to the plethora of opportunities provided through social media [39], we chose a more limited approach, due to the limited scope of the ReConnect platform to structure the social interactions among the youths and adults.

Informant suggestions for forums varied between those just for mentees, mentors, or combinations, and also in conjunction with offline activities. These permutations have implications in terms of how they might enable or limit different forms of social

capital. For example, forums exclusive to mentees were framed as enhancing horizontal bonding types of social capital, that is, connections to other relatively similar youths [23]. At the same time, it was argued that this type of peer support could help youths in navigating and building their relationships with adult mentors. In addition, mentees' expressed desire to *pay it forward* to other mentees could be facilitated through forum connection and belonging [40]. In contrast, a group forum for mentees and mentors might cultivate bridging social capital, allowing connections across boundaries of ethnicity, age, and social status [41]. Mentees argued that this would considerably expand their access to mentors who may have more relevant skills than their own mentor and are therefore better able to advise in terms of goal achievement.

Arguments for including a shared calendar function showing local events were that it would offer a way for mentees to connect with others and participate in activities that they otherwise might not have prioritized or even been aware of. Some of the youths revealed a desire to increase social connections in their life and extend their networks, whereas for a few, this was viewed as less significant than enhancing the individual relationships they had with their mentors.

Both mentors and mentees wanted the toolbox to provide resources for individual knowledge building and goal achievement and to provide tools to guide the mentor-mentee interactions. Completing activities together in person might be even more effective in strengthening individual-level trust when supported by follow-up options through e-mentoring, again, potentially providing a foundation for more generalized trust. The importance of trust, a central aspect of social capital [18,19], was evident in how mentors and mentees discussed requirements for environments that fostered genuine supportive sharing [17,42].

Building trust across categories of ethnicity, culture, age, and socioeconomic class inevitably requires grappling with challenges of power dynamics and both conscious and unconscious preconceptions [43]. In addition to the training the mentors and mentees receive from the nondigital program, a digital platform may aid in reducing these barriers and their adverse impacts on trust and social capital.

Security and Control

Having control over personal information appeared especially important for some of the youths, particularly those who have experienced social control. This was illustrated in the discussions of login processes, where a mentee expressed gratitude for Norway's strict security requirements compared with their country of origin, as well as in terms of information sharing. On the one hand, as shown in other studies [44], anonymity in a digital social setting can allow information and ideas to flow more freely, for example, by decreasing shyness, and on the other hand, it can also result in the lack of accountability or credibility and deindividuation [44]. Confidentiality is important in mentoring as it may influence the amount of information a mentor or mentee reveals in their relationship [45], which, in turn, contributes to the development of trust [46]. However, previous studies also indicate that confidentiality may be more challenging to maintain in the context of virtual mentoring because users can perceive a false sense of security and disclose more than they might otherwise do face to face [47]. Offline activities, thus, remain critical in enhancing social ties and trust [44], which is again supported by the respondents' dismissal of replacing in-person contact with an app.

Limitations and Future Work

Although the Catalysts program endeavors to recruit youths with genuine needs for support, those who participate are self-selected and likely to be more motivated than those who do not show interest for the program. Thus, the participants in this study may have possessed characteristics that distinguish them from the broader population of multicultural youths in Norway. Using ReConnect to facilitate discussions about needs for digital support likely influenced the direction of some interviews. As our point of departure was limited by preparations for an intervention study, which requires a closed and secure environment for social interactions among youths and adults, future research could investigate the benefits and limitations of open social media platforms, compared with more secure and closed environments for mentor-mentee interactions. Also, although not explicitly examined in this study, the individual characteristics of the youths (and adults) in this study (eg, gender, ethnicity, length of time in Norway, and age) can play an important role in their articulated needs and requirements [48]. Future studies would benefit from exploring how these factors influence the reported needs.

Conclusions

This study investigated mentors' and mentees' experiences and needs, providing valuable insights into how to design digital solutions to supplement in-person mentoring of multicultural youth. The desire for connection at the individual and community levels was salient. In addition, a strong emphasis on goal achievement from the stakeholders underscored the motivated character of the youths and the strong desire of mentors to help. Concerns about security and privacy were prevalent in the discussions. Despite the promise of digitalization, mentors and mentees emphasized that the platform should not replace in-person contact. Given attention to the unique needs and preferences of mentoring stakeholders, the features of existing eHealth apps can be adapted to a mentoring program context, potentially enhancing the mentee-mentor relationships and fostering mentee social capital.

XSL•FO

RenderX

Conflicts of Interest

SP is employed by Catalysts Association and Catalysts Technologies, a spin-off from Catalysts Association. Catalysts Technologies holds the rights to commercialize any digital tools resulting from this research. The other coauthors have no conflicts of interest.

Multimedia Appendix 1
Supplementary quotations illustrating themes and subthemes (spreadsheet).
[XLSX File (Microsoft Excel File), 23 KB - formative_v4i2e15500_app1.xlsx]

References

1. Freudenberg N, Ruglis J. Reframing school dropout as a public health issue. Prev Chronic Dis 2007 Oct;4(4):A107 [FREE Full text] [Medline: 17875251]
2. Reegård K, Rogstad J, Keeping D. The Drop-outs. Drop-out in upper secondary education.. Oslo, Norway: Gyldendal; 2016.
3. de Ridder KA, Pape K, Johnsen R, Westin S, Holmen TL, Bjørngaard JH. School dropout: a major public health challenge: a 10-year prospective study on medical and non-medical social insurance benefits in young adulthood, the Young-HUNT 1 Study (Norway). J Epidemiol Community Health 2012 Nov;66(11):995-1000. [doi: 10.1136/jech-2011-200047] [Medline: 22315238]
4. Olsen B. Youth with immigrant backgrounds in work and education 2016. Oslo, Norway: Statistics Norway; 2018.
5. Garringer M, Kupersmidt J, Rhodes J, Stelter R, Tai T. MENTOR: The National Mentoring Partnership. Boston, MA; 2015. Elements of Effective Practice for Mentoring URL: https://www.mentoring.org/program-resources/elements-of-effective-practice-for-mentoring/
6. DuBois DL, Portillo N, Rhodes JE, Silverthorn N, Valentine JC. How effective are mentoring programs for youth? A systematic assessment of the evidence. Psychol Sci Public Interest 2011 Aug;12(2):57-91. [doi: 10.1177/1529100611414806] [Medline: 26167708]
7. Raposa EB, Rhodes J, Stams GJ, Card N, Burton S, Schwartz S, et al. The effects of youth mentoring programs: a meta-analysis of outcome studies. J Youth Adolesc 2019 Mar;48(3):423-443. [doi: 10.1007/s10964-019-00982-8] [Medline: 30661211]
8. Loga J, Eimhjellen I, Eschweiler J, Ingstad E, Stokstad S, Winsvold M. Social Entrepreneurs - Partnership for New Solutions. Bergen: Uni Research Rokkansenteret; 2016.
9. Orem SL, Binkert J, Clancy AL. Appreciative Coaching: A Positive Process For Change. San Francisco: Jossey-Bass; 2007.
10. San Martin TL, Calabrese RL. Empowering at‐risk students through appreciative inquiry. Int J Educ Manag 2011;25(2):110-123. [doi: 10.1108/09513541111107542]
11. Snyder CR, Lopez SJ, Edwards LM, Marques SC, editors. The Oxford Handbook of Positive Psychology. Third Edition. Oxford: Oxford University Press; 2016.
12. Gammon D, Strand M, Eng LS. Service users' perspectives in the design of an online tool for assisted self-help in mental health: a case study of implications. Int J Ment Health Syst 2014 Jan 9;8(1):2 [FREE Full text] [doi: 10.1186/1752-4458-8-2] [Medline: 24405765]
13. Gammon D, Strand M, Eng LS, Børøsund E, Varsi C, Ruland C. Shifting practices toward recovery-oriented care through an e-recovery portal in community mental health care: a mixed-methods exploratory study. J Med Internet Res 2017 May 2;19(5):e145 [FREE Full text] [doi: 10.2196/jmir.7524] [Medline: 28465277]
14. Leamy M, Bird V, Le Boutillier C, Williams J, Slade M. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry 2011 Dec;199(6):445-452. [doi: 10.1192/bjp.bp.110.083733] [Medline: 22130746]
15. Garringer M, McQuillin S, McDaniel H. Mentoring. 2017. Examining Youth Mentoring Services Across America. Findings from the 2016 National Mentoring Program Survey URL: https://www.mentoring.org/new-site/wp-content/uploads/2017/07/Mentor-Survey-Report_FINAL_small.pdf [accessed 2019-11-11]
16. Kaufman M. National Mentoring Resource Center. 2017. E-Mentoring URL: http://nationalmentoringresourcecenter.org/images/PDF/E-Mentoring_Model_Review.pdf [accessed 2019-11-11]
17. Rhodes JE, Spencer R, Saito RN, Sipe CL. Online mentoring: the promise and challenges of an emerging approach to youth development. J Prim Prev 2006 Sep;27(5):497-513. [doi: 10.1007/s10935-006-0051-y] [Medline: 16897405]
18. Coleman JS. Social capital in the creation of human capital. AM J Sociol 1988;94:S95-120. [doi: 10.1086/228943]
19. Putnam R. Bowling Alone: The Collapse and Revival of American Community. New York: Simon & Schuster; 2000.
20. Rostila M. The facets of social capital. J Theory Soc Behav 2010;41(3):308-326. [doi: 10.1111/j.1468-5914.2010.00454.x]
21. Schwartz SE, Kanchewa SS, Rhodes JE, Gowdy G, Stark AM, Horn JP, et al. 'I'm Having a Little Struggle With This, Can You Help Me Out?': examining impacts and processes of a social capital intervention for first-generation college students. Am J Community Psychol 2018 Mar;61(1-2):166-178. [doi: 10.1002/ajcp.12206] [Medline: 29178300]
22. Bourdieu P. The forms of capital. In: Richardson JG, editor. Handbook of Theory and Research for the Sociology of Education. New York: Greenwood Press; 1986:241-258.

