JMIR Formative Research

Impact Factor (2022): 2.2
Volume 4 (2020), Issue 4 ISSN 2561-326X Editor in Chief: Gunther Eysenbach, MD, MPH, FACMI

Contents

Original Papers

Viewpoint

Original Paper

Evaluation of an Educational Health Website on Infections and Antibiotics in England: Mixed Methods, User-Centered Approach

Rosalie Allison[1], BSc, MSc; Catherine Hayes[1], BSc, MPH; Vicki Young[1], BSc, PhD; Cliodna A M McNulty[1], MBBS, FRCP

Public Health England, Gloucester, United Kingdom

Corresponding Author:
Rosalie Allison, BSc, MSc
Public Health England
Twyver House
Gloucester, GL1 1DQ
United Kingdom
Phone: 44 208 495 3258
Email: rosie.allison@phe.gov.uk

Abstract

Background: e-Bug, an educational health website for teachers and students, aims to help control antibiotic resistance by educating young people about microbes, hygiene, and antibiotic resistance, reducing the incidence of infection and, therefore, the need for antibiotics. The teachers' section of the e-Bug website has not been evaluated since it was launched in 2009, and worldwide page views have been steadily decreasing since 2013.

Objective: This study aimed to apply GoodWeb, a comprehensive framework utilizing methodologies and attributes that are relevant to the digital era, to evaluate and suggest improvements to the e-Bug website.

Methods: Electronic questionnaires and face-to-face completion of task scenarios were used to assess content, ease of use, interactivity, technical adequacy, appearance, effectiveness, efficiency, and learnability of the teachers' section of the e-Bug website.

Results: A total of 106 teachers evaluated the e-Bug website; 97.1% (103/106) of them reported that they would use e-Bug, and 98.1% (104/106) of them reported that they would recommend it to others. Participants thought that there was a niche for e-Bug because of the way the resources fit into the national curriculum. Suggestions for improvements included changing the menu indication by highlighting the current page or deactivating links, improving home page indication, and providing a preview of resources when hovering the mouse over hyperlinks. Additional features requested by users included a search function and access to training opportunities.

Conclusions: This paper reports that the GoodWeb framework was successfully applied to evaluate the e-Bug website, and therefore, it could be used to guide future website evaluations in other fields. Results from this study will be used to appraise the current quality and inform any future changes, modifications, and additions to e-Bug.

(JMIR Form Res 2020;4(4):e14504) doi:10.2196/14504

KEYWORDS

user experience; usability; quality; online; science; health

Introduction

Background

According to Song and Zinkhan [1], an excellent website attracts more web users and encourages revisits if the user's interests are carefully considered and incorporated in the design and presentation. User-centered design means that the websites can both fulfill the goals and desires of its users [2] and influence their perception of the organization and overall quality of resources [3].

There is no universally accepted method or technique for website evaluation; various assessment techniques have been employed to evaluate websites [4]. Both qualitative and quantitative measures are appropriate to evaluate website user experience. Questionnaires are the most widely used method for evaluating websites [5-11]. These can be administered remotely or in person [12] and used stand-alone or in combination with other

methodology, such as observed browsing and interviews following completion of task scenarios [13,14]. Allison et al [7] incorporated methods and attributes from 69 studies to suggest a simple but comprehensive guide to evaluating websites, coined GoodWeb, which follows 4 basic steps:

- *Step 1*: What are the important website attributes that affect a user's experience of the chosen website? For example, appearance, content, interactivity, ease of use, and technical adequacy
- *Step 2*: What is the best way to evaluate these attributes? For example, questionnaire and observed completion of task scenarios
- *Step 3*: Who should evaluate the website? For example, users
- *Step 4*: What setting should be used? For example, face-to-face/controlled and remote

The Website

e-Bug [15] is an ongoing international project, operated by Public Health England (PHE), that creates health education resources for teachers and students, covering the subjects of microbes, hygiene, and antibiotic use and resistance [16]. All activities and plans have been designed to complement the national curriculum, particularly Biology and Personal, Social, Health, and Economic [15]. The National Institute for Health and Care Excellence (NICE) recommended that schools could use e-Bug when teaching about antibiotics and infections [17], and e-Bug is a case study in the UK 5-year Antimicrobial Resistance Strategy 2019-2024 [18]. A key component of the e-Bug project is the e-Bug website, established in September 2009 [15]. The e-Bug website comprises 2 microsites: an educator microsite that includes free teaching resources, such as lesson plans and student worksheets [16], and a student microsite that hosts interactive activities, games, and animations, which are developed by graphic designers, researchers, and microbiologists. The e-Bug resources, in particular, the digital media, including the games, have been well evaluated [19,20]. Worldwide page views of the teachers' section have been steadily decreasing since 2013 (282,284 views in 2013-2014, 248,260 views in 2014-2015, 208,540 views in 2015-2016, and 197,740 views in 2016-2017) [21]. The fact that e-Bug is recognized by both NICE [17] and the Department of Health [18] as a recommended tool for teachers highlights the importance of this evaluation to continually strive to improve this part of the website to aid implementation.

In 2015, hard-copy e-Bug resources were sent to all schools in England [22], and the e-Bug team regularly promoted the resources at popular teacher conferences, such as Big Bang and Association for Science Education (ASE) conferences. The circulation of hard-copy resources may be one explanation for the decrease in page views, but a formal evaluation of the teachers' website may provide other explanations and inform improvements.

Aims

This study aimed to apply GoodWeb, a comprehensive guide, to evaluate an educational health website, using recognized methods to evaluate the most crucial website attributes. The results of the evaluation will be used to appraise the current website quality, inform any future modifications or additions to the website, and the possibility of a full-scale evaluation of the whole site. In addition, the evaluation will be used to advise whether GoodWeb is an effective and appropriate website evaluation guide.

Methods

Study Design

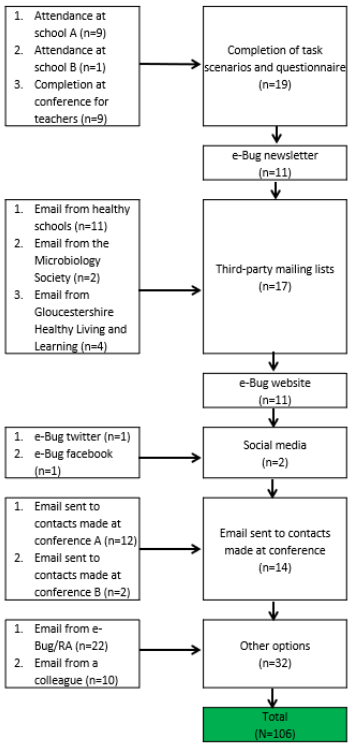
Following GoodWeb, this evaluation was based on the most appropriate existing methodologies and techniques to evaluate websites, determined by a robust review of the current website evaluation literature [7]. The evaluation included a 2-part questionnaire (Multimedia Appendix 1). Part A collected baseline data, whereby users were asked users to rank the importance of different website attributes under the categories appearance, content, interactivity, ease of use, and technical adequacy and then score the e-Bug website against these attributes. Part B of the questionnaire allowed users to rate e-Bug's performance against the attributes ranked in part A. In addition, a subset of users completed website task scenarios to provide more in-depth feedback on the use of the e-Bug website. These data were used to inform prioritization for changes to the e-Bug website.

To assess feasibility and gain suggestions for improvements to the study design, the questionnaire was piloted with 36 users and the task scenarios with 14 users at a 2-day conference. Changes and decisions made based on the pilot and feasibility study include providing the questionnaire in an electronic format rather than a paper-based format, finalizing which website attributes to evaluate, recording the task completion as there was too little time to make comprehensive notes and record mouse clicks at the same time, and providing tick box options to minimize free-text comment boxes.

Recruitment

To increase heterogeneity of participants, for example, to incorporate a range of geographical locations in England, different experiences of e-Bug and to include primary and secondary school teachers, participants were recruited by email from e-Bug contact lists and face-to-face meetings at educational conferences (see Figure 1 for recruitment flowchart).

Figure 1. Recruitment flowchart. RA: Rosalie Allison.



Questionnaire Participant Recruitment

Participants were sent a link to the electronic questionnaire, hosted on SelectSurvey, via teacher mailing lists from e-Bug, healthy schools' leads, At-Bristol Science Centre, Microbiology Society, and Gloucestershire Healthy Living and Learning (GHLL); the e-Bug website's teachers' home page; e-Bug's Twitter and Facebook pages; and teacher contacts that visited the e-Bug advertorial stand at conferences.

Completion of Task Scenarios Recruitment

Teachers visiting the e-Bug stand at 2 conferences (Big Bang 2016 and ASE 2017) were asked whether they would like to participate in an e-Bug website evaluation. Interested teachers were provided with an information sheet, given the opportunity to ask questions, and ensured that they were comfortable with the environment and surroundings, before informed written consent was obtained. Some teachers participated at the conference; others were visited in their own school.

As an incentive to participate in the website evaluation, all participants were offered a £5 (US $6.50) high street gift voucher and a professional certificate for Continuing Professional Development. Participants that completed the additional component of task scenario completion were offered e-Bug resources. Questionnaire participants were entered into a draw to win a set of giant microbes.

Data Collection

Questionnaire Part A: Baseline Questionnaire

Before assessing the e-Bug website, all 106 participants completed an electronic questionnaire (Part A - Multimedia Appendix 1), hosted on SelectSurvey, which asked general questions, including the following:

- Demographic information, such as role, for example, teacher and age groups they teach
- Which teaching resources participants currently use and what devices they use to access resources
- Ranking the 5 main website attributes (appearance, content, interactivity, ease of use, and technical adequacy) in order of importance to themselves as educators. Then within each main attribute, ranking the importance of the subcategories, which ranged from between 2 and 6 subcategories within each main attribute.

Familiarity With e-Bug

All 106 participants were provided with the website link and asked to familiarize or refamiliarize themselves with the teachers' section of the e-Bug website. After 5 min of exploring the e-Bug website, participants completing the questionnaire remotely (n=87) moved on to questionnaire part B.

Completion of Task Scenarios

The 19 participants that were face-to-face with the researchers (RA and CH) at the school or conference were asked to complete task scenarios. Participants were provided with scenarios, typical of users of the e-Bug website, and asked to think aloud [23] as they attempted to complete the task. The screen recording function on Skype for Business was used to record participants' website navigation and capture audio output. Multimedia Appendix 2 shows the task scenarios the participants were asked to complete. Researchers (RA and CH) made observational notes and probed about the ease of completion and ideas for improvement after each task scenario. For the task completion, some teachers participated at a conference; others were visited in their own school.

Questionnaire Part B

All 106 participants completed an electronic questionnaire (Part B - Multimedia Appendix 1), hosted on SelectSurvey. This electronic questionnaire included the following:

- Rating (5-point Likert scale) the e-Bug website against all website attributes previously ranked by them in part A: baseline questionnaire by indicating how strongly they agreed or disagreed with statements about the website, for example, "" (strongly disagree/disagree/neither agree or disagree/agree/strongly agree).
- Overall satisfaction with the website, including ideas for improvement
- Loyalty, measured by their enthusiasm to use e-Bug again and/or inclination to recommend to a colleague/friend
- Importance to participants of suggestions of some additional content, for example, search function; ability to like, share on social media, or email specific e-Bug resources to a friend; and development of an app for users.

Analysis

Task Scenarios

The task scenarios were used to analyze the following [24]:

- Effectiveness (whether it is possible to complete the realistic tasks for end users, which is measured by the percentage of tasks completed)
- Efficiency (whether end users are able to locate the resources using the quickest and most direct route through the website, which is measured by the number of "additional" clicks to locate resources)
- Learnability (whether the structure of e-Bug's website is easy to remember for future use, which is measured by the change in efficiency in the repeated task).

Questionnaire

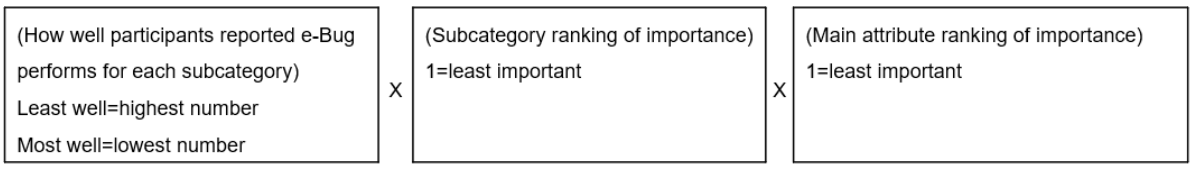
Participants' responses to ranking the importance of website attributes and rating the e-Bug website against these attributes were combined and averaged to provide a prioritization order for implementing change [25]. This was done as follows:

1. Post data collection, the ranking importance given to the 5 main attributes (appearance, content, interactivity, ease of use, and technical adequacy) and their subcategories was reversed so that the least important categories scored only one. This meant that, when calculated, the largest number indicated the highest priority for change.
2. For each subcategory, multiply the rating by the subcategory ranking by the category ranking (Figure 2).
3. Calculate the mean average for the respondents.
4. Divide by the number of subcategories within the main attribute to account for the differing number of subcategories in each main attribute group.
5. Order in descending order—highest number is the highest priority for change.

A random number generator on Excel was used to calculate a value for compatibility with other devices, guidance, sense of community, modern features, and limited use of special plug-ins, as it was not possible for participants to rate these attributes because of the fact that the features did not currently exist, for example, guidance, sense of community, and modern features, or because of the fact that, in a controlled environment, participants were not asked to browse the website on different devices. It was necessary to assign a value to these attributes so as to be included in the overall order for prioritization of change.

Descriptive statistics were used to analyze the remaining questionnaire data, and the free-text answers were collated and qualitatively analyzed, in combination with the transcribed audio output from completion of the task scenarios, using NVivo 10, to provide the main themes. These were discussed and agreed upon by the project team.

Figure 2. Calculation to combine importance of an attribute and performance.



Ethical Approval

This evaluation was approved by the PHE Research Ethics and Governance Group (REGG). After review, the Research Governance Coordinator for PHE confirmed that no ethical approvals were needed.

Results

Overview

A total of 19 participants completed the task scenarios and the questionnaire. Moreover, 87 participants completed the questionnaire only, resulting in a total of 106 heterogeneous participants evaluating the e-Bug website (see Figure 3 for the visual summary of the data collection process and Multimedia

Appendix 3 for an overview of participants and the devices used to access teaching resources).

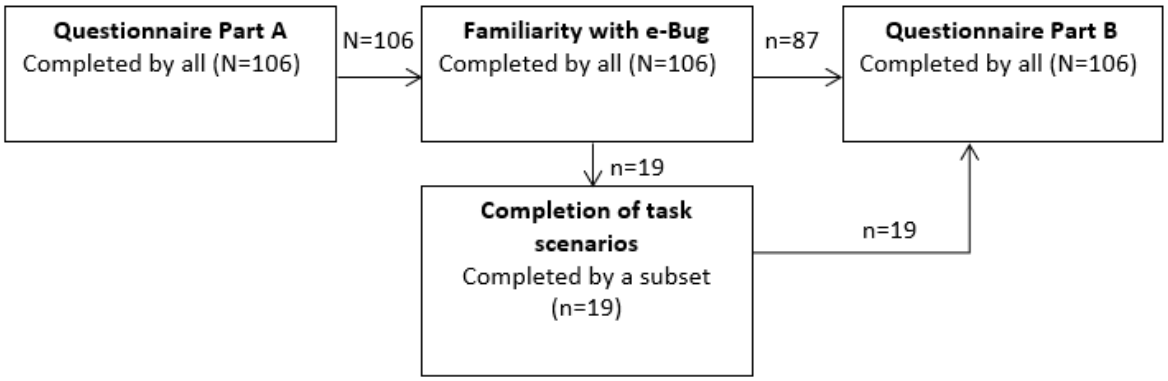
A total of 97.1% (103/106) of the participants would use the e-Bug website in the future, and 98.1% (104/106) of the participants would recommend it to a colleague or friend. Moreover, 65.0% (67/103) of the participants that indicated that they would use e-Bug again were first-time users of the website. A yes or no option button within the questionnaire part B was appropriate for the quantitative element of this attribute, with an additional open comment box for participants to provide reasons for their answers.