23. Ryan L. Migrants' social networks and weak ties: accessing resources and constructing relationships post-migration. Sociol Rev 2011;59(4):707-724. [doi: 10.1111/j.1467-954X.2011.02030.x]

24. McPherson KE, Kerr S, McGee E, Morgan A, Cheater FM, McLean J, et al. The association between social capital and mental health and behavioural problems in children and adolescents: an integrative systematic review. BMC Psychol 2014;2(1):7 [FREE Full text] [doi: 10.1186/2050-7283-2-7] [Medline: 25566380]

25. DuBois DL, Holloway BE, Valentine JC, Cooper H. Effectiveness of mentoring programs for youth: a meta-analytic review. Am J Community Psychol 2002 Apr;30(2):157-197. [doi: 10.1023/A:1014628810714] [Medline: 12002242]

26. Creswell JW. Research Design: Qualitative, Quantitative and Mixed Methods Approaches. Fourth Edition. Los Angeles: Sage; 2014.

27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

28. Edge S, Newbold KB, McKeary M. Exploring socio-cultural factors that mediate, facilitate, & constrain the health and empowerment of refugee youth. Soc Sci Med 2014 Sep;117:34-41. [doi: 10.1016/j.socscimed.2014.07.025] [Medline: 25036014]

29. Birman D, Morland L. Immigrant and refugee youth. In: DuBois DL, Karcher MJ, editors. Handbook of Youth Mentoring. Second Edition. Los Angeles: Sage; 2014:355-368.

30. Stanton-Salazar RD, Spina SU. Informal mentors and role models in the lives of urban Mexican-origin adolescents. Anthropol Educ Q 2003;34(3):231-254. [doi: 10.1525/aeq.2003.34.3.231]

31. Caldarella P, Gomm RJ, Shatzer RH, Wall DG. School-based mentoring: a study of volunteer motivations and benefits. Int Electron J Elem Educ 2010 Mar;2(2):199-216 [FREE Full text]

32. Reeves RB. Inter-cultural mentoring for newcomer immigrants: mentor perspectives and better practices. Int J Evid Based Coach Mentor 2017 Feb;15(1):186-207 [FREE Full text]

33. Dutton H, Bullen P, Deane KL. 'It is OK to let them know you are human too': mentor self-disclosure in formal youth mentoring relationships. J Community Psychol 2019 May;47(4):943-963. [doi: 10.1002/jcop.22165] [Medline: 30714171]

34. Rhodes J, Liang B, Spencer R. First do no harm: ethical principles for youth mentoring relationships. Prof Psychol 2009;40(5):452-458. [doi: 10.1037/a0015073]

35. Karcher MJ, Nakkula MJ, Harris J. Developmental mentoring match characteristics: correspondence between mentors' and mentees' assessments of relationship quality. J Prim Prev 2005 Mar;26(2):93-110. [doi: 10.1007/s10935-005-1847-x] [Medline: 15977044]

36. Cavell TA, Elledge LC. Mentoring and prevention science. In: DuBois DL, Karcher MJ, editors. Handbook of youth mentoring. Second Edition. Los Angeles: Sage; 2014:29-42.

37. Mandarano L, Meenar M, Steins C. Building social capital in the digital age of civic engagement. J Plan Lit 2010;25(2):123-135. [doi: 10.1177/0885412210394102]

38. Skoric MM, Zhu Q, Goh D, Pang N. Social media and citizen engagement: a meta-analytic review. New Media Soc 2016;18(9):1817-1839. [doi: 10.1177/1461444815616221]

39. Fisher JF. Who You Know: Unlocking Innovations That Expand Students' Networks. San Francisco: Jossey-Bass; 2018.

40. Spencer R. 'It's Not What I Expected': a qualitative study of youth mentoring relationship failures. J Adolesc Res 2007;22(4):331-354. [doi: 10.1177/0743558407301915]

41. Gericke D, Burmeister A, Löwe J, Deller J, Pundt L. How do refugees use their social capital for successful labor market integration? An exploratory analysis in Germany. J Vocat Behav 2018;105:46-61. [doi: 10.1016/j.jvb.2017.12.002]

42. Raithelhuber E. 'If we want, they help us in any way': how 'unaccompanied refugee minors' experience mentoring relationships. Eur J Soc Work 2019 Apr 22. [doi: 10.1080/13691457.2019.1606787]

43. Leck JD, Wood PM. Forming trust in e-mentoring: a research agenda. Am J Ind Bus Manag 2013;3(1):101-109. [doi: 10.4236/ajibm.2013.31013]

44. Shiue YC, Chiu CM, Chang CC. Exploring and mitigating social loafing in online communities. Comput Human Behav 2010;26(4):768-777. [doi: 10.1016/j.chb.2010.01.014]

45. Panopoulos AP, Sarri K. E-mentoring: the adoption process and innovation challenge. Int J Inf Manag 2013;33(1):217-226. [doi: 10.1016/j.ijinfomgt.2012.10.003]

46. Evans C. Trust and connection in formal virtual mentoring. Int J Evid Based Coach Mentor 2018;S12:154-164. [doi: 10.24384/000544]

47. Leck J, Orser B. Fostering trust in mentoring relationships: an exploratory study. Equal Divers Incl 2013;32(4):410-425. [doi: 10.1108/EDI-01-2010-0007]

48. Darling N, Bogat GA, Cavell TA, Murphy SE, Sánchez B. Gender, ethnicity, development, and risk: mentoring and the consideration of individual differences. J Commun Psychol 2006;34(6):765-780. [doi: 10.1002/jcop.20128]

Abbreviations

CV: curriculum vitae
eHealth: electronic health

e-mentoring: electronic mentoring
NAV: Norwegian Labor and Welfare Administration
NORCE: Norwegian Research Center

XSL·FO
RenderX

Original Paper

The Diabetes and Emotional Health Handbook and Toolkit for Health Professionals Supporting Adults With Type 1 and Type 2 Diabetes: Formative Evaluation

Jennifer A Halliday[1,2], BHealthSciences (Hons); Jane Speight[1,2,3], PhD; Andrea Bennet[1,2], PhD; Linda J Beeney[4,5], PhD; Christel Hendrieckx[1,2], PhD

[1]School of Psychology, Deakin University, Geelong, Australia

[2]The Australian Centre for Behavioural Research in Diabetes, Diabetes Victoria, Melbourne, Australia

[3]Applied Health Psychology Research, Hornchurch, Essex, United Kingdom

[4]Sydney Medical School, The University of Sydney, Sydney, Australia

[5]Diabetes & Medical Psychology Services, Normanhurst, Australia

Corresponding Author:
Jennifer A Halliday, BHealthSciences (Hons)
School of Psychology
Deakin University
Locked Bag 20000
Geelong, 3220
Australia
Phone: 61 (0) 3 92445278
Email: j.halliday@deakin.edu.au

Abstract

Background: Health professionals have expressed unmet needs, including lacking the skills, confidence, training, and resources needed to properly attend to the psychological needs of people with diabetes.

Objective: Informed by needs assessments, this study aimed to develop practical, evidence-based resources to support health professionals to address the emotional needs of adults with type 1 or type 2 diabetes.

Methods: We developed a new *handbook* and *toolkit* informed by formative evaluation, including literature reviews, stakeholder consultation and review, and a qualitative study. In the qualitative study, health professionals participated in interviews after reading sections of the *handbook* and *toolkit*.

Results: The literature review uncovered that psychological problems are common among adults with diabetes, but health professionals lack resources to provide related support. We planned and drafted resources to fill this unmet need, guided by stakeholder consultation and an Expert Reference Group (ERG). Before finalizing the resources, we implemented feedback received from stakeholders (ERG, health professionals, academics, and people with diabetes). The resulting resources were the practical, evidence-based *Diabetes* and *Emotional Health* handbook and toolkit. A total of 19 health professionals took part in the qualitative study about the *handbook* and *toolkit*. They viewed the resources favorably, felt empowered to support people with diabetes experiencing psychological problems, and felt motivated to share the resources with others. Some gave examples of how they had used the *handbook* in clinical practice. A perceived highlight was the inclusion of a process model outlining 7 steps for identifying and supporting people with emotional problems: the 7 A's model. With funding from the National Diabetes Services Scheme (NDSS), more than 2400 copies of *Diabetes and Emotional Health* have been distributed. It is freely available on the Web. The NDSS is an initiative of the Australian Government administered with the assistance of Diabetes Australia.

Conclusions: The new evidence-based resources are perceived by stakeholders as effective aids to assist health professionals in providing emotional support to adults with diabetes. The 7 A's model may have clinical utility for routine monitoring of other psychological and health-related problems, as part of person-centered clinical care.

(JMIR Form Res 2020;4(2):e15007) doi:10.2196/15007

XSL•FO
RenderX

Introduction

Diabetes and Emotional Health

Diabetes is a serious chronic condition affecting more than 415 million people worldwide, and this number is rising [1]. The 2 most common forms are type 1 and type 2 diabetes. Diabetes is characterized by high blood glucose levels, which over many years increase the risk of diabetes-related complications. Therefore, it is important to manage the condition to maintain optimal glucose levels (and minimize other risk factors) [2]. This is highly reliant on daily self-care (eg, taking medications, checking glucose levels, and maintaining a healthy lifestyle). Living with and managing diabetes can place considerable psychological burden on the person and their family [3]. Psychological problems can be both diabetes-specific (eg, diabetes distress and fear of hypoglycemia) and generic (eg, depression and anxiety) [4]. They are relatively common among adults with both types of diabetes, with many similarities in terms of the nature and prevalence of presenting problems as well as the strategies to address them. Preservation of psychological well-being is important in its own right [5] but impaired well-being is also associated with suboptimal diabetes self-care, biomedical outcomes (eg, HbA_{1c}), and quality of life.

Recommendations for Holistic Diabetes Care and Unmet Needs of Health Professionals

Clinical guidelines recommend person-centered, holistic diabetes care, including routine monitoring for psychological problems [6-9]. However, such guidelines are rarely implemented in clinical practice and the emotional health needs of people with diabetes often go unmet [10-13]. Few evidence-based, comprehensive, practical resources exist for health professionals about how to implement psychological care, beyond book chapters [14,15], books highly specific to 1 psychological problem [16,17], and student texts [18]. Although these have value, they are not freely available to health professionals and are limited in their practical application. For example, they are often heavily text-based or do not provide a step-by-step guide, including the practical elements that health professionals need, such as practice points, suggestions for open-ended questions

and responses, and copies of validated screening tools for psychological problems. Health professionals have expressed unmet needs, including lacking skills, confidence, training, and resources, to properly attend to the psychological needs of people with diabetes [19-21]. Recommendations have been made recently for improved communications between health professionals and people with diabetes and for additional training and improved skills among health professionals to assist with overcoming these barriers [22].

Aim and Objectives

Therefore, our aim was to address health professionals' unmet needs, using a formative evaluation approach, by developing a practical, evidence-based resource: *Diabetes and Emotional Health: A handbook and toolkit for health professionals supporting adults with type 1 or type 2 diabetes* [23]. The objectives of the *handbook* and *toolkit* were to raise awareness, provide practice points and tools, and foster skill development, for identifying, communicating about, and addressing psychological problems experienced by adults with diabetes. This paper describes the formative evaluation of the *handbook* and *toolkit*.

Methods

Overview

To develop the *handbook* and *toolkit*, we used a formative evaluation approach, that is, "a set of activities designed to develop, identify and test program materials and methods … [which] occurs as a part of program planning and occurs before any elements of the program are implemented" [24]. We selected this approach to enable us to shape the resources in accordance with the literature, best practice clinical recommendations, and stakeholder consultations, to meet the needs of diabetes health professionals. The formative evaluation steps we applied (ie, review the problem, understand the target population, and pretest the intervention material) [24] are shown in Figure 1 and described in the following sections. The methods of each phase were informed by the results of the previous phase (ie, an iterative process). The decisions made after each phase are described in each corresponding results section.

Figure 1. Summary of the formative evaluation of the Diabetes and Emotional Health handbook and toolkit.
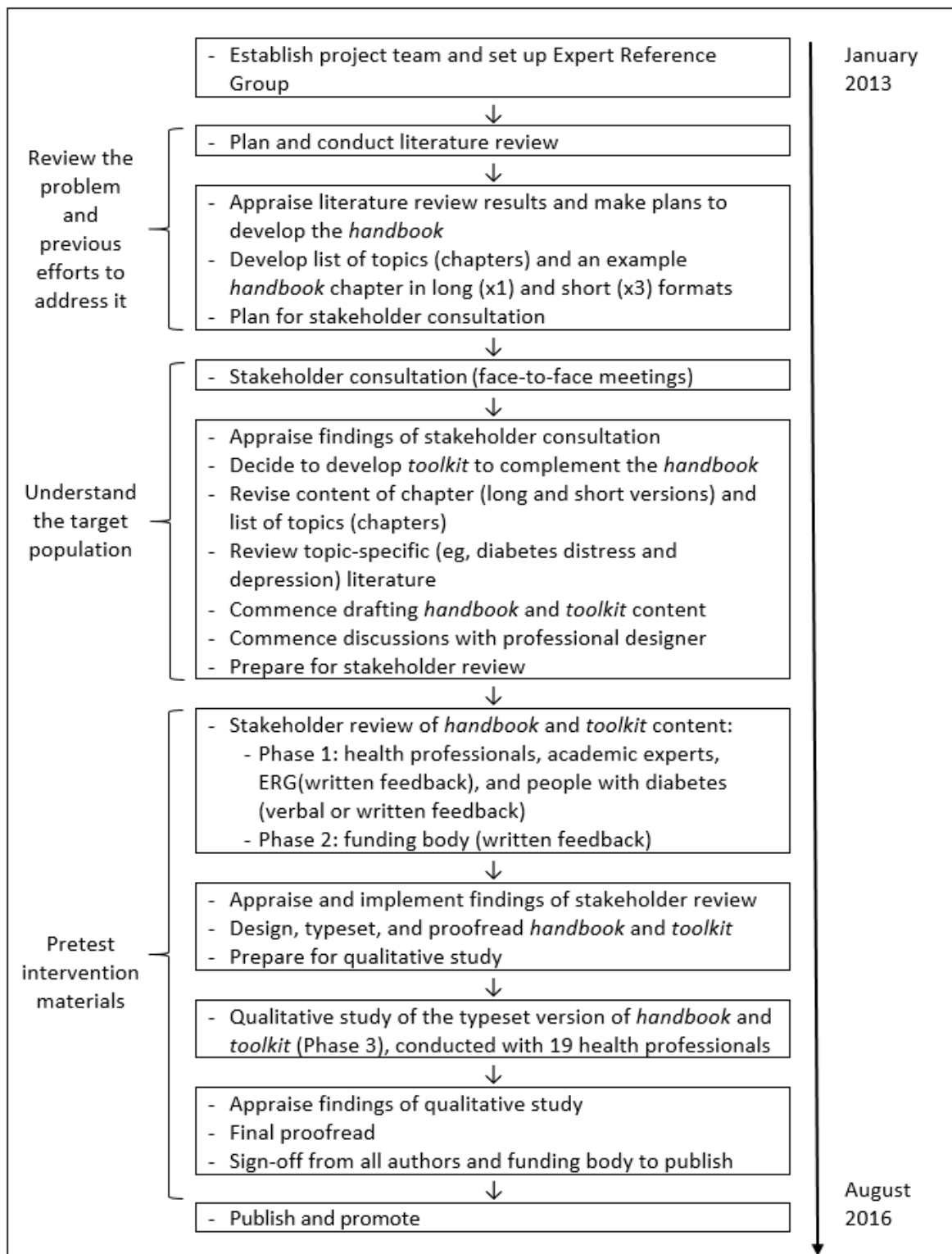
Establish Project Team and Set Up Expert Reference Group

To begin, we established a project team that included health and clinical psychologists and researchers with expertise in the psychological aspects of diabetes. We also established a multidisciplinary Expert Reference Group (ERG), representing various stakeholders (eg, general practitioner, endocrinologist, diabetes nurse educator, psychologist, key organizations, and people with diabetes). The project team worked collaboratively throughout the project, meeting regularly, holding workshops to discuss and progress ideas, and constructively reviewing each other's work. We engaged the ERG at least quarterly to discuss plans and progress.

XSL•FO
RenderX

Review the Problem and Previous Efforts to Address It: Literature Review

In 2013, we conducted a narrative literature review to investigate 7 research questions (see Multimedia Appendix 1). We developed these research questions using a stepwise approach, to help formulate an evidence base for developing a resource to support diabetes health professionals with providing psychological support. We engaged our ERG in the process of formulating the research questions. We searched MEDLINE, Scopus, Google (for gray literature), and the reference lists of relevant literature. Peer-reviewed narrative and systematic reviews and meta-analyses, empirical research (quantitative and qualitative), expert commentaries, reports, and clinical guidelines were consulted. The search was limited to literature relevant to adults with type 1 or type 2 diabetes, published in English from the year 2000 onward. Search terms varied according to the research question (eg, Diabetes and Psychological). Truncations and synonyms were used as appropriate.

The literature review confirmed the need for practical, user-friendly, and evidence-based resources; thus, we decided to develop a handbook to meet this need (See "Review the Problem and Previous Efforts to Address It: Literature Review" in the Results for further information).

Understand the Target Population: Stakeholder Consultation

From late 2013 to early 2014, JH and LB consulted with each ERG member individually to (1) gain a deeper understanding of their current practice related to diabetes-related psychological care and (2) ascertain their needs (individual and professional, content and design) for the *handbook*. The consultations were unstructured to ensure the relevance to all ERG members and used funneling (moving from broad to specific topics). Before the consultations, the project team agreed on a general list of topics to discuss, but they also agreed to allow the stakeholders to expand on any relevant topic (consistent with an unstructured consultation approach). For example, after the person had spoken, unprompted, about their experiences and needs, they were shown pre prepared examples of draft *handbook* content, in long (*chapter*) and short (3 *summary versions*—5 A's model, flowchart, and factsheet) formats. We explored their preferred version, positive and negative aspects of each version, how each version would meet their needs (or not), and how we could adapt each version to better suit their needs. We also built on topics raised by other ERG members (eg, "Someone else suggested X, what do you think?") and presented a draft list of topics (chapters) to check relevance and completeness. We made comprehensive notes during the consultations. JH examined all notes for commonalities and differences in views, then summarized the results (see "Understand the Target Population: Stakeholder Consultation" in the Results) and discussed them with the team. We reported back to the ERG to validate the findings and seek agreement for proposed plans and decisions, which included a decision to develop both a *handbook* and a *toolkit*, to meet the varying needs of different health professional disciplines (see "Understand the Target Population: Stakeholder Consultation" in Results).

Pretest Intervention Methods and Materials: Phases 1–3

Phase 1: Review by Professionals and People With Diabetes

A review of each completed draft chapter and its corresponding summary and questionnaire cards was conducted to ensure that the content was consistent with best practice and recent evidence and met the needs of the intended audience. The factsheet (part of the *toolkit*) development is outside the scope of this study, but stakeholders also reviewed these resources. In total, four professionals reviewed each chapter (plus the corresponding summary and questionnaire cards): one health professional, one academic expert, and two ERG members. We invited the reviewers on the basis of their clinical or academic experience and relevant publication record. The reviewers were multidisciplinary, including psychologists, endocrinologists, general practitioners, dietitians, and credentialed diabetes educators. An adult with type 1 or type 2 diabetes also reviewed each chapter, to ensure they deemed the content as appropriate (eg, language use, references to people with diabetes, their experiences, and suggested strategies or techniques).

Owing to the significant time and work required to review the chapters (estimated minimum 3 hours), we offered remuneration to all reviewers. We provided the reviewers with background information (eg, aim, scope, content, and target population), instructions (eg, scope for the review), and a nontypeset copy of the chapter and summary (both annotated with guiding questions). JH prepared and provided guiding questions relevant to each reviewer's background to ensure the reviews were within the scope of the individual's expertise (eg, medical content was not the responsibility of mental health professionals or people with diabetes). The professionals provided written feedback, while the people with diabetes provided written or verbal feedback. JH met (face-to-face or telephone) with those who opted to provide verbal feedback, making comprehensive notes during and immediately after the conversation. Once all reviewers provided feedback, we consolidated the feedback for each chapter into a single document for team review and implementation.