Participants' commented as follows:

> *Great website which has been highly effective in supporting the new science GCSE.* [Teacher, Secondary]
>
> *Definitely a niche in the market for it...useful to have a site that does all and that makes it applicable and user friendly for children.* [Assistant Head Teacher, Primary]

Multimedia Appendix 4 shows users' comments about the e-Bug website.

Figure 3. Visual summary of data collection process.



Completion of Task Scenarios

Table 1 shows the effectiveness (whether it is possible to complete the realistic tasks for end users, which is measured by the percentage of tasks completed) and Table 2 shows the efficiency (whether end users are able to locate the resources using the quickest and most direct route through the website, which is measured by the number of "additional" clicks to locate resources) of the e-Bug website. Moreover, 100% (19/19) of the participants were able to locate the full pack of resources (task 1) and the link to the national curriculum (task 5). Participants found it most difficult to return to the e-Bug home page when they had navigated away from it, as only 59% (10/17) of the participants were able to locate this, suggesting that this is an obvious area for improvement of the website.

Table 1. Effectiveness of the e-Bug website. Tasks are arranged from most effective at the top to least effective at the bottom (N=17).

Task number	Description of task	Effectiveness of participants that were able to complete the task, n (%)
Task 1	Full pack of resources	17 (100)
Task 5	National curriculum	17 (100)
Task 7	Task 1 again—full pack of resources	16 (94)
Task 3	Vaccinations timeline	14 (82)
Task 4	Antibiotic worksheet	14 (82)
Task 2	Hand hygiene complete pack	13 (77)
Task 6	e-Bug home page	10 (59)

Table 2. Efficiency of the e-Bug website. Tasks are arranged from most efficient at the top to least efficient at the bottom.

Task number	Description of task	Efficiency (average number of additional clicks participants needed to complete the task)
Task 3	Vaccinations timeline	0.1
Task 7	Task 1 again—full pack of resources	0.2
Task 5	National curriculum	0.2
Task 2	Hand hygiene complete pack	0.3
Task 4	Antibiotic worksheet	0.7
Task 1	Full pack of resources	0.8
Task 6	e-Bug home page	1.4

Most participants were able to complete scenarios typical of an e-Bug user. The following feedback was included:

> *Really easy to navigate around. I'm sure even a technophobe would be able to do it.* [Teacher, Secondary]

> *Possibly the fact I that had looked at the website previously, even just for a couple of moments, it is quite intuitive because each page is laid out in the same way. I know that I can get the whole pack and then individual resources: teacher's sheets, pupil's sheets. Funny, because I put consistency as quite low [referring to part A: baseline questionnaire], but seeing how consistent e-Bug is proves that this is actually quite important to me.* [Teacher, Primary]

> *Very useful resources and tools. Nice and teacher friendly. Doesn't require a massive amount of time browsing to find the information you need.* [Teacher, Secondary]

Participants were able to locate the vaccination timeline (task 3) most efficiently, with minimal unnecessary clicks around the website. Conversely, participants found locating the e-Bug home page most difficult (task 6). Those that were able to locate it (only 10/17, 59%) took, on average, an extra 1.4 clicks compared with the most efficient route, with 0-4 additional clicks needed to locate the e-Bug home page.

Feedback and suggestions for improvement included the following:

> *Would never think to click on the logo. Have a pop-up description when hover over with mouse. Call teacher's home page, teacher's hub.* [Teacher, Secondary]

> *Everything else is words so wasn't looking for a symbol.* [Teacher, Primary]

> *Not consistent with "young adult page." When scrolling over, it doesn't come up with "hand image"- noticed when surfing the websites. And this does not take you back to the e-Bug home page, so very confusing.* [Teacher, Secondary]

> *Expected home link to take to the e-Bug home.* [Teacher, Secondary]

In comparison with other task scenarios, participants also struggled to find the most efficient route to the full pack of resources (task 1—required an additional 0.8 clicks on average, range of 0-3 additional clicks) and the antibiotic worksheet (task 4—required an additional 0.7 clicks on average; range of 0-8 additional clicks). When finding the full pack of resources, participants were unaware that two separate links took them to the same page, and therefore, they often clicked both options. A suggestion to alleviate this included the following:

> *It would be useful that, when on a page, the link in the menu goes a different colour. Or when on a certain page, deactivate the link in the menu, so that you know you've actually gone to that page.* [Teacher, Secondary]

To increase the efficiency of finding specific worksheets, such as the student worksheet on antibiotics (task 4), suggestions included:

> *Perhaps when hover mouse over link, a dialogue box with info on each link could come up.* [Teacher, Secondary]

> *Would be useful to have a print-screen or image of what the worksheet looked like. Or rollover image...interactive feature.* [Teacher, Secondary]

Owing to the small sample size and the high efficiency and effectiveness of task scenario completion first time around, it was not possible to statistically analyze learnability (whether the structure of e-Bug's website is easy to remember for future use), which is measured by change in effectiveness and efficiency in repeated tasks (comparing tasks 1 and 7).

However, qualitative feedback included:

> *Having done one, it then became clear what to do with the other ones. Easy to learn.* [Teacher, Secondary]

> *A LOT easier than previous time. Would remember in a month's time. Practice made navigation easier.* [Teacher, Secondary]

Ranking Importance of Website Attributes

On average, participants ranked the content (68/106, 64.1% participants) of the website as the most important website attribute to themselves as educators and appearance (40/106, 37.7% participants) as the least important attribute (Figure 4). For ranking of all subcategories, see Multimedia Appendix 5.

Reasons given for ranking content as the most important attribute included:

It doesn't matter how flashy the website, if the content is wrong then it's no good. [Teaching Assistant, Secondary]

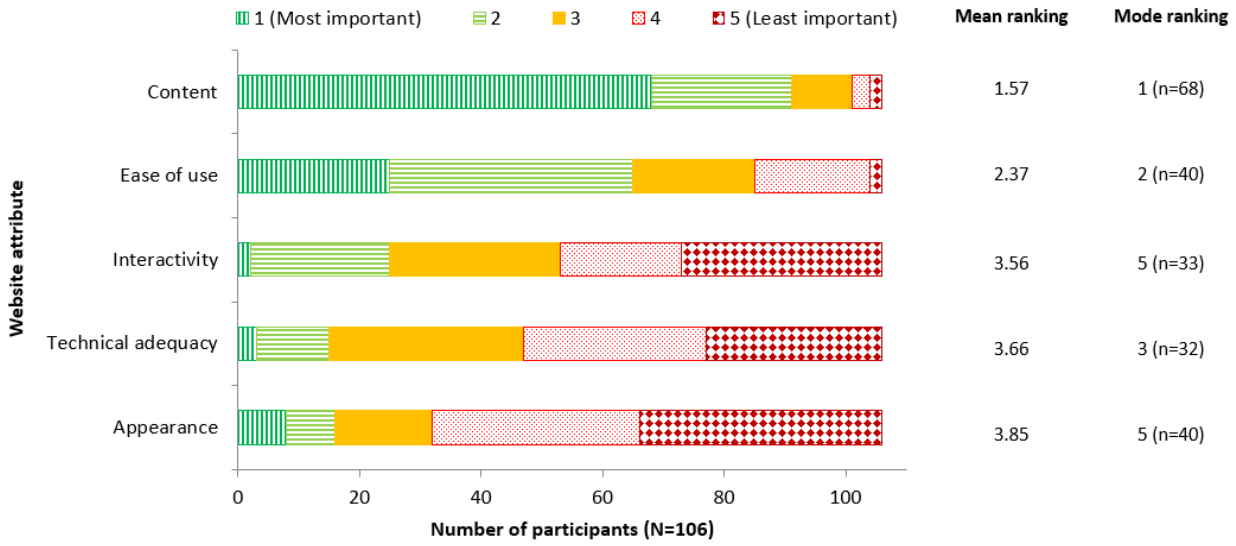
If the content is not correct, relevant, or detailed enough, then it is useless to me despite how nice it may look, or how easy it is to use. [Teacher, Secondary]

Reasons given for ranking appearance as the least important attribute included:

Whilst the appearance is important to attract people, I am more interested in appropriate, easily accessed information. [Teacher, Secondary]

I don't really care what it looks like, so long as I can adapt resources. [Teacher, Secondary]

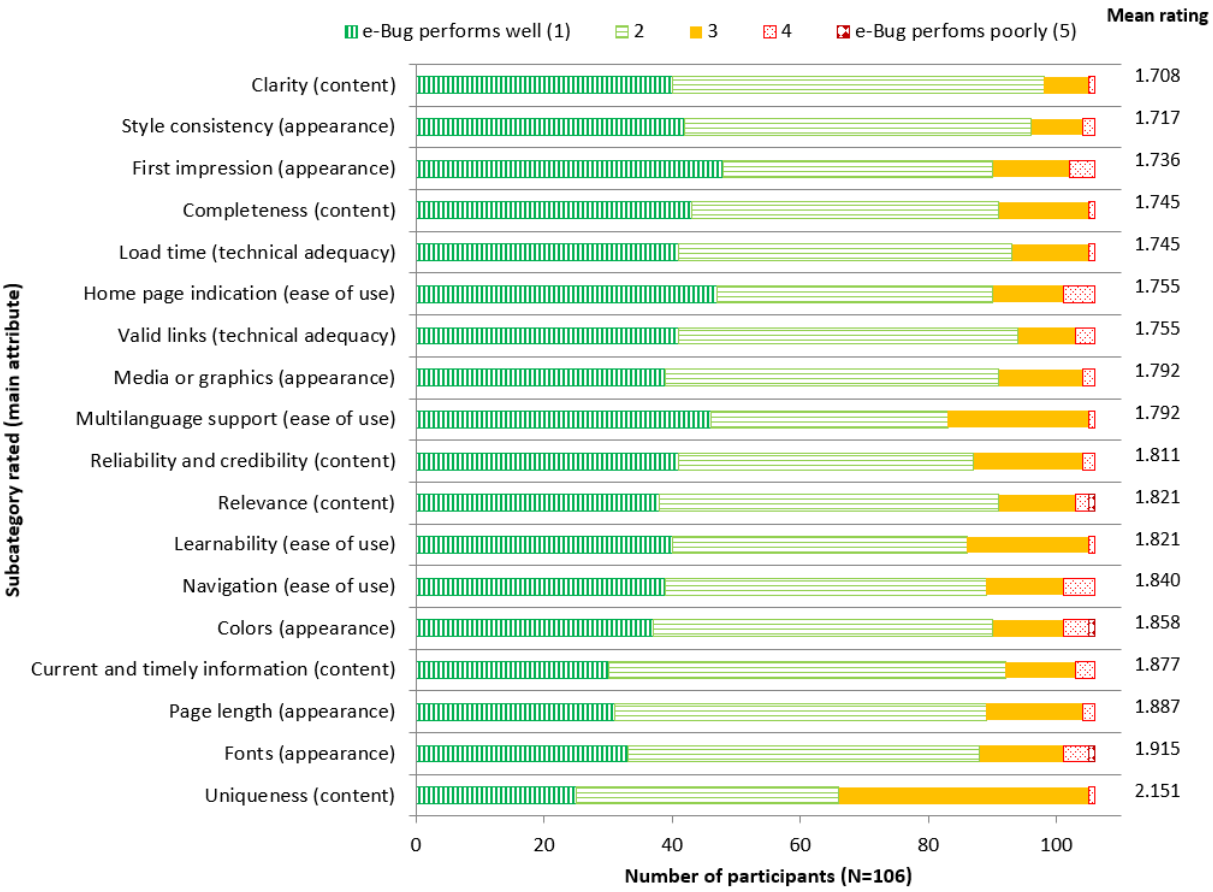
Figure 4. Importance of the five main website attributes to teachers.



Rating e-Bug Against the Website Attributes Previously Ranked

Figure 5 shows that overall, e-Bug performs very well, as all averages of the 18 attributes evaluated ranged between 1.708 and 2.151, with 1 being the best and 5 being the worst on the Likert scale. Clarity of content, style consistency, and first impression were rated as e-Bug's best qualities. Attributes that e-Bug was rated lowest against included uniqueness of content, fonts, and page length, but they were still rated positively on the Likert scale (<3).

Figure 5. Users' rating of the teachers' section of the e-Bug website in regards to different website attributes.



Prioritization for Change

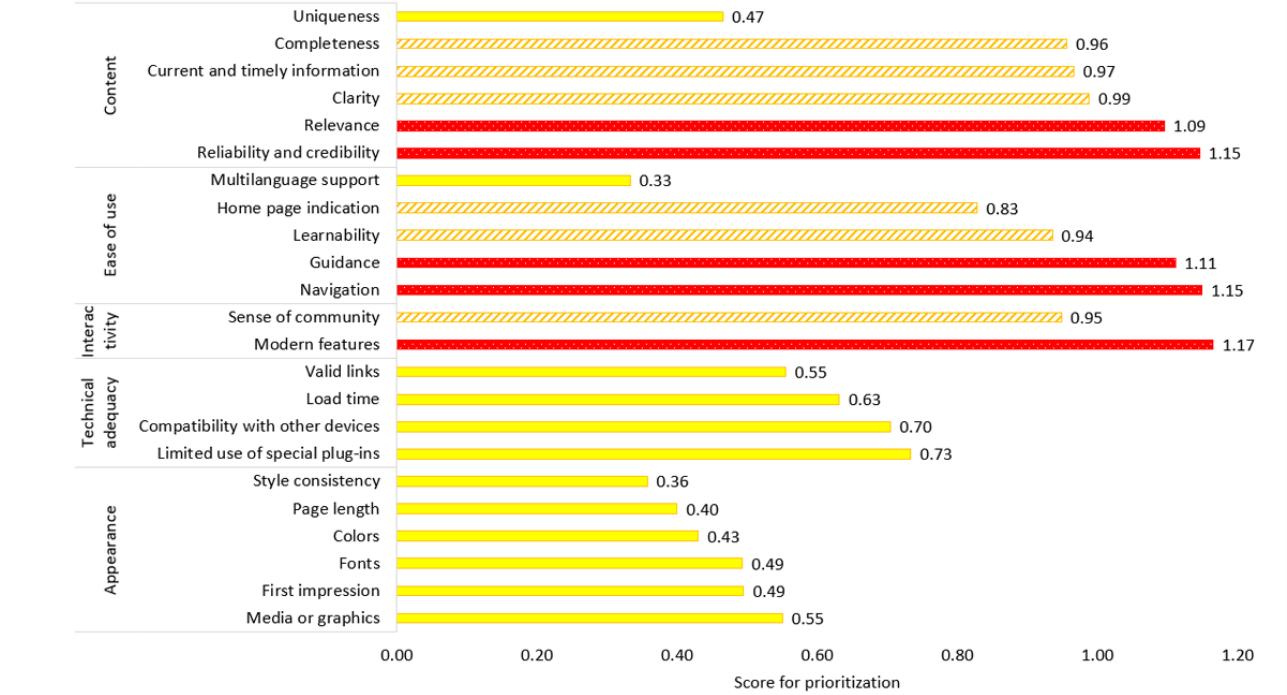
Figure 6 shows the priority score for implementing changes, based on users' perception of importance of attributes combined with how well e-Bug performed against these attributes. The highest priorities for change, highlighted in red in Figure 6, included modern features: the educational website reflects the most current trend or trends, eg, Twitter feeds visible and blog posts; navigation: navigating the educational website is intuitive, and it is easy to find the desired information; reliability and credibility: the educational website provides information that is trustworthy; guidance: the educational website provides help for users in recovering from common errors or assists them in the completion of tasks, eg, frequently asked questions, help option, and search tool; and relevance of content: the educational website offers content that is relevant to educators. The attribute of least priority is multilanguage support (the educational website supports its users' language preferences), which corresponds with the fact that e-Bug is available in 23 different languages.

Suggestions to improve reliability and credibility of e-Bug's content include:

> *Cite where the information given has come from. Clearly show the year the information was updated.* [Science Technician, Secondary]

See Multimedia Appendix 6 for participants' suggestions for improvement, using the prioritization order for change as a framework.

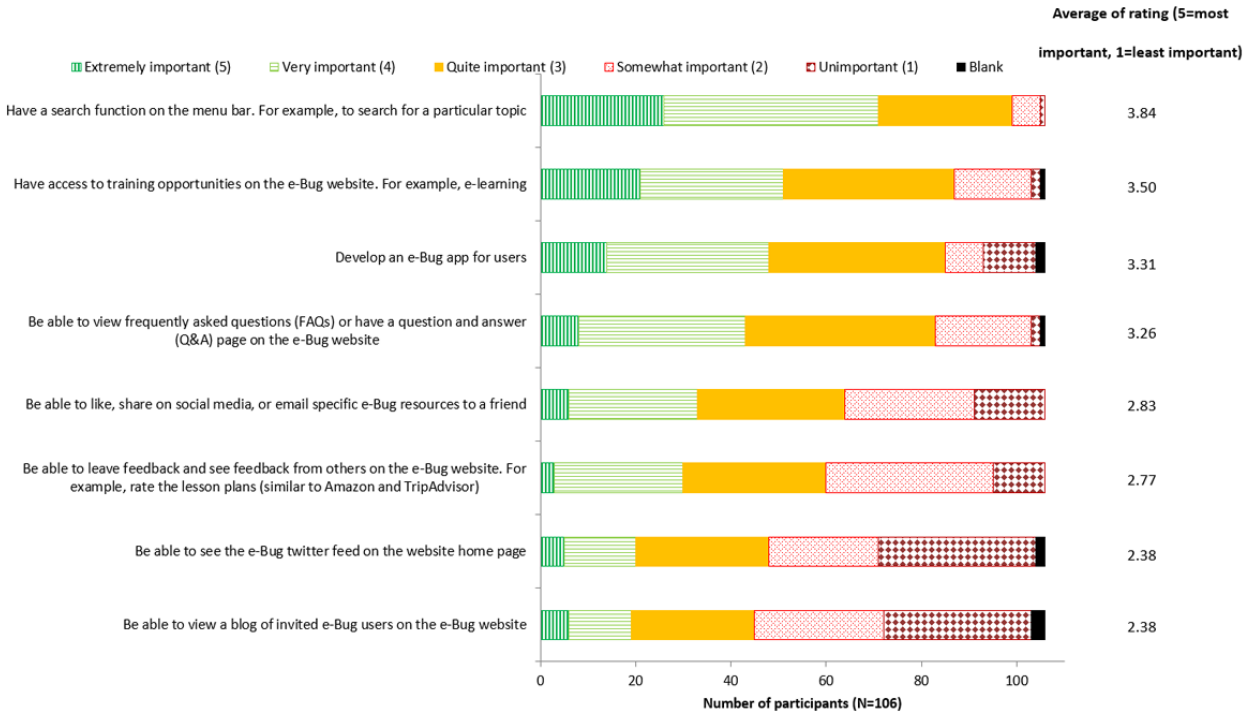
Figure 6. Prioritization score combining users' ranking of attributes' importance and rating of e-Bug's performance with red dots representing the highest priority for change, orange diagonal stripes lower priority, yellow block lowest priority.



Prioritization for Additional Features

Figure 7 shows the priority order for additional features to the teachers' section of the e-Bug website. Of features that e-Bug does not currently have, the highest priorities, rated by the users, include a search function (rated as "extremely important" or "very important" by 66.9% [71/106] of participants) and access to training opportunities (rated as "extremely important" or "very important" by 48.1% [51/106] of participants).

Figure 7. Importance of additional features to the teacher's section of the e-Bug website.



Discussion

Principal Findings

Overall, teachers liked the e-Bug website; 97.1% (103/106) of them would use it themselves, and 98.1% (104/106) of them would recommend e-Bug to others. Feedback showed that there are a lot of websites with science resources, but there was a niche for e-Bug because of the bulk of resources that fit into the national curriculum.

This study found that, for educational websites, users value the content of the website over other attributes, such as appearance (Figure 4). Teachers reported that they often take the content from educational health websites and locally modify it to fit their purpose or ability of the audience. Website designers of health care websites should be mindful of this when designing future platforms.

Findings from the completion of typical task scenarios with users of the e-Bug website were that it is generally easy to navigate and find resources, even with little previous experience on the website. Participants were able to locate the vaccination timeline efficiently, with minimal unnecessary clicks around the website, suggesting optimal information architecture [2] of this page. However, as participants were less able to return to the home page, they may have missed the whole student's section of the website, which includes games, revision guides, and quizzes for students to play, learn from, and complete as well as resources for communities and training for educators.

In terms of other improvements to e-Bug, features such as having a search function; developing an e-Bug app; and being able to like, share on social media, or email specific e-Bug resources to a friend could improve the "modern features" attribute, voted as the highest priority for change, and the "sense of community," ranked in the ninth position for priority for implementation. e-Bug was originally developed in 2009, and this study has clarified that the technology is out of date as there has been a shift in the digital market [26]. Further research is needed in this area to see how new media can better support teachers and educational providers.

Strengths and Limitations of GoodWeb

After piloting GoodWeb to evaluate the e-Bug website, it can be said that a strength of the framework is the easy step-by-step guide, which is adaptable and flexible to the website and objective of the evaluation.

Previous evaluations of health websites focus specifically on the quality of the content [27]. This study takes a holistic approach assessing the quality of the website as a whole, facilitated by the step-by-step guide of GoodWeb [7], which highlighted areas for improvement that would not otherwise have been identified.

A possible limitation is that the study design did not allow for "learnability" to be measured, because of the small sample size and the high efficiency and effectiveness of task scenario completion first time around. It would be advised that a larger sample size is required for this attribute, or conversely, more complex task scenarios, although this was not appropriate for e-Bug as most of the typical tasks are relatively simple.

A major strength of this study is the methodology chosen of both ranking the importance of attributes and then rating how well e-Bug performed against these attributes, which meant that it was possible to calculate an overall prioritization order for change. In addition, using a range of end users to evaluate the website, including those who were familiar with e-Bug and first-time users as well as participants with varying levels of computer literacy, means that e-Bug can be tailored to the needs of all end users.

A possible limitation of the methodology is that only a subset of participants completed the task scenarios. This decision was made because of the time taken for task completion and the logistics of the researchers observing, in person. If the study were to be repeated, the researchers could use a combination of observing the task scenarios in person and remotely, for example, utilizing the "share desktop" and "record" functions of Skype for Business. In this study, screen capture with audio output was essential to capture the task scenario data, as it allowed the researcher to assess efficiency, effectiveness, and learnability at a later time, and pull out and compare key themes from the discussion between the user and researcher, without compromising the situation at the time with excessive note taking and counting of mouse clicks. This process provided such rich data from just a subset of participants, that it was not deemed essential for all participants to complete.

A possible limitation of this evaluation is the small sample size, in comparison with larger studies [12]. However, there is still uncertainty over how many participants are needed to assess usability [28-32]. With the addition of the in-depth observational data during and interviewing post completion of task scenarios, data collected were rich and valuable for appraising quality and suggesting modifications for implementation, and the data were not dissimilar to other studies [33-38].

A constraint of this methodology is that subcategories could not be compared with subcategories from other categories, that is, ranking of clarity of content could not be compared with ranking of compatibility with other devices as they are within different categories (content compared with technical adequacy). Feedback from the feasibility study suggested that participants were unlikely to rank the 23 attributes assessed using the questionnaire, in order of importance, and therefore, the chosen methodology reflects the advice of users to successfully attain a high completion rate. This highlights the importance of patient and public involvement [39,40] throughout the evaluation process.

Furthermore, a random number generator was used to rate the attributes for features that did not currently exist, which could have introduced inaccuracies. If the study were to be repeated, it is advised that an average of other ratings is used instead, to account for this.

Recommendations

Areas that users ranked as most important and where e-Bug is currently not delivering as well as it could include modern

features, navigation, reliability and credibility of content, guidance, and relevance of content.

The highest priority for additional features that e-Bug does not currently have include a search function and access to more training opportunities, which may improve the subcategory "modern features." As a result, e-Bug has already reacted to these findings and now advertises e-learning modules and face-to-face e-Bug–approved educator training, through the e-Bug website. This has been successful as e-Bug has now trained over 100 educators (2016-2018), as advertising on the website and page views have started to increase (220,045 views in 2017-2018).

Users' suggestions to make the website even easier to navigate include highlighting or deactivating the current page in the menu bar, improving home page indication, and a preview of resources when hovering the mouse over hyperlinks.

Suggestions to improve reliability and credibility of content were to cite where the information had come from and clearly show the year that the content had last been updated.

By implementing the suggested changes and continuing to promote e-Bug, it is hoped that the trend of reduced use of the teachers' pages will be reversed and current and new users will be retained. It is recommended, therefore, that the e-Bug website is evaluated again, following the implementation of the suggested modifications.

Finally, GoodWeb, the comprehensive framework, was successfully applied to evaluate this educational health website and, therefore, could be used to guide future website evaluations in other fields.

Acknowledgments

This work was supported by the Primary Care Unit, PHE. The authors thank all teachers who participated in the evaluation of the teachers' pages of the e-Bug website. The authors also thank their collaborators who forwarded their advert on to their mailing lists and Neville Verlander for advising on the statistical analysis plan.

Authors' Contributions

RA wrote the protocol with input from CH, CM, and VY. RA and CH conducted the website analysis with users. RA wrote the final manuscript with input from CH, CM, and VY. All authors reviewed and approved the final manuscript.

Conflicts of Interest

At the time of this evaluation, CM was the Head of the Primary Care Unit that operates the e-Bug project; CH and VY both worked on the e-Bug project. RA, who led the research, has no conflicts of interest.

Multimedia Appendix 1
e-Bug website evaluation questionnaire.
[DOCX File , 69 KB - formative_v4i4e14504_app1.docx]

Multimedia Appendix 2
Task scenarios.
[DOCX File , 686 KB - formative_v4i4e14504_app2.docx]

Multimedia Appendix 3
Overview of participants and devices used to access teaching resources.
[DOCX File , 15 KB - formative_v4i4e14504_app3.docx]

Multimedia Appendix 4
User's comments about the e-Bug website.
[DOCX File , 17 KB - formative_v4i4e14504_app4.docx]

Multimedia Appendix 5
Ranking importance of sub-categories within categories to teachers.
[DOCX File , 14 KB - formative_v4i4e14504_app5.docx]

Multimedia Appendix 6
Participants' suggestions for improvement, using the prioritisation order for change as a framework.
[DOCX File , 20 KB - formative_v4i4e14504_app6.docx]

References

1. Song J, Zinkhan G. Features of web site design, perceptions of web site quality, and patronage behavior. 2003 Presented at: ACME 2003; March 4-8, 2003; Houston, Texas.
2. Mashable. 22 Essential Tools for Testing Your Website's Usability URL: http://mashable.com/2011/09/30/website-usability-tools/#cNv8ckxZsmqw [accessed 2020-08-16]
3. Gonzalez ME, Quesada G, Davis J, Mora-Monge C. Application of quality management tools in the evaluation of websites: The case of sports organizations. Qual Manage J 2015;22(1):30-46. [doi: 10.1080/10686967.2015.11918417]
4. Tsai WH, Chou WC, Lai CW. An effective evaluation model and improvement analysis for national park websites: A case study of Taiwan. Tour Manag 2010;31(6):936-952. [doi: 10.1016/j.tourman.2010.01.016]
5. Barnes SJ, Vidgen RT. Data triangulation and web quality metrics: A case study in e-government. Inf Manag 2006;43(6):767-777. [doi: 10.1016/j.im.2006.06.001]
6. Thielsch MT, Blotenberg I, Jaron R. User evaluation of websites: From first impression to recommendation. Interact Comput 2014;26(1):89-102. [doi: 10.1093/iwc/iwt033]
7. Allison R, Hayes C, McNulty CA, Young V. A comprehensive framework to evaluate websites: literature review and development of GoodWeb. JMIR Form Res 2019 Oct 24;3(4):e14372 [FREE Full text] [doi: 10.2196/14372] [Medline: 31651406]
8. Hartmann J, Sutcliffe A, Angeli AD. Towards a theory of user judgment of aesthetics and user interface quality. ACM Trans Comput-Hum Interact 2008;15(4):1-30. [doi: 10.1145/1460355.1460357]
9. Bargas-Avila JA, Hornbæk K. Old Wine in New Bottles or Novel Challenges: A Critical Analysis of Empirical Studies of User Experience. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.: ACM; 2011 Presented at: CHI'11; May 7 - 12, 2011; Vancouver BC Canada p. 2689-2698. [doi: 10.1145/1978942.1979336]
10. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Inf Technol 2006;25(2):91-97. [doi: 10.1080/01449290500330331]
11. Aranyi G, van Schaik P. Testing a model of user-experience with news websites. J Assn Inf Sci Tec 2016 Jul;67(7):1555-1575. [doi: 10.1002/asi.23462]
12. Elling S, Lentz L, de Jong M, van den Bergh H. Measuring the quality of governmental websites in a controlled versus an online setting with the 'Website Evaluation Questionnaire'. Gov Inf Q 2012;29(3):383-393. [doi: 10.1016/j.giq.2011.11.004]
13. Chase D, Trapasso E, Tolliver R. The perfect storm: examining user experience and conducting a usability test to investigate a disruptive academic library web site redevelopment. J Web Librarianship 2016;10(1):28-44. [doi: 10.1080/19322909.2015.1124740]
14. Petrie H, Precious J. Measuring User Experience of Websites: Think Aloud Protocols and an Emotion Word Prompt List. In: Proceedings of the 28th Annual CHI Conference on Human Factors in Computing Systems. 2010 Presented at: CHI'10; April 10-15, 2010; Atlanta, GA p. 3673-3678. [doi: 10.1145/1753846.1754037]
15. e-Bug. e-Bug Teachers Area URL: http://e-bug.eu/eng_home.aspx?cc=eng&ss=1&t=Welcome%20to%20e-Bug [accessed 2020-01-28]
16. McNulty CA, Lecky DM, Farrell D, Kostkova P, Adriaenssens N, Herotová TK, e-Bug Working Group. Overview of e-Bug: an antibiotic and hygiene educational resource for schools. J Antimicrob Chemother 2011 Jun;66(Suppl 5):v3-12. [doi: 10.1093/jac/dkr119] [Medline: 21680584]
17. Regis T, Stone J. The National Institute for Health and Care Excellence. 2017 Jan 25. Children and Young People Should Be Taught Simple Hygiene Measures to Help Curb the Spread of Infections, Says Nice URL: http://indepth.nice.org.uk/children-and-young-people-should-be-taught-simple-hygiene-measures-to-curb-the-spread-of-infections-says-nice/index.html [accessed 2018-08-01]
18. Department of Health and Social Care. UK Government. 2019 Jan 24. Tackling Antimicrobial Resistance 2019-2024: The UK's Five-Year National Action Plan URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/773130/uk-amr-5-year-national-action-plan.pdf [accessed 2019-11-01]
19. Eley CV, Young VL, Hayes CV, Verlander NQ, McNulty CA. Young people's knowledge of antibiotics and vaccinations and increasing this knowledge through gaming: mixed-methods study using e-Bug. JMIR Serious Games 2019 Feb 1;7(1):e10915 [FREE Full text] [doi: 10.2196/10915] [Medline: 30707096]
20. Hale AR, Young VL, Grand A, McNulty CA. Can gaming increase antibiotic awareness in children? a mixed-methods approach. JMIR Serious Games 2017 Mar 24;5(1):e5 [FREE Full text] [doi: 10.2196/games.6420] [Medline: 28341618]
21. Young VL, Rajapandian V, Eley CV, Hoekstra BA, Lecky DM, McNulty CA. Monitoring web site usage of e-Bug: a hygiene and antibiotic awareness resource for children. JMIR Res Protoc 2015 Nov 13;4(4):e131 [FREE Full text] [doi: 10.2196/resprot.4049] [Medline: 26567127]
22. Eley CV, Young VL, Hoekstra BA, McNulty CA. An evaluation of educators' views on the e-Bug resources in England. Journal of Biological Education 2017;52(2):166-173. [doi: 10.1080/00219266.2017.1285808]
23. Yen PY, Bakken S. A comparison of usability evaluation methods: heuristic evaluation versus end-user think-aloud protocol - an example from a web-based communication tool for nurse scheduling. AMIA Annu Symp Proc 2009 Nov 14;2009:714-718 [FREE Full text] [Medline: 20351946]