Phase 2: Review by Funding Body

The National Diabetes Services Scheme (NDSS; funding body) required its Medical, Education, and Scientific Advisory Council (MESAC) to review and approve the resources before publication. We submitted a standardized form with background information in addition to the nontypeset *handbook* and *toolkit* drafts. The MESAC reviewed the draft *handbook* and *toolkit* (summary and questionnaire cards) to provide tracked changes and comments. The factsheets were also subject to MESAC review, but further details of their development and review is beyond the scope of this paper.

Phase 3: Interviews With Health Professionals

As a final step, we undertook a qualitative study of the *typeset versions* of the *handbook* and *toolkit* (summary and questionnaire cards). We aimed to (1) collect feedback about the *handbook's* content, structure, and usability and (2) understand health professionals' perspectives on implementing

the strategies described in the *handbook*. We developed a Web-based screening survey and a semistructured telephone interview schedule. The survey included demographic and clinical characteristics and brief questions, such as confidence to talk about and assist with the emotional aspects of diabetes (measured on a Likert scale from 1=not at all confident to 5=very confident). The interview schedule included questions about the resources (eg, overall impressions of *handbook*, what they learned, and implementation plans).

We promoted the study via newsletter advertisements and direct emails to health professionals who had previously registered their interest. Participants were eligible if they worked in Australia as a general practitioner, endocrinologist, credentialed diabetes educator, nurse, or dietitian and consulted with at least 10 adults with diabetes weekly. About 2 weeks before the interview, we sent consenting participants a typeset copy of the *handbook* and summary cards, details of the study procedure, and a summary interview guide. They were asked to read the "Introduction;" "How to use this *handbook* and *toolkit;*" and their choice of at least one of the chapters focused on a particular psychological problem (ie, Chapters 3 to 8) and the associated summary card(s).

A researcher who did not develop the *Diabetes and Emotional Health* resources (AB) conducted the interviews. The interviews were audio-recorded and transcribed professionally. Subsequently, AB summarized responses and categorized quotes from the transcripts according to the interview questions. CH checked the categorization and summaries, then prepared a report of the findings, which AB, JH, JS, and the ERG reviewed and approved.

Results

Review the Problem and Previous Efforts to Address It: Literature Review

The literature review demonstrated that psychological problems are common among Australian adults with type 1 and type 2 diabetes [25-27]. Furthermore, Australian adults with diabetes attending tertiary diabetes clinics respond well to the use of psychological screening questionnaires [27,28]. International guidelines recommend attention to and routine screening for psychological problems as a part of routine care [29] yet health professionals report lacking resources, training, and confidence to do so [19,30]. We presented these results in a report that the ERG reviewed and approved. These findings and "Lessons Learned and Actions" are summarized in Multimedia Appendix 1.

On the basis of the literature review findings (see Multimedia Appendix 1), we decided to develop a practical, user-friendly, and evidence-based resource (ie, the *Diabetes and Emotional Health handbook*). The *handbook* would cover a range of diabetes-specific and general psychological problems common to adults with type 1 or type 2 diabetes, which affect their diabetes self-care and outcomes and quality of life. It would provide health professionals with practical tools and strategies for identifying, communicating about, and addressing psychological problems. The ERG supported our proposed plans. They also pointed out that health professionals often have limited time and would appreciate short (eg, 1 page) summaries, with an option to access more detail and background information in the *handbook*.

The literature review informed the *handbook* planning and content. We began by brainstorming a list of proposed topics (ie, chapters; see Figure 2) and chapter content (proposed headings and sections). Drawing on content-specific literature, we prepared a draft chapter, proposing a standardized format with the following headings (structure): key messages, prevalence, risk factors, impacts, what to look for, treatment or management, case studies, resources, and a copy of a validated screening questionnaire. We also investigated formats for the summaries and prepared drafts in three formats (5 A's model, flowchart, and factsheet). We selected these potential formats because they are common layouts in other resources (eg, the 5 A's model appeared in existing Australian health professional guidelines) [31]. The 5 A's model is a well-cited adaptation of the 4 A's model (see Figure 3) [32-34].

Figure 2. Overview of the Diabetes and Emotional Health handbook and toolkit.

Diabetes and Emotional Health is an evidence-based, clinically informed, practical *handbook* and *toolkit* (described below) to support health professionals in meeting the psychological health needs of adults with type 1 or type 2 diabetes. It is available in hardcopy (limited supply) and electronically (ndss.com.au). It offers strategies and tools for how to recognise and have conversations about psychological problems, and for providing appropriate support.

The *handbook* includes:
- How to use this handbook
- Chapter 1: Communication and engagement
- Chapter 2: Facing life with diabetes
- Chapter 3: Diabetes distress
- Chapter 4: Fear of hypoglycaemia
- Chapter 5: Psychological barriers to insulin use
- Chapter 6: Depression
- Chapter 7: Anxiety disorders
- Chapter 8: Eating problems
- Chapter 9: Referring to a mental health professional
- Appendix A: Diabetes Australia's position statement 'A new language for diabetes' [35]
- Appendix B: Peer support
- Appendix C: Examples of strategies for diabetes distress
- Appendix D: Examples of strategies for psychological barriers to insulin use.
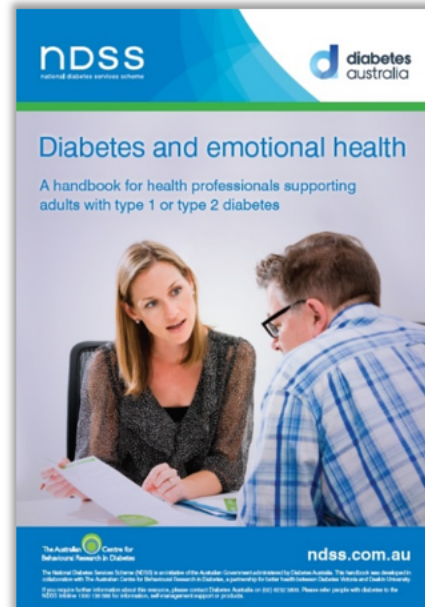
Chapters 3-8 follow a consistent chapter structure using the 7 A's model (AWARE, ASK, ASSESS, ADVISE, ASSIST, ASSIGN and ARRANGE; Figure 3). Each includes: key messages and practice points, background information (e.g. what it is, how common, consequences), how to identify problems (e.g. signs to look for, suggested open-ended questions, validated screening tool), how to support (e.g. conversation points, strategies, referral options), case studies, additional resources (for people with diabetes and health professionals), and references.

The *toolkit* contains practical resources to facilitate implementation of the *handbook*:
- Summary cards: double-sided A4 cards summarising Chapters 3-8 for 'quick reference' (e.g. kept in drawer).
- Questionnaire cards: double-sided A4 cards with screening questionnaires and scoring instructions, to facilitate easy photocopying for routine use.
- Factsheets for people with diabetes: to facilitate conversations about the emotional problems in the *handbook*.
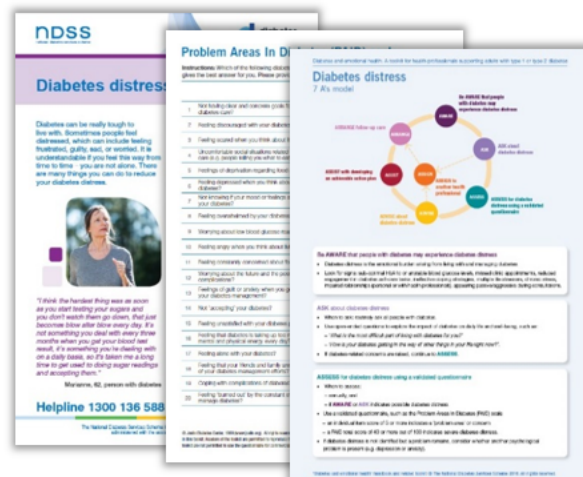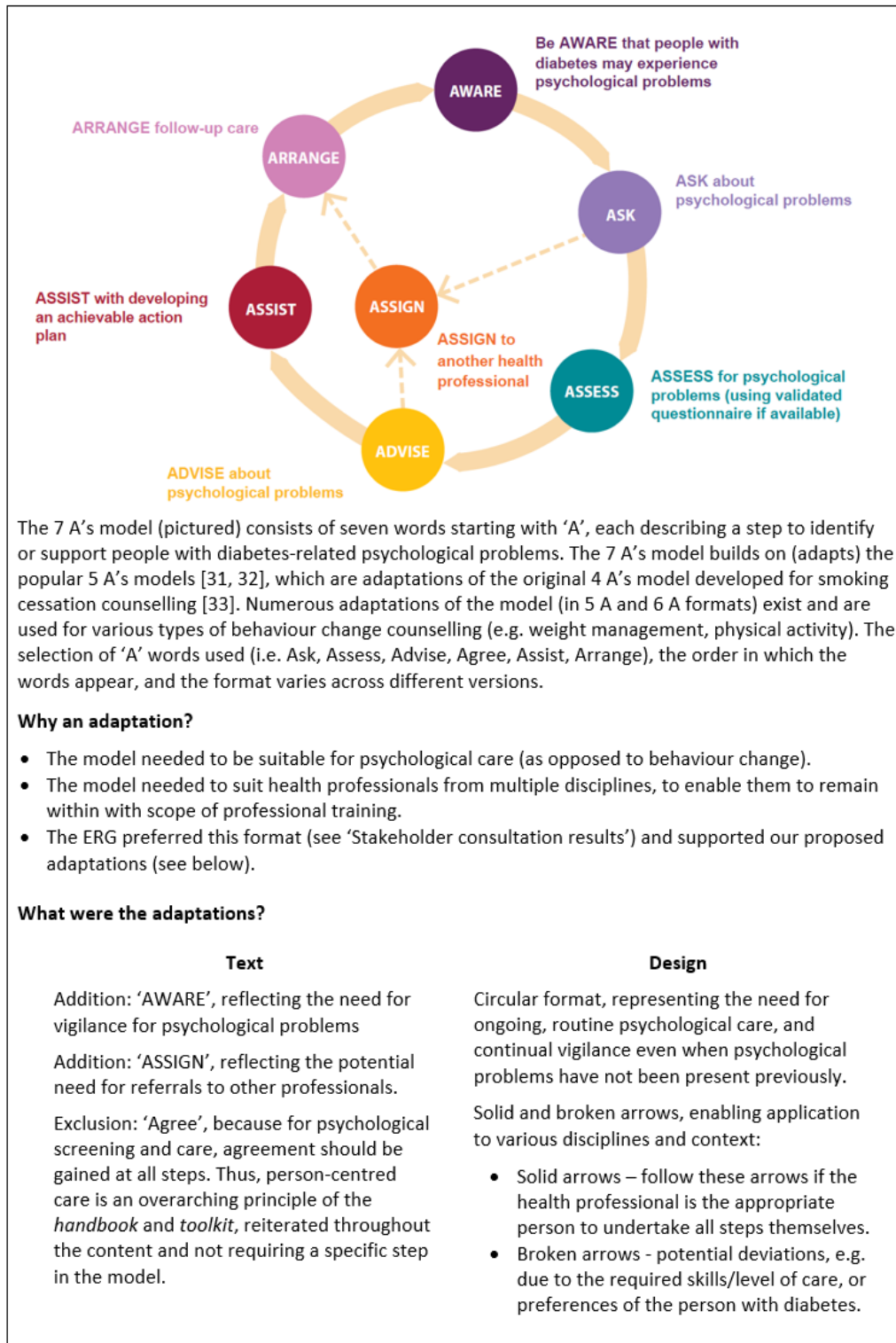
Figure 3. Overview of the 7 A's model.