24. Arrue M, Fajardo I, Lopez JM, Vigo M. Interdependence between technical web accessibility and usability: its influence on web quality models. Int J Web Eng Technol 2007;3(3):307. [doi: 10.1504/ijwet.2007.012059]
25. Agarwal R, Venkatesh V. Assessing a firm's Web presence: a heuristic evaluation procedure for the measurement of usability. Inf Syst Res 2002;13(2):168-186. [doi: 10.1287/isre.13.2.168.84]
26. Budiu R, Nielsen J. Semantic Scholar. Fremont, California: Nielsen Norman Group; 2015. User Experience for Mobile Applications and Websites: Design Guidelines for Improving and Usability of Mobile Sites and Apps URL: https://pdfs.semanticscholar.org/6d75/e07170a76954cc7c78990d744b2a3119c955.pdf [accessed 2020-02-18]
27. Eysenbach G, Powell J, Kuss O, Sa E. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. J Am Med Assoc 2002;287(20):2691-2700. [doi: 10.1001/jama.287.20.2691] [Medline: 12020305]
28. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003 Aug;35(3):379-383. [doi: 10.3758/bf03195514] [Medline: 14587545]
29. Hwang W, Salvendy G. Number of people required for usability evaluation: the 10±2 rule. Commun ACM 2010;53(5):130. [doi: 10.1145/1735223.1735255]
30. Macefield R. How to specify the participant group size for usability studies: a practitioner's guide. J Usability Stud 2009;5(1):34-45 [FREE Full text]
31. Nielsen J. Usability Engineering. Cambridge, Massachusetts: Academic Press; 1994.
32. 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]
33. Muhtaseb R, Lakiotaki K, Matsatsinis N. Applying a multicriteria satisfaction analysis approach based on user preferences to rank usability attributes in E-tourism websites. J Theor Appl Electron Commer Res 2012;7(3):28-48. [doi: 10.4067/s0718-18762012000300004]
34. Moreno JM, del Castillo JM, Porcel C, Herrera-Viedma E. A quality evaluation methodology for health-related websites based on a 2-tuple fuzzy linguistic approach. Soft Comput 2009;14(8):887-897. [doi: 10.1007/s00500-009-0472-7]
35. Leite P, Gonçalves J, Teixeira P, Rocha A. Towards a model for the measurement of data quality in websites. New Rev Hypermedia Multimedia 2014;20(4):301-316. [doi: 10.1080/13614568.2014.968638]
36. Flavián C, Guinalíu M, Gurrea R. The role played by perceived usability, satisfaction and consumer trust on website loyalty. Inf Manag 2006;43(1):1-14. [doi: 10.1016/j.im.2005.01.002]
37. Emmert M, Adelhardt T, Sander U, Wambach V, Lindenthal J. A cross-sectional study assessing the association between online ratings and structural and quality of care measures: results from two German physician rating websites. BMC Health Serv Res 2015 Sep 24;15:414 [FREE Full text] [doi: 10.1186/s12913-015-1051-5] [Medline: 26404452]
38. Deng J, Zhang T, Li M. The Empirical Analysis of Design Quality Measurement System Concerning Tour Electronic Commerce Website in China. In: Proceedings of the 2011 International Conference on E-Business and E-Government. 2011 Presented at: ICEE'11; May 6-8, 2011; Shangai, China URL: https://doi.org/10.1109/ICEBEG.2011.5881858 [doi: 10.1109/ICEBEG.2011.5881858]
39. Brett J, Staniszewska S, Mockford C, Herron-Marx S, Hughes J, Tysall C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. Health Expect 2014 Oct;17(5):637-650 [FREE Full text] [doi: 10.1111/j.1369-7625.2012.00795.x] [Medline: 22809132]
40. NIHR Involve. Briefing Note Three: Why Involve Members of the Public in Research? URL: https://www.invo.org.uk/posttyperesource/why-should-members-of-the-public-be-involved-in-research/ [accessed 2018-11-01]

Abbreviations

NICE: National Institute for Health and Care Excellence
PHE: Public Health England

Edited by G Eysenbach; submitted 26.04.19; peer-reviewed by C Eley, C Brown, J Lander; comments to author 14.06.19; revised version received 09.10.19; accepted 22.10.19; published 06.04.20.

Please cite as:
Allison R, Hayes C, Young V, McNulty CAM
Evaluation of an Educational Health Website on Infections and Antibiotics in England: Mixed Methods, User-Centered Approach
JMIR Form Res 2020;4(4):e14504
URL: https://formative.jmir.org/2020/4/e14504
doi:10.2196/14504
PMID:32203932

Original Paper

A Mobile Sexual Health App on Empowerment, Education, and Prevention for Young Adult Men (MyPEEPS Mobile): Acceptability and Usability Evaluation

Brittany Gannon[1,2], RN, PhD; Rindcy Davis[3], PhD; Lisa M Kuhns[4,5], PhD; Rafael Garibay Rodriguez[1], BA; Robert Garofalo[4,5], MPH, MD; Rebecca Schnall[1], RN, PhD

[1]School of Nursing, Columbia University, New York, NY, United States

[2]HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute, New York City, NY, United States

[3]Gertrude H Sergievsky Center, Columbia University, New York, NY, United States

[4]Lurie Children's Hospital, Chicago, IL, United States

[5]Ann & Robert H Lurie Children's Hospital of Chicago, Northwestern University Feinberg School of Medicine, Chicago, IN, United States

Corresponding Author:
Rebecca Schnall, RN, PhD
School of Nursing
Columbia University
560 West 168th Street
New York, NY, 10032
United States
Phone: 1 2123426886
Email: rb897@columbia.edu

Abstract

Background: HIV incidence among young adult men who have sex with men (MSM), particularly among black and Latino men, continues to rise. As such, continued HIV prevention interventions for young MSM of color are of utmost importance. Male Youth Pursuing Empowerment, Education and Prevention around Sexuality (MyPEEPS) Mobile is a comprehensive HIV prevention and sexual health education smartphone app initially created to promote sexual health and HIV prevention among adolescent sexual minority young men aged 13 to 18 years.

Objective: The objective of this study was to critically appraise the acceptability and usability of MyPEEPS Mobile for young adult MSM aged 19 to 25 years.

Methods: Study participants used the mobile app, completed usability questionnaires and in-depth interviews, and reported their experience using the app. Analysis of interview data was guided by the Unified Theory of Acceptance and Use of Technology (UTAUT) to better understand the usability and acceptability of this intervention for young adults. Interview data were coded using the following constructs from the UTAUT model: performance expectancy, effort expectancy, and social influence.

Results: A total of 20 young adult MSM (n=10 in Chicago, Illinois, and n=10 in New York, New York) were enrolled in the study. Participants reported that MyPEEPS Mobile was free of functional problems (Health Information Technology Usability Evaluation Scale scores and Post-Study System Usability Questionnaire scores consistent with high usability), easy to use, and useful, with an engaging approach that increased acceptability, including the use of avatars and animation, and inclusive representation of the diverse identities by race and ethnicity, gender identity, and sexual orientation. Recommended areas for improving MyPEEPS Mobile for the target demographic included more adult-oriented graphics, advanced educational content, scenarios for youth with more sexual experience, and search function to increase accessibility of key content.

Conclusions: Overall, young adult MSM aged 19 to 25 years described the MyPEEPS Mobile as educational, informative, and usable for their sexual health education and HIV prevention needs, and they provided actionable recommendations to optimize its use and applicability for this age group.

(JMIR Form Res 2020;4(4):e17901) doi:10.2196/17901

KEYWORDS

young adults; usability; HIV; mHealth; young men; mobile phone

Introduction

Background

Of the 38,739 new HIV diagnoses in the United States and dependent areas, young men, aged 13 to 24 years, accounted for 18% of new diagnoses of HIV in 2017 [1]. Despite continued advances in HIV prevention due, in part, to the widespread promotion of treatment as prevention [2,3] for HIV-positive individuals and pre-exposure prophylaxis (PrEP) [4,5] for high-risk HIV-negative individuals, young men who have sex with men (YMSM) continue to bear the burden of new HIV infections. Of the young men diagnosed with HIV in 2017, 21% were aged 15 to 19 years and 79% were aged 20 to 24 years, with 93% of male cases attributed to male-to-male sexual contact [6]. Further highlighting the severity of the epidemic among subpopulations, young black/African American and Hispanic/Latino youth account for the highest proportion of new HIV diagnoses [6]. Given the racial/ethnic disparities in the epidemic, access to culturally and developmentally appropriate HIV prevention tools for YMSM of color is vital to *Ending the HIV Epidemic: A Plan for America* initiative in the United States, as proposed by the Department of Health and Human Services [7].

Sexual risk factors among YMSM identified in the literature include low rates of HIV testing; high rates of sexually transmitted infections (STIs), particularly among youth of color aged 20 to 24 years; substance use; as well as low levels of consistent condom and PrEP use [6,8]. YMSM face additional multilevel challenges to accessing effective sexual health education and services, including low risk awareness, intersectional stigma, poor access to care, high cost of care, medical mistrust, homelessness, and lack of culturally appropriate care [9-12].

According to the US 2020 National HIV/AIDS Strategy, public health initiatives should focus on increasing HIV and STI testing, linkage to care, universal viral suppression, and increasing PrEP awareness and access [13]. In addition to these strategies, access to HIV prevention information via Web and mobile platforms has been shown to improve knowledge, acceptability, and utilization of health prevention interventions and services [14].

In consideration of the 2020 National HIV/AIDS Strategy, several culturally tailored mobile health (mHealth) technology interventions have been designed in areas of primary HIV prevention (eg, myDEx, *Male Youth Pursuing Empowerment, Education and Prevention around Sexuality* [MyPEEPS], and HealthMindr), HIV testing and PrEP uptake (eg, P3: Prepared, Protected, emPowered; Mychoices; and mLab), and *linkage to care* (eg, LYNX and Get Connected) and are in various stages of testing for feasibility, acceptability, and efficacy [14-20]. Advantages of mHealth apps include the ability to rapidly and cost-effectively disseminate information broadly while also addressing the stated needs and desires of men who have sex with men (MSM) [14,21,22]. Moreover, use of mobile apps for sexual health education may facilitate privacy for YMSM while accessing sensitive and often stigmatized health information [23].

Objective

The *MyPEEPS* curriculum was originally developed as an in-person group-based intervention for racially and ethnically diverse YMSM to improve HIV-related risk [24]. For the group-based intervention, MyPEEPS was manualized and consisted of six interactive sessions focusing on HIV and STI epidemiology in YMSM, building knowledge and skills for safer sex, minority stress, emotion regulation, interpersonal and substance-related risk factors, developing risk reduction plans, and condom negotiation. MyPEEPS Mobile was adapted from the in-person, group-based intervention to a mobile app via an iterative process, including expert panel reviews, in-depth interviews with adolescent YMSM aged 13 to 18 years [25], a rigorous usability evaluation [26], and pilot testing [18].

MyPEEPS Mobile is a mobile, responsive Website that is viewable on small screens and usable with touch screens. MyPEEPS Mobile provides educational information about STIs and HIV for YMSM, builds skills for condom use, and raises awareness of minority stress. The app content is guided by 4 peeps: Tommy, Philip, Nico, and Artemio, or composite characters/avatars, who relay content through comics, animation, and scenarios delivered through 21 brief activities in four sequential modules (see Multimedia Appendix 1). A running theme throughout the intervention is the "Bottom Line," in which participants can set goals about their sexual risk reduction and commit to how much sexual risk they are willing to undertake. Privacy is protected via password and automatic log-off of the app after 20 min of inactivity.

In this study, we sought to assess the usability of MyPEEPS Mobile among an older age group with increased risk for HIV infection. To do so, we conducted in-depth interviews about the app content with young adult MSM aged 19 to 25 years.

Methods

Study Period and Participant Inclusion

This study was approved by the Institutional Review Board (IRB) at Columbia University Medical Center, which served as the single IRB of record. Data were collected in New York, New York, from August 2018 to October 2018 and in Chicago, Illinois, from December 2018 to February 2019. An a priori sample size of approximately 20 was estimated to provide saturation of acceptability and usability themes [27]. Individuals were recruited for participation via advertisement on social media platforms (ie, Instagram, Grindr, etc), via flyer distribution at local community locations and events (New York City and Chicago), and among research participants from other studies at the two sites. Inclusion criteria were as follows: (1) aged 19 to 25 years, (2) male sex assigned at birth, (3) self-identified as male or gender nonconforming/nonbinary, (4) comfortable speaking and reading in English, (5) living within the metropolitan area of New York City or Chicago, (6) anal or oral sex with another male in the past 12 months, and (7) self-reported HIV-negative or unknown status. Interested individuals were screened for participation via a Web or phone screening survey and, if eligible, enrolled in a subsequent in-person study visit.

Data Collection

Participants met the study staff at each study site for a single visit, lasting approximately 3 to 4 hours. The interviews were conducted in private conference rooms at the Lurie Children's Hospital and the Columbia University School of Nursing. After initial informed consent and completion of a computer-assisted self-interview, including demographic and behavioral questions, participants were instructed to complete intervention activities in the MyPEEPS Mobile app. While using MyPEEPS Mobile, participants recorded notes on their perceptions of app design and content. After completing the modules, participants completed the following usability questionnaires: (1) Health Information Technology Usability Evaluation Scale (Health-ITUES) [28] and (2) Post-Study System Usability Questionnaire (PSSUQ) [29,30]. Participants then completed face-to-face, audio-recorded, qualitative interviews, facilitated by study staff, using a semistructured interview guide. The following questions were included in the guide: (1) Thinking back about the information you learned from the MyPEEPS app, how would you apply this information/lessons/activities in your own life?; (2) How do the MyPEEPS activities reflect your cultural beliefs, norms, values?; (3) How do you perceive this app would be of relevance to other young adult MSM aged 19 to 25 years?; and (4) How would you modify these activities, if needed, to make them more relevant to young adult MSM aged 19 to 25 years? Data were collected until saturation was reached.

Theoretical Model

The Unified Theory of Acceptance and Use of Technology (UTAUT) model was originally developed as a conceptual framework to explain individuals' intention to adopt and use technological innovations. In this study, we draw from this theory to describe the applicability and likelihood of use of the MyPEEPS Mobile app among young adult MSM aged 19 to 25 years [31-33]. Originally developed to explain employee adoption of technology in the workplace, the UTAUT model has been extended to new contexts, such as health information systems and new populations (eg, consumers and health care professionals) [34,35]. Here, we use the UTAUT model to evaluate the applicability of MyPEEPS for young adult MSM in community settings where they access and use mobile apps. The following key theoretical constructs from the UTAUT model guided the analysis of the in-depth interview data: (1) performance expectancy, (2) effort expectancy, and (3) social influence [32,33].