The 7 A's model (pictured) consists of seven words starting with 'A', each describing a step to identify or support people with diabetes-related psychological problems. The 7 A's model builds on (adapts) the popular 5 A's models [31, 32], which are adaptations of the original 4 A's model developed for smoking cessation counselling [33]. Numerous adaptations of the model (in 5 A and 6 A formats) exist and are used for various types of behaviour change counselling (e.g. weight management, physical activity). The selection of 'A' words used (i.e. Ask, Assess, Advise, Agree, Assist, Arrange), the order in which the words appear, and the format varies across different versions.

Why an adaptation?

- The model needed to be suitable for psychological care (as opposed to behaviour change).
- The model needed to suit health professionals from multiple disciplines, to enable them to remain within with scope of professional training.
- The ERG preferred this format (see 'Stakeholder consultation results') and supported our proposed adaptations (see below).

What were the adaptations?

Text	Design
Addition: 'AWARE', reflecting the need for vigilance for psychological problems	Circular format, representing the need for ongoing, routine psychological care, and continual vigilance even when psychological problems have not been present previously.
Addition: 'ASSIGN', reflecting the potential need for referrals to other professionals.	Solid and broken arrows, enabling application to various disciplines and context:
Exclusion: 'Agree', because for psychological screening and care, agreement should be gained at all steps. Thus, person-centred care is an overarching principle of the *handbook* and *toolkit*, reiterated throughout the content and not requiring a specific step in the model.	• Solid arrows – follow these arrows if the health professional is the appropriate person to undertake all steps themselves. • Broken arrows - potential deviations, e.g. due to the required skills/level of care, or preferences of the person with diabetes.

Understand the Target Population: Stakeholder Consultation

The ERG consultations facilitated greater understanding of the roles of different health professional disciplines and their needs. A summary of the key results is included in Table 1. The consultations highlighted the diverse needs (eg, learning styles and scope of practice) and preferences of the ERG members. While some preferred the detail of the chapter, others preferred the brevity of the summaries. The 5 A's model was the preferred summary format. The ERG suggested to develop complementary factsheets for people with diabetes about the emotional problems in the *handbook*.

Table 1. Results and actions arising from stakeholder (Expert Reference Group) consultations

Topic	Results	Actions
Assessment of psychological health in clinical practice settings	Each practicing health professional had a different preferred style, for example: • Enquiring through conversation and open-ended questions, because they believe a structured questionnaire might divert the focus of the conversation from the agenda of the person with diabetes to the agenda of the health professional. • Routine screening using short questionnaires, because they believed it is easy to miss problem areas in conversation, eg, the person with diabetes might not raise it themselves. They believed that asking people to complete the screening questionnaires in the waiting room (before the consultation) is appropriate. They agreed that introduction of the questionnaire to the person with diabetes is important (eg, to explain the purpose and that completion is optional). • Annual assessment (eg, for diabetes distress using the PAID[a] scale), because it forms part of an annual holistic approach to care and is acceptable to people with diabetes.	• As there is no evidence to support one approach over another, the *handbook* caters to these different styles. It provides examples of different ways to incorporate psychological screening into routine clinical practice.
Feedback about the *long version* (sample chapter) and *summary versions* (ie, sample 5 A's model, flowchart, and factsheet)	*Long version (sample chapter)*: • Some thought the amount of detail (length) in the example chapter was appropriate and that there was a good balance of bullet points and sentences. • They believed that some health professionals would want this level of detail, but that others may not have time to read it. So, it would be important to find a balance between detailed information and 1-page summaries. • They offered various suggestions for presentation (eg, develop electronic and printed versions, use a PDF rather than CD for the electronic version). *Summary versions*: • Of the 3 short formats provided, the 5 A's model was preferred because it was a simple, step-by-step guide and would be familiar to many health professionals. They made minor suggestions for improvement (eg, reducing text and design elements). • Summary cards would be useful to put on wall or in top drawer (for quick accesses). • They suggested and agreed that factsheets for people with diabetes would be useful (to facilitate conversations to distribute in consultations and waiting rooms).	• The *handbook* contains detailed information, while the *toolkit* contains double-sided summaries, copies of validated questionnaires, and factsheets for adults with diabetes. It caters to all expressed needs and preferences. • The 5 A's model (later adapted to a 7 A's model, see Figure 3) is a key element of the *handbook* and summary cards (consistent structure). • Electronic (PDF) and hardcopy versions of the *handbook* and *toolkit* are available (from ndss.com.au).
Topic-specific feedback	• They offered suggestions regarding topics to include and exclude and questionnaires to include and exclude.	• We considered and included the suggestions as appropriate (eg, within scope of project and supporting evidence).
Language considerations	• Use plain language in communications with and for people with diabetes. • Define commonly used words and use terms consistently. • Avoid referring to "patients."	• We made efforts to ensure appropriate and consistent language use, and in accordance with published recommendations [35]. • The *handbook* includes a Glossary of terms. • The *handbook* and *toolkit* were professionally copyedited and proofread. For the factsheets, this included plain language editing.
Ideas for future stakeholder consultation	• Include people with diabetes in the next phase of consultation. • A suggestion was offered for implementing stakeholder consultation (based on a process which worked well for another organization when developing factsheets)—to prepare unformatted drafts and email it to stakeholders (eg, professional bodies and consumers), for written responses or tracked changes. Give a few guiding questions in the cover letter then leave it "open".	• We adapted (to suit the project) and implemented the suggested consultation method. • Health professionals, academic experts, and people with diabetes reviewed the *handbook* and *toolkit* (see "Phase 1: Review by Professionals and People with Diabetes" in the Methods and Results).

Topic	Results	Actions
Suggestions for dissemination and future work	• Consider attaining endorsement of the *handbook* by relevant professional bodies. • Consider developing training to complement the *handbook* and offer professional development points.	• We did not pursue endorsements, owing to complexity of the process and because the *handbook* and *toolkit* would adopt the strong brand and endorsement of the funding body. • We informed relevant professional bodies of the resources (enabling promotion to their members). • We pursued funding opportunities to enable development of Web-based health professional training.

^aPAID: Problem Areas In Diabetes.

On the basis of the consultation results, we decided to develop both a *handbook* and *toolkit* to meet the diverse needs of the various health professionals involved in routine diabetes care. The *handbook* would provide detailed information, with each chapter using a consistent format to enable quick reference (eg, consistent headings and structure, guided by the 5 A's model, which was the preferred *summary* format) [31]. The *toolkit* would provide practical resources for use in clinical consultations, that is, 1-page summary cards of each chapter (focused on the 5 A's model), copies of validated questionnaires, and factsheets for people with diabetes.

Drafting the *handbook* and *toolkit* was a collaborative process. We divided the topics among the individual authors (CH, JH, LB, and JS) to lead preparation. For some chapters, the writing was shared (eg, by multiple authors or with a contributor from outside the core authorship team). Overall, a minimum of 3 authors contributed to each draft chapter. We drew upon clinical practice guidelines, published peer-reviewed literature, the findings of the ERG consultation, our own clinical and research experience, and iterative ongoing discussions (with the ERG and team). We held writing workshops regularly to discuss writing progress and content and to make adjustments for consistency between chapters.

Pretest Intervention Methods and Materials: Phases 1–3

Phase 1: Review by Professionals and People With Diabetes

The reviewers (n=37) provided positive and constructive feedback. Many suggestions for improvement were minor and most was chapter specific (eg, add a reference, adapt a suggested strategy, or change a word). Several reviewer suggestions were outside the project scope (eg, include information about complex psychiatric conditions or for carers of people with diabetes) or included elsewhere in the *handbook* (eg, information about cultural and linguistic diversity). Some suggestions conflicted with those of other reviewers; this diversity reflected the multidisciplinary nature of the content and review, and individuals' views and areas of interest. A common criticism was the long chapter length, but there was no consensus among the reviewers about what information to remove; they considered all content important and had differing opinions about the most valuable content.

We held several whole-day team workshops to compare, discuss, and review the feedback. This included discussion of conflicting feedback and views. To overcome conflict, we discussed and collaboratively made pragmatic decisions (eg, could we add a text box or bullet point, or refer the reader to another page or chapter?). We sought clarification from reviewers as needed (eg, we wanted clarification or further information). Significant revisions were required for 2 chapters: "Communication" and "Facing life with diabetes." Finally, the author who drafted the chapter implemented the agreed revisions.

Where relevant, we implemented feedback across multiple chapters of the *handbook and toolkit*. Examples include the following:

- Structured case studies—we had developed 2 versions of the first 2 chapters reviewed ("Psychological barriers to insulin use" and "Depression"), one "structured" in accordance with the 7 A's model and one "unstructured." The reviewers indicated that the structured version was preferable, so we structured all other case studies in this manner.
- Editing—we developed a style guide to ensure consistency across the handbook in reducing unnecessary wordiness, improving readability or understanding, and improving the language (to be more positive, empowering, inclusive, supportive, and consistent).
- Quotes from people with diabetes—we added these to demonstrate examples of concepts in the handbook.