Performance expectancy is defined as the degree to which an individual believes that using the technology will help them attain gains in the outcome of interest, that is, work performance in the original conceptualization, and herein, health protective behavior. Effort expectancy is the degree of ease associated with use of the technology. Finally, social influence is the extent to which an individual perceives that important others believe they should use the technology. These constructs are theorized to drive behavioral intention, which leads to use behavior. That is, if users expect use of the technology to improve the outcome of interest, find it easy to use, and perceive that important others believe they should use it, use of the technology will follow. Several constructs that stem from organizational contexts used in the original model (ie, facilitating conditions and extrinsic motivation) were eliminated in this study, as they were not applicable because use of the app does not depend on an organizational structure to drive use.

In addition to the primary constructs, the UTAUT model also includes moderators that are theorized to impact the pathway from primary constructs' technology use, based on individual characteristics. For this study, we sought to identify potential contextual factors that might also impact the pathway to use, based on analysis of interview data.

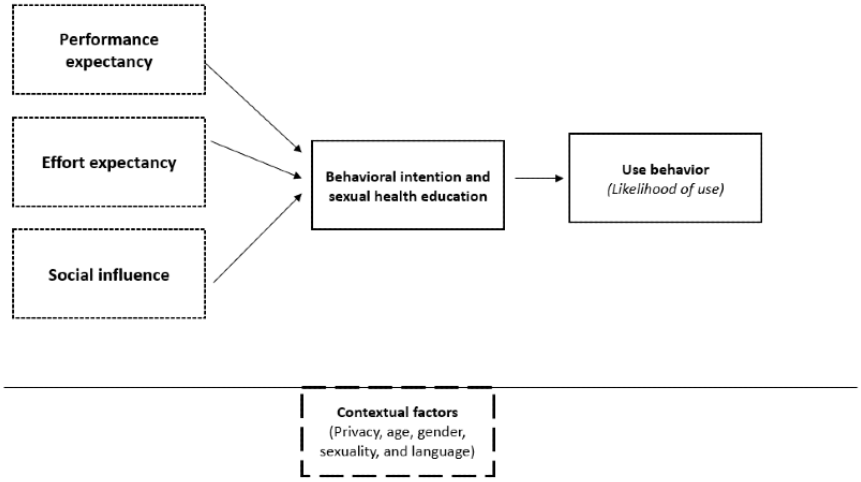
Qualitative Analyses

The qualitative analysis process consisted of two rounds of coding, including primary, open coding of emergent themes and secondary, thematic coding of those themes onto UTAUT constructs [36]. Audio recordings were transcribed verbatim and coded by 2 independent coders (BG and RD) in NVivo qualitative data analysis software (QSR International Pty Ltd, version 12, 2018). Before coding, both reviewers wrote a reflexivity statement of their background, preconceived notions, and/or engagement in HIV research and lesbian, gay, bisexual, transgender, queer affairs. This form of self-appraisal is utilized to ensure rigor of the analyses and brings awareness to the interpretive lens through which a qualitative analysis is being performed [37].

Initial content analyses via line-by-line coding, a form of open coding, was performed by the 2 coders independently to identify emergent codes [36]. The coders then met to debrief and discuss emergent codes derived from this process [38]. Secondary content analyses were then performed by applying the UTAUT as a semistructured framework; see Figure 1 [32,39]. Individual coders were tasked to apply themes within the UTAUT model—incorporating main constructs and subconstructs.

The coders met after completion of secondary analyses to discuss the fit of their codes and any variances in the application to the UTAUT model. Cognitive mapping of secondary codes (development of a graphic map that represents the relationships between concepts) was then performed in NVivo to facilitate discussion, integration, and agreement for the development of a preliminary version of the adapted model. Moreover, 3 of the 4 macro-constructs were agreed upon and utilized in the adapted model to explain the applicability and likelihood of use of MyPEEPS among young adult MSM, including (1) performance expectancy, (2) effort expectancy, and (3) social influence. Several constructs of the UTAUT model (ie, facilitating conditions, extrinsic motivation, etc) were not apparent in the interview data, and thus, they were not applied in the adapted model. To facilitate credibility and reliability, the study team, comprising the principal investigator, primary coders, and the study coordinator, then convened for consensus coding of UTAUT-related content. Applicability of constructs and codes were discussed with any discrepancies among codes negotiated and reached by majority [40,41]. An adapted version of the UTAUT model is presented in Figure 1, highlighting the thematic structure of the end user's experience with MyPEEPS.

Figure 1. Adapted Unified Theory of Acceptance and Use of Technology model.



Results

Study Sample

The study sample comprised 20 YMSM, aged 19 to 25 years, recruited from New York (n=10) and Chicago (n=10), with a median age of 22.5 years. Regarding sexual orientation, 90% (18/20) identified as gay, 5% (1/20) identified as bisexual, and 5% (1/20) identified as pansexual. In terms of race/ethnicity, 50% (10/20) identified as white, 35% (7/20) black/African American, 10% (2/20) multiracial, and 5% (1/20) other (Puerto Rican). Of these participants, 45% (9/20) had completed college, 25% (5/20) had completed some college, 20% (4/20) had completed high school, 5% (1/20) was currently enrolled in high school, and 5% (1/20) had completed a master's degree.

Usability Ratings

Overall, participants rated the app as highly usable on the usability instruments. On the PSSUQ, all subscales, including system quality, information quality, and interface quality, fell below 2 on average, indicating high usability (see Table 1). The overall mean (SD) PSSUQ score on all 16 items was 1.83 (0.43). The subconstructs of the Health-ITUES, including system impact, perceived usefulness, perceived ease of use, and user control, were all rated above 4 on average, indicating that participants rated the app as highly usable (see Table 2). The mean (SD) Health-ITUES score on all 20 items was 4.41 (0.23).

Table 1. Post-Study System Usability Questionnaire scores (N=20).

PSSUQ[a] constructs	Score[b], mean (SD)
System quality	1.49 (0.15)
Information quality	1.98 (0.42)
Interface quality	1.48 (0.10)

[a]PSSUQ: Post-Study System Usability Questionnaire.

[b]Scores range from 1=strongly agree to 7=strongly disagree.

Table 2. Health Information Technology Usability Evaluation Scale scores.

Health-ITUES[a] constructs	Score[b], mean (SD)
System impact	4.37 (0.24)
Perceived usefulness	4.29 (0.16)
Perceived ease of use	4.67 (0.08)
User control	4.38 (0.28)

[a]Health-ITUES: Health Information Technology Usability Evaluation Scale.

[b]Scores range from 1=strongly disagree to 5=strongly agree.

Qualitative Findings

Qualitative findings are arranged by the constructs of the UTAUT model. Quotes from study participants are presented below followed by their city of residence and their age in years.

Performance Expectancy

Participants described their intention to use the app within their everyday lives. In particular, one participant said:

> *I think it's applicable to anyone's life who is sexually active. Obviously, the messaging was catered to a*

> *gay demographic. So, identifying as a gay male, I did find it applicable to my life; more specifically, my sex life. So, yeah, I did think that the information that was shared is something that I can use in my everyday life.* [Chicago, 24 years]

Performance expectancy and plans to use the MyPEEPS app were further operationalized through the three subconstructs of the UTAUT model: (1) perceived usefulness, (2) outcome expectations, and (3) relative advantage [32].

Perceived usefulness of MyPEEPS was ascertained from end users' summative experiences and perceived knowledge attained while using the app. Most participants noted that the app was informative and useful beyond standard sexual health education, with statements such as:

> *I feel like even I learned something new today, so we could all learn something new.* [New York, 21 years]

Participants expressed a foreseeable usefulness of MyPEEPS for YMSM who live in different geographic regions (ie, outside of urban settings):

> *Obviously, in more conservative areas you still have like very strong conservative values on that. Like, the portion of sex ed that dealt with gay sex, was like five minutes, you know. So, I think it would be applicable for people who don't have the same type of sex education because there's still plenty of places that their sex education is still abstinence first, which is just not reality, you know.* [New York, 19 years]

On the other hand, participants who had more educational or life experience perceived the app as less useful. For example, one user noted:

> *For me it felt redundant because it didn't change mine [bottom line] at all, but maybe for someone who is younger or has less sexual experience or education it could still...it could be helpful for them.* [Chicago, 24 years]

Another participant commented that the app was less useful to him because of his educational background:

> *Coming from a student background, especially in a school of public health, this is an activity that we had done before, so it was kind of a little bit repetitive for me just because I have seen these questions before, and I know what we're doing here.* [New York, 23 years]

When probed by the interviewer regarding usefulness of the app, one end user commented:

> *So, I guess it's important to clarify – what is the purpose of the app? Is it to be a one-time resource, or is it to be a go-back-to-it resource?* [Chicago, 25 years]

Outcome expectations were expectations regarding the impact of the use of the app on health behavior. After participants used the app, one user reflected his intent to change behavior with his partner:

> *Me and my lover have to plan. I know I will be looking at the app. When I get home, I am going to tell him all about it.* [New York, 19 years]

Furthermore, use of culturally representative avatars and animation surpassed participants' expectations relative to expected traditional sexual health didactic content. One participant stated:

> *I really like "Testing with Tommy" where it literally...you go through what it's like to go to an HIV clinic and go through that experience. Like, you take a number, and there are a bunch of people there. It's like it's not a big deal.* [Chicago, 25 years]

One of the major expectations "not met" by participants was the availability and accessibility of sexual health resources. A representative comment about this was as follows:

> *I really wish that all the references and everything could have been provided somewhere versus going back inside the modules.* [New York, 23 years]

Although sexual health–related external references and hotlines were provided in activities throughout the app, this did not meet end user's expectation of having a single "go-to" resource tool, which was a content recommendation.

Relative advantage is the extent to which using the technology is perceived as better than using its precursor [32]. In this study, relative advantage included participants' perceptions of the use of MyPEEPS relative to the current standards of sexual health education (ie, provided by either parent(s)/guardian, health care provider, school system, etc). Uniquely, the use of relatable avatars was lauded relative to the standard provision of sexual health education. Moreover, end users conveyed that MyPEEPS was useful for those with little to no sexual health education:

> *I think it's also nice because it's very educational and kind of starts easy, like you don't have to have any prior knowledge, like any sexual education, before. The app really provides you with a lot of baseline sexual education, which is really important.* [Chicago, 24 years]

Acknowledging the individual variability in sexual experience, end users found the app to be useful for those who have less sexual experience or knowledge:

> *I thought it did a great job about...talking about the bottom line, like standing by the things that are important to you. I could see as someone who didn't have much information about sexual education and maybe didn't know what the risks of different sexual activities were they would get more information and that would change their bottom line.* [Chicago, 24 years]

Effort Expectancy

Effort expectancy is defined as the degree of ease associated with the use of MyPEEPS, including *ease of use* and *complexity*. Most end users described the mobile app as *easy to use*. For example, one participant said:

> *I felt the flow of the app was really easy to follow and set up in a very nice way.* [New York, 19 years]

Most end users found the overall interface of MyPEEPS to be intuitive, approachable, and easy to use:

> *Well it's pretty straightforward and easy to use. The logo is cute.* [New York, 20 years]

The MyPEEPS activity map was designed as an urban city street, reflecting participant's environment. Overall, participants at these study sites related to the map design. Reflecting on the ease of navigation, one participant noted:

> *I loved the platform of scrolling through the street. I thought that was really cool because you can see what's still to come. While they're only brief little titles you can still see what's going on, you can see there is going to be a point in which you're going to see the experience of a clinic and things of that sort that I enjoyed. It was, of course, really easy to go back and click on other things if I did want to go back.* [Chicago, 24 years]

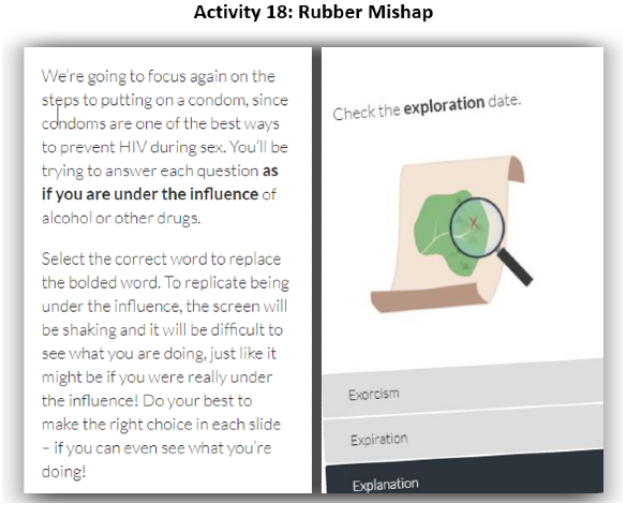
On the other hand, one user noted:

> *I didn't like how it was so linear.* [Chicago, 25 years]

Furthermore, although there were very few negative comments about the gamification aspects of the app, most end users found activity 18 (Figure 2), an interactive feature designed to simulate the experience of intoxication, to be hard to use with the game aspect distracting from the educational purpose. For example, one end user commented:

> *With 18 the thing started shaking so I couldn't read the thing properly. I think they should change that. If you ever have it shake, don't have it shake too much.* [New York, 19 years]

Figure 2. App activity with the screen shaking to simulate being under the influence of drugs or alcohol.



Social Influence

Social influence was applied in this study as the degree to which an individual perceives that important others believe that they would benefit from the use of MyPEEPS app [32]. Subconstructs included (1) *subjective norms* and (2) *social factors*.

Subjective norms are the end users' perception that most people who are important to him think that he should or should not use the MyPEEPS app. Several end users detailed that they would share this mobile app with their peer(s) and/or current sexual partner(s):

> *Me and my lover have to plan. I know I will be looking at the app. When I get home, I am going to tell him all about it.* [New York, 19 yeas]

Conversely, other participants noted that they would not share this app or content with their partner(s) or friend(s) because of embarrassment, implied stigma, or discrimination:

> *I feel like for friends I would be okay telling them. If it's like a partner I might be embarrassed. It seems a little bit like I don't know what I am doing so I have to use this app to help support me make all my decisions. For friends I feel like I'd be more open to share that with them.* [New York, 23 years]

> *Social factors are the end user's perspective of their subjective culture (ie, YMSM) and physical environment [32]. This construct was adapted to include emergent subthemes derived from the analyses that may influence the core determinants of usage intention and behavior, including privacy, age, gender, sexuality, and language.*

Privacy

Participants' reported use of the app was influenced by their physical and social environment and sense of privacy while completing app activities. Privacy was a salient concern because of the sensitivity of sexual health content and associated concern for being "outed":

> *Also, the video was really blunt and direct, and so if I was not listening to that without headphones...it was like not safe for work kind of thing. And if my parents were around, I would feel super uncomfortable. Yeah. And just like making sure that I have privacy. I mean, this stuff is pretty blunt. And I mean, that's good, but*

> *at the same time being aware of who's around is kind of important.* [New York, 23 years]

Age

Most end users found the applicability of the content depending on age of the user. A representative quote is as follows:

> *Yes, and I think specifically younger adults would benefit from this. If I were a college student still in my freshman year, this would be perfect for some sort of training. I think it would be really, effective for young people who are going out into the world, like 18/19 years old...It gives you strategies. I think that a lot of young people probably don't get the sex education they need until college, and not everyone goes to college.* [Chicago, 25 years]

Similarly, some participants commented that the scenarios are most relevant for those still living with parents or caregivers:

> *It wouldn't really be as relevant if I'm in my mid-twenties...because you are in different environments. You may not be living with your parents...and you're navigating different other social circles, and networking in other communities. Some of the scenarios could be transferred over.* [New York, 20 years]

Another participant reflected:

> *I think maybe for the younger crowd; the visual style can maybe be updated, to be a little more grown-up.* [New York, 21 years]

Some end users commented about the app's graphics being tailored to a younger audience, using descriptors such as "juvenile," "cartoonish," and "elementary." One participant commented on the following for improving the interface for young adults aged 19 to 25 years:

> *It's just tricky, because I appreciated the creativity there, but it did feel a little elementary.* [Chicago, 24 years]

Gender/Sexuality

Participants discussed the variability of gender expression among YMSM, which was reflected in the study sample, including participants who described themselves as nonbinary, two-spirit, masculine of center, and feminine of center. One participant noted:

> *Something I also appreciated about the app was it was representative of a lot of different identities that's you could have, and it also talked about the intersection of identities, like gender identity and sexual orientation and race and how all those can play together and affect someone's experience.* [Chicago, 24 years]

Thus, the match of key characteristics in the MyPEEPS Mobile app is likely to resonate with participants.

In addition, several participants noted the desire for a narrative or an avatar who was in the process of "coming out," amid the stages of disclosing one's sexuality and/or gender. For instance:

> *How would you address coming out to parents, or addressing to your friends that I think I might be into guys or girls or look, I have these feelings, like how to navigate those effective communicating strategies.* [New York, 21 years]

Language

Most participants related to language used in the delivery of content:

> *Very sex positive, for sure, using vocabulary that we use; And this is all really very accessibly worded...if that's a word...simply worded, simply put.* [New York, 23 years]

In contrast, several participants did not relate to the use of vernacular language:

> *I felt like, at that moment, some of the vocabulary I thought was intrusive.* [New York, 22 years]

Overall, participants found the language used by avatars during quizzes and in case scenarios to be appropriate.

Discussion

Principal Findings

As mHealth interventions become increasingly available for consumers, it is critical to ensure that mobile technologies are designed and targeted to meet end users' needs [34]. Given that this app was designed for YMSM aged 13 to 18 years, a rigorous evaluation of its usability in YMSM aged 19 to 25 years is important to understand whether it is an acceptable intervention for an older demographic age group.

Overall, end users of this study found the mobile app to be highly usable, as indicated through the survey data (Health-ITUES and PSSUQ), with no major bugs or functional problems reported and no issues with flow of activities. However, the qualitative analyses, based on the UTAUT framework [32,33], provided important insight into nuances of both the strengths and limitations of the app content, including the overall intervention approach, activities, and images for young adult MSM. Thus, emergent qualitative findings provided context and further allowed for in-depth evaluation from the user's perspective.

The basic sexual health information in the app was deemed useful overall, with limitations for those with preexisting sexual health knowledge or sexual experience. The app was perceived to be most useful among those with limited sexual health knowledge and experience. To optimize content for a broader group of young adult MSM, more advanced educational content and social scenarios may need to be added.

Users reported the relative advantage of the MyPEEPS app over standard sexual health approaches because of the use of avatars and animation to aid in the understanding and absorption of content. The use of the avatars also provided a relative advantage, providing material that is salient and relatable to YMSM.

Regarding the UTAUT construct of effort expectancy, the app menu, which was depicted as an urban city street in which

participants advanced along activities sequentially, moving horizontally along the street scene, was largely perceived as both easy to navigate and interesting. Although there were no functional problems with app activities, one activity in which the screen shakes to simulate intoxication was perceived as distracting, with the related text difficult to read, limiting its effectiveness.

Social influence encompassed subconstructs including social factors and subjective norms. Although many participants lauded the "sex positive" nature of the app, basic educational information, and frank language about sexual behavior, they found that these may result in discomfort on the part of the user to complete the activities in front of others because of the stigma associated with the content, reflecting the subjective norms of those around them. Although users are encouraged to use the app in private, this was reinforced given these comments.

The participants were universally enthusiastic about communicating and recognizing different gender presentations and identities in the app and the recognition of intersecting multiple identities, including gender, race/ethnicity, and sexual orientation. Participants recommended a case scenario of an avatar who is in the process of "coming out" disclosure of one's gender and/or sexuality. Although "coming out" is a common thread in the community, the process of disclosing one's sexual and gender identity is dependent on one's circumstance and social context and intersecting social identities [9]. This process may be explained as YMSM nascent from adolescence into young adulthood; an emergence of self-concept and independence is developed through (1) identity formation—awareness and exploration of one's sexual and gender identity—and (2) identity integration—involvement in, comfort with, and disclosure of one's sexual and/or gender identity [42]. As these processes are tangential and nonlinear, the majority of participants responded that MyPEEPS is a foreseeably useful tool for MSM who do not have access to or are impeded by other contextual social barriers to comprehensive sexual health education.

Although there are privacy and security concerns with the use of mobile technology for storing personal health information [43], educational mobile apps such as MyPEEPS may be useful for protecting privacy for highly stigmatized topics such as sexual health for MSM, HIV/AIDS education, gender identity, and mental health [21,23,44]. In fact, MyPEEPS can be very useful for YMSM who are fearful of disclosing their sexuality to family or providers and want to access information about how to protect their health. Thus, MyPEEPS may be particularly useful to those students who do not want to disclose their sexual or gender identities or to those who are living in regional areas with less access to comprehensive sexual health education (ie, varying abstinence policies at the state and school district level) [45].

Limitations

Our study presents some limitations. First, the purposive sample of young adult MSM from urban settings may not generalize to suburban or rural settings and individuals therein. In addition, this study did not include transgender men (assigned female at birth) who have sex with men, and therefore, findings cannot be generalized to that group.

Conclusions

Critical to the uptake and use of eHealth interventions is the rigor in which they are appraised before implementation [26,46,47]. Taking these factors into consideration, this study aimed to rigorously appraise the potential use of the MyPEEPS Mobile intervention, designed for adolescent MSM, to young adult MSM, aged 19 to 25 years, with the goal of future application and/or adaptation to this older group. The perception of usability and acceptability of MyPEEPS Mobile among this demographic of young adult MSM was overall favorable but with key recommendations to improve the applicability of the intervention material for this group with more sexual education and sexual experience.

Acknowledgments

Research reported in this publication was supported by National Institute of Minority and Health Disparities of the National Institutes of Health under award number U01MD011279 and the National Institute of Nursing Research under award number K24NR018621. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. BG was supported by the NIMH Center Grant (P30-MH43520; PI: Robert H Remien, PhD) and NIMH training grant (T32-MH19139 Behavioral Sciences Research in HIV Infection; PI: Theodorus GM Sandfort, PhD).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots from the Male Youth Pursuing Empowerment, Education and Prevention Around Sexuality Activities app.
[DOCX File , 1954 KB - formative_v4i4e17901_app1.docx]

References

1. Centers for Disease Control and Prevention. 2018 Nov. HIV Surveillance Report 2017. Volume 29 URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2017-vol-29.pdf [accessed 2020-03-15]

2. Rodger A, Cambiano V, Bruun T, Vernazza P, Collins S, Degen O, PARTNER Study Group. Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive partner taking suppressive antiretroviral therapy (PARTNER): final results of a multicentre, prospective, observational study. Lancet 2019 Jun 15;393(10189):2428-2438 [FREE Full text] [doi: 10.1016/S0140-6736(19)30418-0] [Medline: 31056293]
3. Bavinton BR, Pinto AN, Phanuphak N, Grinsztejn B, Prestage GP, Zablotska-Manos IB, Opposites Attract Study Group. Viral suppression and HIV transmission in serodiscordant male couples: an international, prospective, observational, cohort study. Lancet HIV 2018 Aug;5(8):e438-e447. [doi: 10.1016/S2352-3018(18)30132-2] [Medline: 30025681]
4. Grohskopf LA, Chillag KL, Gvetadze R, Liu AY, Thompson M, Mayer KH, et al. Randomized trial of clinical safety of daily oral tenofovir disoproxil fumarate among HIV-uninfected men who have sex with men in the United States. J Acquir Immune Defic Syndr 2013 Sep 1;64(1):79-86. [doi: 10.1097/QAI.0b013e31828ece33] [Medline: 23466649]
5. 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]
6. Centers for Disease Control and Prevention. 2019. HIV and Youth URL: https://www.cdc.gov/hiv/group/age/youth/index.html [accessed 2020-03-15]
7. The US Department of Health and Human Services (HHS). Centers for Disease Control and Prevention. 2019. Ending the HIV Epidemic: A Plan for America URL: https://www.cdc.gov/endhiv/index.html [accessed 2020-03-18]
8. Siegler AJ, Mouhanna F, Giler RM, Weiss K, Pembleton E, Guest J, et al. The prevalence of pre-exposure prophylaxis use and the pre-exposure prophylaxis-to-need ratio in the fourth quarter of 2017, United States. Ann Epidemiol 2018 Dec;28(12):841-849 [FREE Full text] [doi: 10.1016/j.annepidem.2018.06.005] [Medline: 29983236]
9. Fields E, Morgan A, Sanders RA. The intersection of sociocultural factors and health-related behavior in lesbian, gay, bisexual, and transgender youth: experiences among young black gay males as an example. Pediatr Clin North Am 2016 Dec;63(6):1091-1106 [FREE Full text] [doi: 10.1016/j.pcl.2016.07.009] [Medline: 27865335]
10. Levy ME, Wilton L, Phillips G, Glick SN, Kuo I, Brewer RA, et al. Understanding structural barriers to accessing HIV testing and prevention services among black men who have sex with men (BMSM) in the United States. AIDS Behav 2014 May;18(5):972-996 [FREE Full text] [doi: 10.1007/s10461-014-0719-x] [Medline: 24531769]
11. Marcell AV, Morgan AR, Sanders R, Lunardi N, Pilgrim NA, Jennings JM, et al. The socioecology of sexual and reproductive health care use among young urban minority males. J Adolesc Health 2017 Apr;60(4):402-410 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.11.014] [Medline: 28065520]
12. Stahlman S, Lyons C, Sullivan PS, Mayer KH, Hosein S, Beyrer C, et al. HIV incidence among gay men and other men who have sex with men in 2020: where is the epidemic heading? Sex Health 2017 Feb;14(1):5-17. [doi: 10.1071/SH16070] [Medline: 27491699]
13. HIV - Federal Government. National HIV/AIDS Strategy: Updated to 2020 URL: https://www.hiv.gov/federal-response/national-hiv-aids-strategy/nhas-update [accessed 2020-03-15]
14. Sullivan PS, Driggers R, Stekler JD, Siegler A, Goldenberg T, McDougal SJ, et al. Usability and acceptability of a mobile comprehensive HIV prevention app for men who have sex with men: a pilot study. JMIR Mhealth Uhealth 2017 Mar 9;5(3):e26 [FREE Full text] [doi: 10.2196/mhealth.7199] [Medline: 28279949]
15. Bauermeister JA, Golinkoff JM, Horvath KJ, Hightow-Weidman LB, Sullivan PS, Stephenson R. A multilevel tailored web app-based intervention for linking young men who have sex with men to quality care (get connected): protocol for a randomized controlled trial. JMIR Res Protoc 2018 Aug 2;7(8):e10444 [FREE Full text] [doi: 10.2196/10444] [Medline: 30072358]
16. Bauermeister JA, Tingler RC, Demers M, Harper GW. Development of a tailored HIV prevention intervention for single young men who have sex with men who meet partners online: protocol for the myDEx project. JMIR Res Protoc 2017 Jul 19;6(7):e141 [FREE Full text] [doi: 10.2196/resprot.7965] [Medline: 28724513]
17. Biello KB, Marrow E, Mimiaga MJ, Sullivan P, Hightow-Weidman L, Mayer KH. A mobile-based app (MyChoices) to increase uptake of HIV testing and pre-exposure prophylaxis by young men who have sex with men: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2019 Jan 7;8(1):e10694 [FREE Full text] [doi: 10.2196/10694] [Medline: 30617042]
18. Ignacio M, Garofalo R, Pearson C, Kuhns LM, Bruce J, Batey DS, et al. Pilot feasibility trial of the MyPEEPS mobile app to reduce sexual risk among young men in 4 cities. JAMIA Open 2019 Jul;2(2):272-279 [FREE Full text] [doi: 10.1093/jamiaopen/ooz008] [Medline: 31294422]
19. LeGrand S, Knudtson K, Benkeser D, Muessig K, Mcgee A, Sullivan PS, et al. Testing the efficacy of a social networking gamification app to improve pre-exposure prophylaxis adherence (P3: Prepared, Protected, emPowered): protocol for a randomized controlled trial. JMIR Res Protoc 2018 Dec 18;7(12):e10448 [FREE Full text] [doi: 10.2196/10448] [Medline: 30563818]
20. Liu A, Coleman K, Bojan K, Serrano PA, Oyedele T, Garcia A, et al. Developing a mobile app (LYNX) to support linkage to HIV/sexually transmitted infection testing and pre-exposure prophylaxis for young men who have sex with men: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Jan 25;8(1):e10659 [FREE Full text] [doi: 10.2196/10659] [Medline: 30681964]

21. Muessig KE, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman LB. A systematic review of recent smartphone, internet and web 2.0 interventions to address the HIV continuum of care. Curr HIV/AIDS Rep 2015 Mar;12(1):173-190 [FREE Full text] [doi: 10.1007/s11904-014-0239-3] [Medline: 25626718]
22. Sullivan PS, Grey JA, Rosser BR. Emerging technologies for HIV prevention for MSM: what we have learned, and ways forward. J Acquir Immune Defic Syndr 2013 Jun 1;63(Suppl 1):S102-S107 [FREE Full text] [doi: 10.1097/QAI.0b013e3182949e85] [Medline: 23673879]
23. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. JMIR Mhealth Uhealth 2019 Apr 16;7(4):e11223 [FREE Full text] [doi: 10.2196/11223] [Medline: 30990458]
24. Hidalgo MA, Kuhns LM, Hotton AL, Johnson AK, Mustanski B, Garofalo R. The MyPEEPS randomized controlled trial: a pilot of preliminary efficacy, feasibility, and acceptability of a group-level, HIV risk reduction intervention for young men who have sex with men. Arch Sex Behav 2015 Feb;44(2):475-485 [FREE Full text] [doi: 10.1007/s10508-014-0347-6] [Medline: 25135064]
25. Schnall R, Kuhns LM, Hidalgo MA, Powell D, Thai J, Hirshfield S, et al. Adaptation of a group-based HIV RISK reduction intervention to a mobile app for young sexual minority men. AIDS Educ Prev 2018 Dec;30(6):449-462 [FREE Full text] [doi: 10.1521/aeap.2018.30.6.449] [Medline: 30966769]
26. Cho H, Powell D, Pichon A, Thai J, Bruce J, Kuhns LM, et al. A mobile health intervention for HIV prevention among racially and ethnically diverse young men: usability evaluation. JMIR Mhealth Uhealth 2018 Sep 7;6(9):e11450 [FREE Full text] [doi: 10.2196/11450] [Medline: 30194060]
27. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Nov 21;18(1):148 [FREE Full text] [doi: 10.1186/s12874-018-0594-7] [Medline: 30463515]
28. Brown W, Yen P, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
29. Fruhling AL, Lee S. Assessing the Reliability, Validity and Adaptability of PSSUQ. In: Proceedings of the 11th Americas Conference on Information Systems conference. 2005 Presented at: AMCIS'05; August 11-14, 2005; Omaha, Nebraska, USA p. 463-488. [doi: 10.1080/10447318.2002.9669130]
30. Lewis J. Psychometric evaluation of the PSSUQ using data from five years of usability studies. Int J Hum-Comput Int 2002;14(3-4):463-488. [doi: 10.1080/10447318.2002.9669130]
31. Venkatesh V, Davis FD, Morris MG. Dead or alive? The development, trajectory and future of technology adoption research. J Assoc Inf Syst 2007;8(4):267-286. [doi: 10.17705/1jais.00120]
32. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
33. Venkatesh V, Thong JY, Xu X. Unified theory of acceptance and use of technology: a synthesis and the road ahead. J Assoc Inf Syst 2016;17(5):328-376. [doi: 10.17705/1jais.00428]
34. Chang IC, Hwang HG, Hung WF, Li YC. Physicians' acceptance of pharmacokinetics-based clinical decision support systems. Expert Syst Appl 2007 Aug;33(2):296-303. [doi: 10.1016/j.eswa.2006.05.001]
35. Sambasivan M, Esmaeilzadeh P, Kumar N, Nezakati H. Intention to adopt clinical decision support systems in a developing country: effect of physician's perceived professional autonomy, involvement and belief: a cross-sectional study. BMC Med Inform Decis Mak 2012 Dec 5;12:142 [FREE Full text] [doi: 10.1186/1472-6947-12-142] [Medline: 23216866]
36. Charmaz K. Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. Thousand Oaks, CA: Sage Publications Ltd; 2006.
37. Koch T, Harrington A. Reconceptualizing rigour: the case for reflexivity. J Adv Nurs 1998 Oct;28(4):882-890. [doi: 10.1046/j.1365-2648.1998.00725.x] [Medline: 9829678]
38. Creswell JW, Poth CN. Qualitative Inquiry And Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: Sage Publications Inc; 2016.
39. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
40. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
41. Saldana J. Fundamentals Of Qualitative Research. New York: Oxford University Press; 2011.
42. Rosario M, Schrimshaw EW, Hunter J. Different patterns of sexual identity development over time: implications for the psychological adjustment of lesbian, gay, and bisexual youths. J Sex Res 2011 Jan;48(1):3-15 [FREE Full text] [doi: 10.1080/00224490903331067] [Medline: 19916104]
43. Schnall R, Higgins T, Brown W, Carballo-Dieguez A, Bakken S. Trust, perceived risk, perceived ease of use and perceived usefulness as factors related to mHealth technology use. Stud Health Technol Inform 2015;216:467-471 [FREE Full text] [Medline: 26262094]