Phase 2: Review by Funding Body

The MESAC viewed the *handbook and toolkit* favorably. The feedback was mostly minor and chapter specific (eg, rephrasing a sentence or changing a word), reflecting the rigorous development and review process. MESAC also commented on the *handbook* chapter length but could not make specific suggestions for shortening them.

We implemented the requested changes as appropriate (eg, rewriting some sections to reduce wordiness), resulting in a *final version* (described in Figure 2). Some comments were outside the scope of the project and could not be included in the *handbook* and *toolkit*. For example, suggestions to include information for "carers" (ie, family and friends of people with diabetes) were not implemented, as it would have required different information and consultations. At least three authors (CH, JH, and JS) reviewed this final version. We documented the changes made and provided the MESAC with a written

response and copy of the final version. The MESAC approved the final version.

Phase 3: Interviews With Health Professionals

We interviewed 19 participants, of 25 health professionals who volunteered. The participants included 9 nurses and 6 dietitians (of whom 7 and 1 were credentialed diabetes educators, respectively), 2 general practitioners, and 2 endocrinologists. The participants worked in urban or metropolitan (9/19, 47%) and regional or rural (10/19, 53%) settings. Most (16/19, 84%) worked in a multidisciplinary health service that did not include a mental health professional in the team. They reported varying levels of confidence to talk about (median 3, range 2-5) and assist with (median 3, range 2-5) with diabetes-related emotional problems; 37% (7/19) had used a questionnaire in clinical care to assess emotional health. The 6 who did not participate could not be reached at the time of interview. They had similar confidence to talk about and assist with emotional problems, compared with the participants, but none had used questionnaires.

The chapters selected by participants were fear of hypoglycemia (n=7), diabetes distress (n=5), eating problems (n=4), psychological barriers to insulin use (n=1), depression (n=1), and anxiety disorders (n=1). Most also read other chapters of the *handbook*.

Participant quotes are included in Table 2. Overall, the *handbook* and *toolkit* were viewed favorably as well-written, easy to read and understand, and easy to navigate, with consistent chapter structure and good design elements (eg, use of colors, fonts, and boxes). Perceived *highlights* were practical elements, such as the examples of open-ended questions, validated questionnaires, case studies, 7 A's model, and summary cards. For example, 1 nurse practitioner described how she valued the summary cards in the *toolkit*, "Loved the summary cards and felt that this was the handbook's biggest strength." Participants gave very few suggestions for improvement; the most common criticism was the chapter length, but they were unable to offer suggestions to shorten it and they appreciated its comprehensiveness.

The participants reported how the *handbook* raised their awareness about the role of health professionals in attending to the psychological aspects of diabetes, which encouraged *self-reflection* on their practice. This is demonstrated well by a Nurse Practitioner-Credentialed Diabetes Educator who said, "I think it's a really good way to actually reflect back on...your own thoughts and about how you actually engage in asking about diabetes distress." The *handbook* also helped to build *self-confidence*, affirming for some that their current clinical practice was *on track*. For example, a dietitian found validation

in the *handbook's* message that health professionals are often best placed to provide support for diabetes distress, "And, the other comment that's interesting is that diabetes distress is best managed within the context of diabetes care...[it] improves my confidence and I think 'yeah – that's what I do with people.'"

The participants commented how the *handbook* could *influence clinical practice*—some had already implemented aspects of the *handbook* (eg, asking more frequently about how the person with diabetes feels), whereas others planned to make changes (eg, implementing routine assessment for diabetes distress and trying new strategies for observed problems). For instance, 1 Credentialed Diabetes Educator-Registered Nurse gave an example of how she asked about a person's well-being after reading the *handbook* and subsequently provided a listening ear and psychology referral, "I was reading the handbook before the patient came in two days ago and I said to her, 'How are you going? How are you feeling about everything? Are you managing things?'...I realized it was a real can of worms...I did refer her on..."

The participants identified some possible *barriers to implementation*, including a lack of referral options (to mental health professionals, particularly with diabetes expertise) and costs of intensified follow-up (eg, calls or text messages, when needed). For example, an endocrinologist commented "I think the problem with the psychology referral is that...people with a specific interest in diabetes are few and far between." Notably, the participants felt empowered not only to implement the resources but also to *spread the word*, by recommending it to other health professionals or even training others to use it. For example, one dietitian explained, "I've talked to my boss and I wouldn't mind doing a bit of an information session to other dietitians and coaches about it." They also offered suggestions for future work (eg, *promotion* and *training*), such as this from a Credentialed Diabetes Educator-Registered Nurse, "I think it needs to go into the Graduate Certificate for Diabetes Education. I think it needs to go to medical students and...pharmacy students as well..." Some felt that the *handbook* was sufficient as a stand-alone resource to supplement participants' existing professional skills and experience. Others considered that training would be useful to help build confidence or to enhance specific skills (eg, introducing, scoring, and responding about questionnaires).

Given the positive feedback and lack of major concerns, CH, JH, and JS finalized the content and typesetting. Electronic and hardcopy versions of the *Diabetes and Emotional Health* handbook and toolkit were published in print and on the Web [23]. Promotion of the resources commenced in August 2016, at an Australian diabetes conference, and is ongoing.

Table 2. Qualitative study evaluation of the handbook and toolkit by health professionals.

Topic	Example quotes
Perceived highlights	• "Most useful section was the ASK section as it gave practical tips on asking questions that will bring up any issues." (Dietitian, male, eating problems) • "I really liked the quotes from practice nurses and quotes from health professionals and quotes from consumers. I thought that was a nice value add. And, I liked the 'ABCs of effective communication.'" (Dietitian, female, diabetes distress) • "Loved the summary cards and felt that this was the *handbook*'s biggest strength." (Nurse practitioner, female, depression) • "Irena [case study]…that would be a very common scenario…I thought it was a great example, and it's brief, it's not too lengthy. And then it sets out the 7 A's framework in each section of what you can do…. so it's easy for the practitioner to work with that." (Registered nurse-diabetes educator, female, fear of hypoglycemia) • "I'm quite visual in my learning, so, when I'm trying to think about structures in my head, I think visually and I like colors… I like the approach of it." (Registered nurse, female, diabetes distress) • "The structure is well laid out and it certainly follows the 7 A's model very well. All the chapters have done that, which is exactly what you want. You want consistency so that the reading is streamlined and everyone who reads it can access what they want very quickly." (GP, male, fear of hypoglycemia)
Role of health professionals	• "I think that the one thing that the *handbook* says… that comes out in every chapter, at every appointment, ask about their well-being. Don't just assume they're okay …ask the question." (Credentialed diabetes educator-registered nurse, female, fear of hypoglycemia) • "For me, this book has really normalized that our role is working with a whole person, their emotional health and their physical health… sometimes in the rush of everything, and especially in a tertiary setting, you can get quite focused on the physical aspect of somebody who's acutely unwell… I found it a really useful tool to just ensure that we all keep in mind that we're working with a person." (Registered nurse, female, diabetes distress) • "I think the point that kept coming through… most people actually do want to go to a member of their diabetes team to talk about this. So, if my question was, you know, is this my role? Well, yeah, it is. They're seeing it as my role." (Dietitian, female, diabetes distress)
Encourage self-reflection	• "I think it's a really good way to actually reflect back on… your own thoughts and about how you actually engage in asking about diabetes distress." (Nurse practitioner-credentialed diabetes educator, female, diabetes distress) • "I used the Diabetes Distress chapter more for like, a concrete reflection on my practice." (Registered nurse, female, diabetes distress) • "I discussed it with a Type 1 client and we've made an arrangement that next time that she has an appointment that we will discuss her views on this, because it's something that I feel that I could do better on it as a clinician." (Credentialed diabetes educator-registered nurse, female, fear of hypoglycemia)
Build self-confidence	• "And, the other comment that's interesting is that diabetes distress is best managed within the context of diabetes care… [it] improves my confidence and I think 'yeah – that's what I do with people' and just talking through their feelings and what's an issue for them – that I'm doing the right thing… Which is such a feeling of relief really." (Dietitian, female, diabetes distress) • "That 7A's thing, it really worked. I was nicely surprised. Yeah. And, it felt like I didn't do anything different to what I normally do, but, I had a language for it. So, that I could work it through for myself." (Registered nurse, female, diabetes distress)
Influence clinical practice	• "I was reading the *handbook* before the patient came in two days ago and I said to her, 'How are you going? How are you feeling about everything? Are you managing things?...I realized it was a real can of worms … it was not just about her diabetes. It was about feeling overwhelmed…She actually ended up saying to me, 'The only reason why this conversation's coming out is because you asked me how I am. Had you not asked me, I wouldn't tell you that I'm getting depressed… I am noticing that I am mood swinging and I don't know whether it's my stage of life, and problems which are hormonal, or whether it's depression'...In that situation I did refer her on because I feel that she's been missed..." (Credentialed diabetes educator-registered nurse, female, fear of hypoglycemia) • "The 7A's tool itself, in working with this client – it let me look at structuring how to work both with … the physical aspects and the emotional responses together. It was really useful." (Registered nurse, female, diabetes distress) • "I really think that part of my practice I need to do more of regularly is start to use the PAID scale and particularly for everyone, not just picking some people but just doing it on everyone, not all the time, but doing it. But, I really think it will actually bring out a lot of things that both sides of the party didn't realize or think about." (Nurse practitioner-credentialed diabetes educator, female, diabetes distress) • "I can think of a number of people I see where I've recognized fear of hypoglycemia and there's information within that chapter which may help me work through that fear with them." (Endocrinologist, male, fear of hypoglycemia)
Spread the word	• "I actually showed one of my neighbors who is a pharmacist and I was discussing with him, on chapter 5, how good I thought it was to do decisional balancing with someone who is reluctant to start insulin." (Credentialed diabetes educator-registered nurse, female, fear of hypoglycemia) • "I'll be actively promoting it as a resource and I'll be wanting our GPs to have it available to them and I'll be promoting it heavily." (Credentialed diabetes educator-registered nurse, female, eating problems) • "I've talked to my boss and I wouldn't mind doing a bit of an information session to other dietitians and coaches about it… summarize the importance of dealing with the emotional aspect of diabetes and then point to some of these summary cards and… the *handbook*." (Dietitian, female, diabetes distress)

Topic	Example quotes
Barriers to implementation	• "I think the problem with the psychology referral is that… people with a specific interest in diabetes are few and far between… But, it would be really nice to know which psychologists for instance, in our area, are interested in dealing with patients with diabetes and the same with psychiatrists. I think that's where I struggle clinically." (Endocrinologist, male, fear of hypoglycemia) • "You know, unless you are lucky enough to have a psychologist that works and is employed within a practice and they don't charge a gap, people can't afford that." (Nurse practitioner-credentialed diabetes educator, female, diabetes distress) • "Regarding diabetes distress, it's [*handbook*] encouraging us as health professionals to be able to get on the telephone to follow up, or to arrange all this follow up, but that's not remunerated under Medicare… Under the Medicare system and under the care planning arrangements that we have, I might have one visit allocated, so it doesn't give me necessarily a follow-up visit without the patient then being out of pocket..." (Credentialed diabetes educator-registered nurse, female, fear of hypoglycemia) • "We have 20 minute appointments and have to deal with the medical side of diabetes and then to deal with psychological side too, we probably can designate maybe five minutes – ten minutes, if we're lucky… we need the numbers and the resources and things at our fingertips… I think it's fantastic and it definitely is needed but we need to have that information right there or in a website form where we can just go click." (Endocrinologist, female, fear of hypoglycemia)
Suggestions for promotion	• "Getting it out there, is the key… a good advertising campaign, and that's probably best orchestrated through presentation of some of the work in the book at the various diabetes-related clinical meetings that occur around the countryside..." (Endocrinologist, female, fear of hypoglycemia) • "I think it needs to go into the Graduate Certificate for Diabetes Education. I think it needs to go to medical students and…pharmacy students as well…" (Credentialed diabetes educator-registered nurse, female, fear of hypoglycemia)
Training needs and ideas	• "… actually getting people to physically do the questionnaire themselves or give them a case study scenario so that they're doing it from the perspective of their person… it's really important to be familiar with the tool. But, also scoring. So, it's one thing to give people a questionnaire to fill out but to then be able to score it on the spot and give them some feedback." (Dietitian, female, diabetes distress) • "…as part of the training something along the lines of a mentoring or relationship...you might, as part of the training process, bring along how you do things, case studies, that part of it is meeting and talking with a peer." (Nurse practitioner-credentialed diabetes educator, female, diabetes distress) • "...I think they [chapters] could all be incorporated into one workshop, covering how to ask questions and what to do with the information once you've got it and how to build trust, all that sort of thing." (Nurse-diabetes educator, female, eating problems) • "…practical case studies based. The other thing that works well for medical practitioners is webinars and well, something where people can be at home listening to or do it in their own time or, perhaps type questions in." (Endocrinologist, female, fear of hypoglycemia) • "We need to make the assessment of emotional health a compulsory part of guidelines… because they all tick off these other things on their list. We have to test their cholesterol levels twice a year. We're going to do their HbA_{1c} twice a year… but no one actually asks them about how they are." (Dietitian, female, eating problems)

Discussion

Principal Findings

While guidelines recommending psychological care in diabetes have existed for 25 years [5], to our knowledge, the *Diabetes and Emotional Health handbook* and *toolkit* [23] represents the first attempt to develop evidence-based, clinically informed, freely available, practical resources for multidisciplinary diabetes health professionals supporting the emotional health of adults with type 1 and type 2 diabetes. These resources are a tailored response to the expressed unmet needs of health professionals, who cite lack of resources and confidence to address diabetes-related emotional problems as significant barriers to providing holistic diabetes care. The *handbook* offers strategies and tools for recognizing psychological problems and providing support for them. The *toolkit* contains practical resources to facilitate implementation: chapter summary cards, questionnaires, and factsheets for people with diabetes. The *handbook* and summary cards implement the 7 A's model. A model such as this provides a memorable acronym for application in busy health settings and is consistent with the expectations of people with diabetes about support from health

professionals [34]. The 7 A's model is a useful framework to provide a consistent and logical structure with a clear path to implementation in clinical practice. The reviewers and qualitative study participants favorably viewed its application in the *handbook*.

Formative evaluation is an essential first step for developing high quality and effective interventions that are acceptable to the target population [24]. In this case, formative evaluation helped us to comprehensively explore the problem, while ensuring accountability and quality control. The formative evaluation approach is a key strength of the resources. The inclusion of several stages of end-user (health professionals) and stakeholder (eg, academic experts and people with diabetes) consultation means that we are confident that the resources align with expressed needs and published evidence. Given our combined expertise, we could have developed these resources with less consultation, which would have been less resource and time intensive. However, we would have been less confident with the final product and the rigorous review process was well received by and inspired confidence among stakeholders and potential users [24]. The methods described in this study demonstrate how formative evaluation can inform the

development of high quality, evidence-based resources, and the processes described herein may be valuable for informing the development of similar resources in other areas.

These resources are important stepping stones toward more consistent implementation of clinical practice guidelines and better integration of psychological health into routine diabetes consultations. As described by 1 qualitative study participant, the *handbook* has "really normalized that our role is working with a whole person, their emotional health and their physical health." When we commenced this project, there was no specific Australian guideline for the psychological care of people with diabetes. In addition to developing *Diabetes and Emotional Health*, we concurrently advocated for recognition of psychological care within existing Australian guidelines to align them more closely with international recommendations. Since then, recommendations for routine screening for depression and diabetes distress have been included in the Australian General Practice Management of Type 2 Diabetes guidelines [7,9]; the Australian guidelines for type 1 diabetes [36] are yet to be revised. Importantly, mental health has been recognized in the Australian Government's National Diabetes Strategy: 2016-2020 [37]. Moving forward, we continue promoting these resources via seminars, workshops, conferences, and social media. These resources are part of the training curriculum for the next generation of credentialed diabetes educators (eg, at Flinders University and Deakin University). We have developed, and are currently evaluating, a Diabetes Distress e-Training for health professionals, on the basis of the *handbook* and *toolkit* content. Additionally, we have collaborated internationally on a Diabetes UK adaptation of the *handbook* and *toolkit* to suit the UK health care context, which is now available on the Web [38]. We have also been able to address the resource gap raised by the MESAC (see "Phase 2: Review by Funding Body" in the Results) by subsequently developing a factsheet for family and friends of people with diabetes [39].

We acknowledge that there is more work to do. For example, a limitation of this study was the exclusive focus on adults with type 1 and type 2 diabetes. We selected these 2 groups as they represent the largest populations in need of psychological support, and our previous research had focused largely on adults [25-28]. However, there is a need for similar resources to enhance support for children and adolescents with diabetes (and their families) and women with gestational diabetes. Furthermore, the diabetes-related psychological needs of specific subgroups (such as Aboriginal and Torres Strait Islanders, culturally and linguistically diverse communities, people with psychiatric conditions, and people with disabilities) are under-researched and important areas for future work. Evaluation of the process and impact of the real-world implementation of the resources would also be valuable in the future [24] but was not within our project scope.

Conclusions

Diabetes and Emotional Health is a practical, evidence-based, clinically informed handbook and toolkit developed in consultation with end-users and other stakeholders. The findings of our formative evaluation suggest that the resources are comprehensive yet user friendly, addressing the previously unmet needs of multidisciplinary health professionals, enabling professional development and supporting real-world implementation of clinical practice guidelines related to the psychological care of people with diabetes.

Practice Implications

Diabetes and Emotional Health provides health professionals with practical information and tools required to implement clinical practice guidelines related to the psychological aspects of diabetes. More than 1000 hardcopies have been distributed to Australian health professionals and more than 1400 electronic copies have been downloaded. The resources remain freely available on the Web. [23] Health professionals find the handbook and toolkit useful for their clinical practice, are implementing them, and are taking ownership of them (eg, discussing with others and making plans to train others). These resources are likely to have clinical utility internationally (until international adaptations are developed), owing to the evidence-based content, robust stakeholder review, and shared goals with international clinical practice guidelines (holistic and person-centered care and attention to psychological problems). Similarly, our adapted 7 A's model may have clinical utility for routine screening and monitoring for other problems (psychological or other) as part of a person-centered approach to routine care.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Literature review questions, results, lessons, and actions.
[DOCX File , 30 KB - formative_v4i2e15007_app1.docx]

References

1. International Diabetes Federation (IDF). IDF Diabetes Atlas. Brussels: IDF; 2015. URL: https://www.idf.org/e-library/epidemiology-research/diabetes-atlas/13-diabetes-atlas-seventh-edition.html [accessed 2019-11-26]

2. Speight J. Managing diabetes and preventing complications: what makes the difference? Med J Aust 2013 Jan 21;198(1):16-17. [doi: 10.5694/mja12.11489] [Medline: 23330749]

3. Funnell MM, Anderson RM. Empowerment and Self-Management of Diabetes. Clin Diabetes 2004;22(3):123-127. [doi: 10.2337/diaclin.22.3.123]

4. Ventura AD, Browne JL, Holmes-Truscott E, Hendrieckx C, Pouwer F, Speight J. Diabetes MILES–2 2016 Survey Report. Melbourne: Diabetes Victoria; 2016. URL: https://acbrd.org.au/wp-content/uploads/2018/06/C292_DV_miles-2_web_final.pdf [accessed 2019-11-26]

5. Bradley C, Gamsu D. Guidelines for encouraging psychological well-being: report of a Working Group of the World Health Organization Regional Office for Europe and International Diabetes Federation European Region St Vincent Declaration Action Programme for Diabetes. Diabet Med 1994;11(5):510-516. [doi: 10.1111/j.1464-5491.1994.tb00316.x] [Medline: 8088133]

6. American Diabetes Association (ADA). (4) Foundations of care: education, nutrition, physical activity, smoking cessation, psychosocial care, and immunization. Diabetes Care 2015 Jan(38 Suppl):S20-S30. [doi: 10.2337/dc15-S007] [Medline: 25537702]