44. Holloway IW, Winder TJ, Lea CH, Tan D, Boyd D, Novak D. Technology use and preferences for mobile phone-based HIV prevention and treatment among black young men who have sex with men: exploratory research. JMIR Mhealth Uhealth 2017 Apr 13;5(4):e46 [FREE Full text] [doi: 10.2196/mhealth.6436] [Medline: 28408360]
45. Santelli JS, Kantor LM, Grilo SA, Speizer IS, Lindberg LD, Heitel J, et al. Abstinence-only-until-marriage: an updated review of US policies and programs and their impact. J Adolesc Health 2017 Sep;61(3):273-280 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.05.031] [Medline: 28842065]
46. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: mixed methods approach. JMIR Mhealth Uhealth 2019 Apr 26;7(4):e11656 [FREE Full text] [doi: 10.2196/11656] [Medline: 31025951]
47. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]

Abbreviations

Health-ITUES: Health Information Technology Usability Evaluation Scale
IRB: Institutional Review Board
mHealth: mobile health
MSM: men who have sex with men
MyPEEPS: Male Youth Pursuing Empowerment, Education and Prevention around Sexuality
PrEP: pre-exposure prophylaxis
PSSUQ: Post-Study System Usability Questionnaire
STI: sexually transmitted infection
UTAUT: Unified Theory of Acceptance and Use of Technology
YMSM: young men who have sex with men

Edited by G Eysenbach; submitted 27.01.20; peer-reviewed by J Schensul, TH Kwan; comments to author 18.02.20; revised version received 20.02.20; accepted 29.02.20; published 07.04.20.

Please cite as:
Gannon B, Davis R, Kuhns LM, Rodriguez RG, Garofalo R, Schnall R
A Mobile Sexual Health App on Empowerment, Education, and Prevention for Young Adult Men (MyPEEPS Mobile): Acceptability and Usability Evaluation
JMIR Form Res 2020;4(4):e17901
URL: https://formative.jmir.org/2020/4/e17901
doi:10.2196/17901
PMID:32254043

Original Paper

Optimizing Child Nutrition Education With the Foodbot Factory Mobile Health App: Formative Evaluation and Analysis

Jacqueline Marie Brown[1], MHSc; Robert Savaglio[2], BIT; Graham Watson[2], BIT; Allison Kaplansky[3], MSc; Ann LeSage[4], PhD; Janette Hughes[4], PhD; Bill Kapralos[2], PhD; JoAnne Arcand[1], PhD, RD

[1]Faculty of Health Sciences, University of Ontario Institute of Technology, Oshawa, ON, Canada

[2]Faculty of Business & Information Technology, University of Ontario Institute of Technology, Oshawa, ON, Canada

[3]Faculty of Community Services, Ryerson University, Toronto, ON, Canada

[4]Faculty of Education, University of Ontario Institute of Technology, Oshawa, ON, Canada

Corresponding Author:
JoAnne Arcand, PhD, RD
Faculty of Health Sciences
University of Ontario Institute of Technology
2000 Simcoe Street North
Science Building Rm 3016
Oshawa, ON, L1G 0C5
Canada
Phone: 1 (905) 7218668 ext 3796
Fax: 1 (905) 7213179
Email: joanne.arcand@uoit.ca

Abstract

Background: Early nutrition interventions to improve food knowledge and skills are critical in enhancing the diet quality of children and reducing the lifelong risk of chronic disease. Despite the rise of mobile health (mHealth) apps and their known effectiveness for improving health behaviors, few evidence-based apps exist to help engage children in learning about nutrition and healthy eating.

Objective: This study aimed to describe the iterative development and user testing of *Foodbot Factory*, a novel nutrition education gamified app for children to use at home or in the classroom and to present data from user testing experiments conducted to evaluate the app.

Methods: An interdisciplinary team of experts in nutrition, education (pedagogy), and game design led to the creation of *Foodbot Factory*. First, a literature review and an environmental scan of the app marketplace were conducted, and stakeholders were consulted to define the key objectives and content of *Foodbot Factory*. Dietitian and teacher stakeholders identified priority age groups and learning objectives. Using a quasi-experimental mixed method design guided by the Iterative Convergent Design for Mobile Health Usability Testing approach, five app user testing sessions were conducted among students (ages 9-12 years). During gameplay, engagement and usability were assessed via direct observations with a semistructured form. After gameplay, qualitative interviews and questionnaires were used to assess user satisfaction, engagement, usability, and knowledge gained.

Results: The environmental scan data revealed that few evidence-based nutrition education apps existed for children. A literature search identified key nutrients of concern for Canadian children and techniques that could be incorporated into the app to engage users in learning. *Foodbot Factory* included characters (2 scientists and Foodbots) who initiate fun and engaging dialogue and challenges (minigames), with storylines incorporating healthy eating messages that align with the established learning objectives. A total of five modules were developed: drinks, vegetables and fruit, grain foods, animal protein foods, and plant protein foods. Seven behavior change techniques and three unique gamified components were integrated into the app. Data from each user testing session were used to inform and optimize the next app iteration. The final user testing session demonstrated that participants agreed that they wanted to play *Foodbot Factory* again (12/17, 71%), that the app is easy to use (12/17, 71%) and fun (14/17, 88%), and that the app goals were clearly presented (15/17, 94%).

Conclusions: *Foodbot Factory* is an engaging and educational mHealth intervention for the Canadian public that is grounded in evidence and developed by an interdisciplinary team of experts. The use of an iterative development approach is a demonstrated

method to improve engagement, satisfaction, and usability with each iteration. Children find *Foodbot Factory* to be fun and easy to use, and can engage children in learning about nutrition.

(JMIR Form Res 2020;4(4):e15534) doi:10.2196/15534

KEYWORDS

mHealth; children; child nutrition sciences; mobile apps; health education

Introduction

Background

Poor diet quality is a leading risk factor for chronic disease, such that the risk of morbidity and mortality from a low-quality diet now surpasses the risk associated with tobacco exposure [1]. Globally, the greatest dietary risk factors are high intakes of sodium and low intakes of whole grains, fruits, and vegetables [1]. In Canada, 92% of children consume excess sodium, and on average, children do not consume the recommended servings of whole grains, vegetables, and fruits, findings that are consistent with most other developed countries [2,3]. These dietary risk factors are extremely relevant for children as they require high-quality diets for growth, development, and academic success [4]. Healthy eating behaviors are established during childhood; therefore, it is important to optimize the adoption of these behaviors early in life so that they will be more likely to be sustained into adulthood [5].

Mobile health (mHealth) innovations present novel opportunities to address public health challenges [6]. Systematic reviews and individual studies demonstrate that mHealth interventions improve a variety of dietary and health-related outcomes, such as nutrition knowledge, overall eating patterns, fruit and vegetable intake, intake of nutrients of public health concern (ie, dietary sodium), body weight, blood pressure, and blood cholesterol levels across various populations and age groups, including children and adolescents [7-14]. In addition to health benefits, mHealth interventions are a promising and an innovative way to facilitate nutrition education as children and adolescents increasingly have access to mobile devices at home and school. In Canada, 47% of children aged 0 to 11 years and 80% of adolescents aged 12 to 17 years own a mobile device [15]. However, few evidence-based mHealth nutrition apps for children are publicly available on the app marketplace [16], and there are currently no known apps available to support educators in facilitating nutrition education in the classroom or clinic settings [17,18]. These data point to a significant gap and an opportunity to address an important public health issue in using evidence-based and engaging mHealth nutrition interventions as a way to engage children in learning about healthy eating.

Objective

The aim of this study was to describe the iterative design, development, and testing of an engaging, evidence-based mHealth nutrition education app, *Foodbot Factory*. The *Foodbot Factory* app was developed as a public health intervention to increase the knowledge and awareness about food and nutrition among healthy children and was designed for use at home and in school. This gamified mHealth app was developed by an interdisciplinary team for children aged 9 to 12 years. The overarching objective of *Foodbot Factory* is to improve food and nutrition knowledge among children.

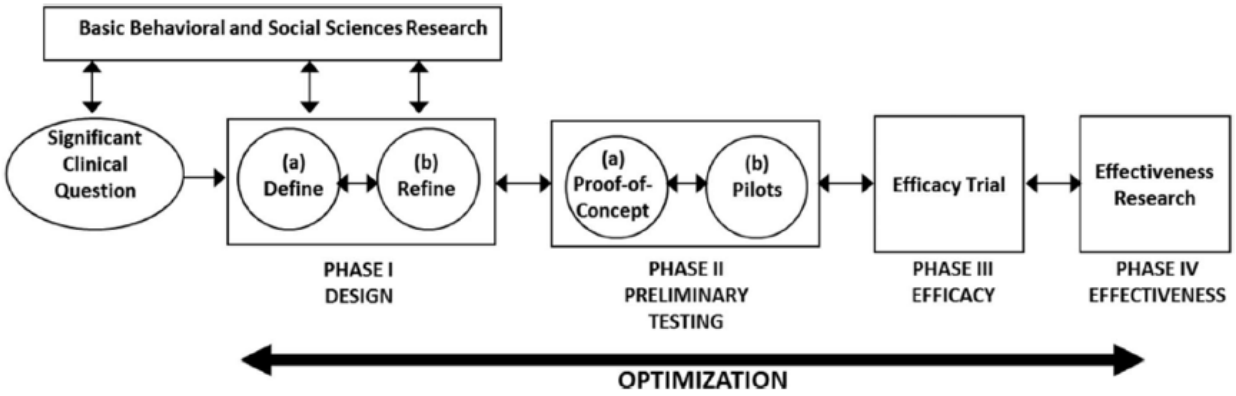
Methods

Study Design

An interdisciplinary team of experts was formed to develop *Foodbot Factory*. This team included experts in nutrition science, behavioral interventions, education, pedagogy and technology, and game design. The team also collaborated with the Office of Nutrition Policy and Promotion at Health Canada who are responsible for improving the health of Canadians through the implementation of nutrition policies and programs.

Foodbot Factory development and testing methodologies were guided by the Obesity-Related Behavioral Intervention Trials (ORBIT) model, which is used for developing behavioral interventions for chronic disease prevention and management (Figure 1) [19]. This model is useful for the development of complex interventions as goals and milestones are established for each phase, and developers are encouraged to revisit previous phases to improve the intervention based on new evidence [19]. The ORBIT model also promotes the integration of behavioral and social sciences research by requiring that the design phase is informed by the latest research, ensuring that high-quality evidence is incorporated into the intervention. The ORBIT model was chosen because it emphasizes the process of developing a behavioral intervention as opposed to the actual content, which will vary across disciplines, populations, and scenarios. It was further considered ideal in the development of *Foodbot Factory* as it emphasizes a data-driven iterative approach to optimize successive iterations of the intervention, so that the intervention will effectively achieve its intended objectives. In addition, the ORBIT model is highly complementary to frameworks used by game developers, such as the Design, Play, and Experience framework [20]. This paper has presented the development and evaluation of *Foodbot Factory* over three phases of the ORBIT model—*Discovery, Phase Ia: Define, and Phase Ib: Refine*[19].

Figure 1. Obesity-Related Behavioral Intervention Trial model.



Discovery Phase

During the *Discovery* phase, the central research question was defined, and educational and behavioral strategies to effectively engage children in learning about nutrition were identified [19]. This phase of development involved background research in behavioral sciences, pedagogy, mHealth app development, and nutritional sciences to identify important points for the intervention and effective methods and techniques for delivery. Background research included a systematic literature search of the PubMed database, conducted in October 2017, to identify peer-reviewed research of mHealth nutrition interventions targeted at a child population. A scoping review was also conducted to identify relevant systematic reviews and literature on the dietary habits of Canadian children. An additional environmental scan of existing mHealth resources was conducted to establish a baseline understanding of the learning tools available for nutrition education [16].

Phase I: Design

The overall goal of *Phase I: Design* was to identify the essential features of the *Foodbot Factory* app, which included identifying the goals and structure and *tinkering* with the nutrition content and educational strategies to refine and improve upon it [19]. This phase required insights on the nutrition education priorities for the target audience, alignment with behavioral theory, and iterative development and testing of *Foodbot Factory*.

Phase Ia: Define

For the *Define* phase, three key components were outlined based on the *Discovery Phase* findings: nutrition content, gamification, and behavior change techniques (BCTs) [19]. Key stakeholders were consulted from February to May 2017 to refine the (1) target age group, (2) isolate the highest priority nutrition education issues, and (3) identify the support needed to implement nutrition curricula. Nutrition-focused stakeholders included public health dietitians who were contacted through Dietitians of Canada, Ontario Dietitians in Public Health, and Health Canada. We also consulted with teachers at a local school board. Stakeholders provided input via a Web-based questionnaire that consisted of 22 items on 7-point Likert scale questions. Open-ended questions were also included to capture additional narrative feedback.

Specific learning objectives incorporating the curriculum and public health priorities were defined based on data and feedback from stakeholders. Considering published evidence on best practices for mHealth apps targeted at children, it was decided that learner engagement with the educational content in *Foodbot Factory* would be enhanced by including (1) gamification, which refers to the application of game design elements to engage and motivate users [21], and (2) evidence-based BCTs, which are "the observable, replicable, and irreducible components of an intervention designed to alter or redirect causal processes that regulate behavior" [22]. The *Foodbot Factory* content was defined and redefined by the development team throughout the iterative development process, incorporating the latest scientific evidence and findings from *Phase Ib* user testing results.

Phase Ib: Refine

The purpose of the *Refine* phase is to ensure that an intervention is robust and efficient at achieving its intended objectives. Therefore, a quasi-experimental mixed method design was used to structure formal user testing sessions to allow for the examination of the user experience and impact of the intervention on learning. These outcomes were assessed with each iteration of *Foodbot Factory*, a method recommended by the Iterative Convergent Design for Mobile Health Usability Testing approach [19,23]. This mixed method approach provided the development team with rich multi-faceted data on the users' experience, which were then used to inform subsequent iterations of the intervention. The *Refine* phase included five iterative user testing sessions with the target audience. As various aspects of *Foodbot Factory* were defined in *Phase Ia*, they were tested with users and refined in *Phase Ib.* Throughout the development process, successive iterations of *Foodbot Factory* moved between *Phases Ia* and *Ib* as the app components were optimized for usability and satisfaction to facilitate engagement in learning.