7. Royal Australian College of General Practitioners (RACGP). General Practice Management of Type 2 Diabetes: 2016-18. East Melbourne: RACGP; 2016. URL: https://www.racgp.org.au/clinical-resources/clinical-guidelines/key-racgp-guidelines/view-all-racgp-guidelines/management-of-type-2-diabetes [accessed 2019-11-26]

8. Robinson DJ, Luthra M, Vallis M, Canadian Diabetes Association. Diabetes and mental health. Can J Diabetes 2013 Apr;37(Suppl 1):S87-S92. [doi: 10.1016/j.jcjd.2013.01.026] [Medline: 24070971]

9. Royal Australian College of General Practitioners (RACGP). General Practice Management of Type 2 Diabetes 2014-15. East Melbourne: RACGP; 2014. URL: https://www.racgp.org.au/download/Documents/Guidelines/Diabetes/2014diabetesmanagement.pdf [accessed 2019-11-26]

10. Li C, Ford ES, Zhao G, Balluz LS, Berry JT, Mokdad AH. Undertreatment of mental health problems in adults with diagnosed diabetes and serious psychological distress: the behavioral risk factor surveillance system, 2007. Diabetes Care 2010 May;33(5):1061-1064 [FREE Full text] [doi: 10.2337/dc09-1515] [Medline: 20185747]

11. Pouwer F, Beekman AT, Lubach C, Snoek FJ. Nurses' recognition and registration of depression, anxiety and diabetes-specific emotional problems in outpatients with diabetes mellitus. Patient Educ Couns 2006 Feb;60(2):235-240. [doi: 10.1016/j.pec.2005.01.009] [Medline: 16442465]

12. Snoek FJ, Kersch NY, Eldrup E, Harman-Boehm I, Hermanns N, Kokoszka A, et al. Monitoring of Individual Needs in Diabetes (MIND): baseline data from the Cross-National Diabetes Attitudes, Wishes, and Needs (DAWN) MIND study. Diabetes Care 2011 Mar;34(3):601-603 [FREE Full text] [doi: 10.2337/dc10-1552] [Medline: 21266654]

13. Halliday JA, Hendrieckx C, Beeney L, Speight J. Prioritization of psychological well-being in the care of diabetes: moving beyond excuses, bringing solutions. Diabet Med 2015 Oct;32(10):1393-1394. [doi: 10.1111/dme.12768] [Medline: 25819747]

14. Butkiewicz E, Carbonem A, Green S, Miano A, Wegner E. Diabetes management. In: Burg MA, Oyama O, editors. The Behavioral Health Specialist in Primary Care: Skills for Integrated Practice. New York: Springer Publishing Company; 2016:51-69.

15. Barnard KD, Lloyd CE, Holt RI. Psychological burden of diabetes: what it means to people with diabetes. In: Barnard KD, Lloyd CE, editors. Psychology and Diabetes Care: A Practical Guide. London: Springer; 2012:1-22.

16. Lloyd CE, Pouwer F, Hermanns H. Screening for Depression and Other Psychological Problems in Diabetes. London: Springer; 2013.

17. Katon W, Maj M, Sartorius N. Depression and Diabetes. West Sussex: Wiley-Blackwell; 2010.

18. Conn JJ, Silberberg CL, Hendrieckx C, Nankervis A, Cheung N, Jenkins A. Enhancing Your Consulting Skills - Supporting Self-management and Optimising Mental Health in People With Type 1 Diabetes. Canberra: National Diabetes Services Scheme; 2014.

19. Mosely K, Aslam A, Speight J. Overcoming barriers to diabetes care: Perceived communication issues of healthcare professionals attending a pilot Diabetes UK training programme. Diabetes Res Clin Pract 2010 Feb;87(2):e11-e14. [doi: 10.1016/j.diabres.2009.12.003] [Medline: 20044163]

20. Holt RIG, Nicolucci A, Burns KK, Lucisano G, Skovlund SE, Forbes A, et al. Correlates of psychological care strategies for people with diabetes in the second Diabetes Attitudes, Wishes and Needs (DAWN2) study. Diabet Med 2016 Sep;33(9):1174-1183. [doi: 10.1111/dme.13109] [Medline: 26939906]

XSL•FO
RenderX

21. Byrne JL, Davies MJ, Willaing I, Holt RIG, Carey ME, Daly H, et al. Deficiencies in postgraduate training for healthcare professionals who provide diabetes education and support: results from the Diabetes Attitudes, Wishes and Needs (DAWN2) study. Diabet Med 2017 Aug;34(8):1074-1083. [doi: 10.1111/dme.13334] [Medline: 28195662]

22. Joensen L, Fisher L, Skinner T, Doherty Y, Willaing I. Integrating psychosocial support into routine diabetes care: perspectives from participants at the Self-Management Alliance meeting 2016. Diabet Med 2019 Jul;36(7):847-853. [doi: 10.1111/dme.13836] [Medline: 30315608]

23. Hendrieckx C, Halliday J, Beeney L, Speight J. Diabetes and Emotional Health: A Handbook for Health Professionals Supporting Adults with Type 1 or Type 2 Diabetes. Canberra: National Diabetes Services Scheme (NDSS); 2016. URL: https://www.ndss.com.au/about-diabetes/resources/find-a-resource/diabetes-and-emotional-health/ [accessed 2019-11-27]

24. Bauman A, Nutbeam D. Evaluation in a Nutshell: A Practical Guide to the Evaluation of Health Promotion Programs. Second Edition. North Ryde: McGraw-Hill Education; 2014.

25. Speight J, Browne JL, Holmes-Truscott E, Hendrieckx C, Pouwer F, On behalf of the Diabetes MILES–Australia reference group. Diabetes MILES–Australia 2011 Survey Report. Canberra: National Diabetes Services Scheme (NDSS); 2011. URL: https://www.ndss.com.au/wp-content/uploads/resources/report-miles-youth-2011.pdf [accessed 2019-11-26]

26. Speight J, Browne JL, Holmes-Truscott E, Hendrieckx C, Pouwer F. Diabetes MILES–Australia (Management and Impact for Long-term Empowerment and Success): methods and sample characteristics of a national survey of the psychological aspects of living with type 1 or type 2 diabetes in Australian adults. BMC Public Health 2012;12:120 [FREE Full text] [doi: 10.1186/1471-2458-12-120] [Medline: 22325032]

27. Hendrieckx C, Halliday JA, Bowden JP, Colman PG, Cohen N, Jenkins A, et al. Severe hypoglycaemia and its association with psychological well-being in Australian adults with type 1 diabetes attending specialist tertiary clinics. Diabetes Res Clin Pract 2014;103(3):430-436. [doi: 10.1016/j.diabres.2013.12.005] [Medline: 24513121]

28. Hendrieckx C, Bowden J, Halliday J, Colman P, Cohen N, Jenkins A, et al. An Audit of Psychological Well-being in Adults With Type 1 Diabetes. 2012 Presented at: Australian Diabetes Society and Australian Diabetes Educators Society Annual Scientific Meeting; August 29-31, 2012; Brisbane, Australia.

29. American Diabetes Association (ADA). Standards of medical care in diabetes–2008. Diabetes Care 2008;31(Suppl 1):S12-S54. [doi: 10.2337/dc08-S012] [Medline: 18165335]

30. Nicholson T, Taylor J, Gosden C, Trigwell P, Ismail K. National guidelines for psychological care in diabetes: how mindful have we been? Diabet Med 2009 Apr;26(4):447-450. [doi: 10.1111/j.1464-5491.2009.02701.x] [Medline: 19388977]

31. Royal Australian College of General Practitioners (RACGP). Smoking, Nutrition, Alcohol and Physical Activity (SNAP): A Population Health Guide to Behavioural Risk Factors in General Practice. South Melbourne, Australia: RACGP; 2004.

32. Glasgow RE, Emont S, Miller DC. Assessing delivery of the five 'As' for patient-centered counseling. Health Promot Int 2006 Sep;21(3):245-255. [doi: 10.1093/heapro/dal017] [Medline: 16751630]

33. Glynn TJ, Manley MW. How to Help your Patients Stop Smoking: A Manual for Physicians. Bethesda, MD: National Institutes of Health; 1989. HathiTrust Digital Library URL: https://catalog.hathitrust.org/Record/009613758 [accessed 2019-11-28]

34. Lawn S, Schoo A. Supporting self-management of chronic health conditions: common approaches. Patient Educ Couns 2010 Aug;80(2):205-211 [FREE Full text] [doi: 10.1016/j.pec.2009.10.006] [Medline: 19931372]

35. Speight J, Conn J, Dunning T, Skinner TC, Diabetes Australia. Diabetes Australia position statement. A new language for diabetes: improving communications with and about people with diabetes. Diabetes Res Clin Pract 2012 Sep;97(3):425-431. [doi: 10.1016/j.diabres.2012.03.015] [Medline: 22513346]

36. Craig ME, Twigg SM, Donaghue KC, Cheung NW, Cameron FJ, Conn J, et al. National Evidence-Based Clinical Care Guidelines for Type 1 Diabetes in Children, Adolescents and Adults. Canberra: Australian Government Department of Health and Ageing; 2011.

37. Australian Government Department of Health. Australian National Diabetes Strategy: 2016-2020. Canberra: Commonwealth of Australia; 2015.

38. Hendrieckx C, Halliday J, Beeney L, Speight J. Diabetes and Emotional Health: A Practical Guide for Healthcare Professionals Supporting Adults with Type 1 or Type 2 Diabetes. London: Diabetes UK: National Diabetes Services Scheme (NDSS); 2019. URL: https://www.diabetes.org.uk/professionals/resources/shared-practice/psychological-care/emotional-health-professionals-guide

39. National Diabetes Services Scheme (NDSS). 2018. Caring for Someone With Diabetes (for Family and Friends) URL: https://www.ndss.com.au/about-diabetes/resources/find-a-resource/caring-for-someone-with-diabetes-for-family-and-friends-fact-sheet/ [accessed 2019-11-28]

Abbreviations

ERG: Expert Reference Group
MESAC: Medical, Education, and Scientific Advisory Council
NDSS: National Diabetes Services Scheme