The five iterative user testing sessions were conducted with student participants in grades 4 to 6 (aged 9-12 years) from a local school board and a science and technology summer camp at Ontario Tech University. User testing sessions were approximately 20 to 30 min in duration and began with an explanation of the study procedures, followed by the collection of informed assent from participants and brief instructions on how to use *Foodbot Factory*. The participants tested the app on an Apple iPad (iOS 12) or a Lenovo tablet (Android 8.1.0) for

10 min, followed by interviews and questionnaires for another 10 min. Each student had their own device; however, participants were able to discuss the app with their peers as they played. Parental consent was obtained before user testing sessions through letters of information and consent forms that were sent home with students before the study. No compensation was provided for parents, students, teachers, or camp counsellors.

The user testing sessions assessed satisfaction, engagement, knowledge, and usability, which were chosen to reflect the core features and components necessary for the app to be successful. Satisfaction refers to the user's subjective experience with *Foodbot Factory* and the appropriateness and suitability of the platform and content for the target audience. Engagement refers to how the target audience interacted with the app, (ie, found the app enjoyable and wanted to play again). Knowledge refers to how well *Foodbot Factory* supported the target audience's acquisition of information about healthy eating. Usability refers to the technical aspects of the app, (ie, easy to use and navigate and clear presentation of content). These metrics were assessed using direct observations, interviews, and questionnaires.

During gameplay, direct observations were used by the research and development team to evaluate usability and engagement with the app. Direct observations included recording notes and completing an internally developed semistructured form where the rater recorded comments about usability and engagement. At this time, perceived engagement with the app dialogue was also rated on a 5-point Likert scale. After gameplay, one-on-one semistructured interviews were conducted with the student participants. The interviews included five questions to assess user satisfaction and knowledge acquisition, with a focus on obtaining feedback for the next iteration of the app, rather than a formal evaluation of learning. Interview guides were created by the development team. The interviews were not audio- or video-recorded, rather verbatim notes were taken. The verbatim notes were reviewed by the development team to identify priority areas for improvement in the next app iteration. Answers to questions regarding nutrition knowledge were summarized based on student responses and have been presented as frequencies. Finally, participants also completed an 8-item (5-point Likert scale) questionnaire to assess engagement and usability. Questions were adapted from a validated questionnaire for adults as no known validated questionnaire was available to assess game engagement in children [24].

Categorical data are reported as frequencies and percentages, and continuous data are reported as means and standard deviations. Likert scale responses ranged from *strongly disagree* to *strongly agree*. For the analysis, *strongly disagree* and *disagree* were pooled to reflect disagreement with a questionnaire item and *strongly agree* and *agree* were pooled to reflect agreement with a questionnaire item. An unpaired 2-tailed *t* test was used to conduct post hoc analyses that compared user satisfaction between boy and girl participants. SPSS (v 25) was used for the statistical analysis. A *P* value of .05 was considered statistically significant.

Parental consent and child assent were obtained before the user testing sessions. In total, five iterative user testing sessions were conducted. An ethics board approval was obtained from both Ontario Tech University and the Durham Catholic District School Board (file numbers: 14426 and 14879).

Results

Discovery

The development process for *Foodbot Factory* began in October 2017, and the research presented in this paper was completed in December 2018. The design process of *Foodbot Factory* was facilitated by an interdisciplinary team who met on a weekly basis to propose ideas, define specific app content, review user testing results, and subsequently refine the app content.

A scoping review identified foods and nutrients of concern for this population and found that the average Canadian child has an inadequate intake of fiber and an excessive intake of saturated fat, sugar, and sodium [25]. Canadian children also consume inadequate amounts of vegetables, fruits, and whole grains and acquire 30% of daily sugar calories from beverages containing free sugars; therefore, the education of these foods and nutrients was considered a priority for the app [3,26].

Existing systematic reviews of mHealth dietary interventions for children and adolescents were identified. These studies indicated that nutrition knowledge acquisition, nutrition-related behaviors, and nutrition-related health outcomes improve with mHealth interventions [7,9], and they highlighted the strengths of gamification as a strategy to implement for child nutrition mHealth apps [27,28]. An additional systematic search, conducted by the *Foodbot Factory* research and development team, focused on identifying other strategies demonstrated to engage children in learning or behavior change through mHealth nutrition apps [10,11,29-35]. Gamification was implemented in two articles and was found to clearly engage children in learning and improving their nutrition knowledge and behaviors and health outcomes [10,31]. In addition, the use of BCTs was also identified as a replicable and potentially effective approach to include in the mHealth app [36]. There was minimal evidence on the most effective BCTs to incorporate for our target audience to achieve the goals of *Foodbot Factory*. Therefore, the development team hypothesized which BCTs would be most effective at facilitating learning among the target audience and evaluated the user experience and learning throughout the iterative development process.

An environmental scan of the app marketplace identified 249 mobile food apps targeted at children [16,37]; however, very few were of high quality, as scored by the Mobile App Rating Scale [38]. Furthermore, gamified apps were more likely to display high-sodium and high-sugar foods and received a higher number of downloads compared with nongamified apps. These results identified gaps in the public app marketplace, where food games are popular with children but are not aligned with the existing dietary recommendations [16,37].

The research conducted in this phase identified research gaps in the development of child nutrition mHealth apps; however, positive findings from the available literature indicated that gamified mHealth apps are a promising avenue to pursue for education. Furthermore, evidence gathered in this phase

demonstrated that mHealth nutrition apps are engaging and can be accessed by the majority of the target population because of the ubiquity of mobile technology in Canada [39]. The *Discovery* phase led to the research question: "Can a mHealth nutrition education app improve nutrition knowledge in children aged 9-12?" This question defined the focus of *Foodbot Factory* to improve users' nutrition knowledge related to foods and nutrients of public health concern [3,25]. With the central research question and strategy identified, the development of *Foodbot Factory* proceeded to *Phase Ia* [19].

Phase Ia: Define

Dietitian and teacher stakeholders (n=21) reported their perceived priorities in relation to the target audience and learning objectives for the app: 81% (17/21) ranked grades 7 to 8 as the first priority age group and 86% (18/21) ranked grades 5 to 6 as the second priority age group for a nutrition education app. The interdisciplinary app development team decided to target students in grades 4 to 6 (aged 9-12 years) for the app as it was believed this age group would be more interested in a gamified nutrition app compared with older students. Overall, 86% (18/21) of stakeholders agreed that increasing water consumption and decreasing sugary beverage consumption was important for this age group, and 57% (12/21) of them also agreed that understanding the nutritional value of foods was important.

In alignment with stakeholder feedback, the app was designed to meet the 2018 Ontario Grade 4 Health and Physical Education Curriculum expectations that students should "identify the key nutrients (eg, fat, carbohydrates, protein, vitamins, minerals) provided by foods and beverages, and describe their importance for growth, health, learning, and physical performance" [40]. This expectation also aligns with the health curriculum expectations from other Canadian provinces and territories, providing teachers across the country with a mobile tool to aid nutrition education in the classroom [41,42].

The goal of *Foodbot Factory* is to improve nutrition knowledge, a component of food skills that has been identified as a gap in the ability of Canadian youth to make healthy eating choices [43]. Although knowledge is often insufficient to initiate dietary behavior change, it is necessary for individuals to change their behavior and is associated with better diet quality in adolescents [44]. Furthermore, the most recent version of Canada's Food Guide (CFG), the leading government resource developed by Health Canada for nutrition guidance, has been disseminated entirely through digital platforms [45]. The CFG is a core component of the school curriculum across Canada. However, to date no tailored resources, messaging, or tools have been created for children and teachers, highlighting a significant gap in nutrition education in Canada [40-42]. As our target audience of 9- to 12-year-old moves into adolescence, they will exercise greater autonomy in their eating decisions. It is hypothesized that the nutrition knowledge obtained from *Foodbot Factory* will empower student users to make healthier dietary choices. Therefore, the goal of this app is to provide the baseline knowledge required to make informed healthy eating choices.

Nutrition Content

To align *Foodbot Factory* with the latest dietary guidelines, the development team consulted with Health Canada to ensure the app's content aligned with the messages in the new 2019 CFG recommendations. The nutrition content in *Foodbot Factory* is structured into five modules, corresponding to the CFG food groups and linking to specific learning objectives [45]. Specific learning objectives were developed for the foods and nutrients relevant for each food grouping, as detailed in Table 1 (Multimedia Appendix 1).

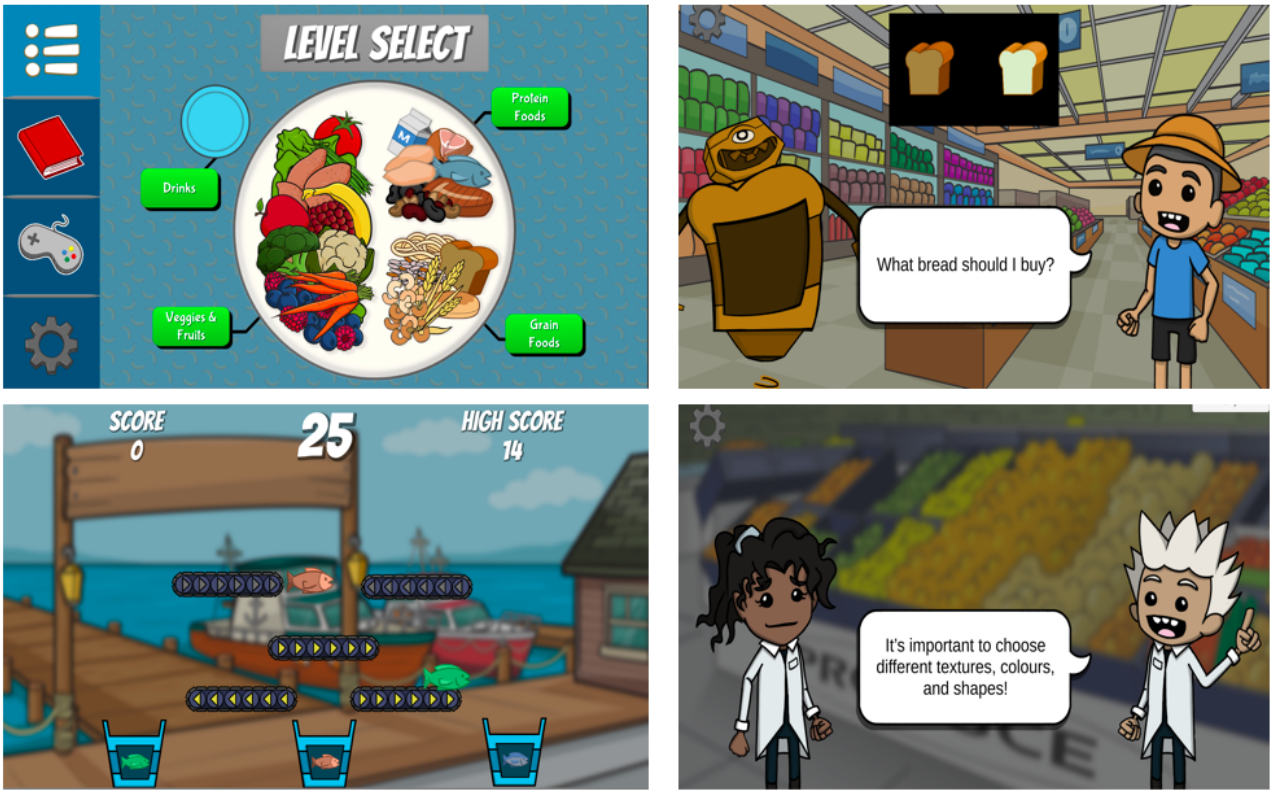
Table 1. Foodbot Factory modules, nutrients of concern, and learning objectives.

Level (food grouping)	Nutrients of concern included	Module learning objectives
Drinks	Sugar	• Recall the best beverage choice for staying hydrated • Describe the health effects of different types of beverages • Recall different types of sugary drinks
Grain foods	Fiber	• Explain the nutritional differences between whole grain and refined grain products • Recall the components of the grain kernel and how grains are refined • Describe why consuming fiber is important for health
Vegetables and fruits	Fiber, vitamins, and minerals	• Recall the amount of vegetables and fruits that should be consumed with a meal • Explain why vegetables and fruits are a healthy choice • Describe which forms of vegetables and fruit are healthiest to consume (ie, canned, frozen, and juice)
Animal protein foods	Protein, saturated fat, and sodium	• Recall that some fats are healthy (unsaturated fats) and unhealthy (saturated fats) • Describe the health effects of excess dietary saturated fat and sodium. • Explain why processed meats should be consumed less often • Describe why fish are a healthy choice
Plant protein foods	Protein and fiber	• Describe what foods are plant protein foods and why they are a healthy choice • Recall that plant protein foods contain fiber

Gamification

The app included three main gamified elements to enhance user interactions, engagement, and motivation to play: quizzes, catching food, and sorting food. *Foodbot Factory* is a story-based app set in a fictional town where residents have nutrition robots (Foodbots) to help them make dietary decisions. The storyline is driven by two nutrition scientists, Robbie and Rebecca, who guide the user on a healthy eating adventure that begins when one of the Foodbots experiences mechanical failure (Figure 2 and Multimedia Appendix 2).

Figure 2. Foodbot Factory screenshots.



Each of the five modules has four unique components: dialogue, gameplay, food quizzes, and food logs (Figure 3). Dialogue communicates the nutrition content and allows for game character development, incorporating elements of humor and banter. A total of three modules contain additional interactive dialogue allowing for the users' choice in the content explored. A Flesch Formula score of 88.9 indicates that the dialogue is easy to read and appropriate for grade 4 readers [46]. Gameplay opportunities are provided intermittently during each module. Gameplay involves collecting or sorting a variety of food items and is varied across the five modules to limit repetition. Food quizzes conclude each module, allowing the users to self-assess their learning. The quizzes allow for self-correction for incorrect responses. The quizzes are interactive, requiring the user to help game characters make healthy eating decisions. The interactive quizzes were designed to simulate real-world dietary choices, which may increase the user's confidence in making real-life healthy eating decisions. The food log (a compendium of foods introduced in the app) is the final component of each module. The food log provides a platform for the user to enhance and expand their learning and explore food nutrition facts, reinforcing information presented in the module.

Figure 3. Foodbot Factory game flow diagram for Drinks module.

Foodbot Factory employs multiple gamification characteristics, as defined by the Taxonomy of Gamification Concepts for Health Apps and highlighted in Multimedia Appendix 3 [47]. In summary, the app utilizes mediated communication where information is delivered through a guided storyline. Rewards are received within the app, and there is no competition with other players or opportunities for collaboration. As *Foodbot Factory* is educational, the content focuses on learning objectives and does not include goal setting, eg, setting a goal for daily water intake. Users receive positive and negative reinforcement during gameplay. Negative reinforcement occurs during gamified components when users do not sort or catch the correct foods or when a question in the food quiz is answered incorrectly. The persuasive intent of *Foodbot Factory* is to positively change user's attitudes and knowledge regarding healthy eating.

Behavior Change Techniques

BCTs were incorporated into the app to illustrate how the knowledge gained can be applied to actual dietary decisions and increase self-efficacy around these decisions. The main BCT categories integrated throughout *Foodbot Factory* are feedback and monitoring, social support, shaping knowledge, natural consequences, and reward and threat [22]. In total, *Foodbot Factory* utilizes seven BCTs across the game (Table 2) [22]. In each module, users have the opportunity to self-assess their knowledge by helping game characters make dietary decisions. The users are then provided with feedback on the accuracy of those dietary decisions. Users also learn about the natural health consequences that result from consuming certain foods and nutrients. At the end of each module, users are rewarded by unlocking new food items in the food log, where they can review additional nutrition information.

Table 2. Behavior change techniques implemented in Foodbot Factory.

Module	Behavior change techniques implemented
Drinks	2.2 Feedback on behavior
	4.1 Instruction on how to perform the behavior
	4.2 Information about antecedents
	5.1 Information about health consequences
	10.3 Nonspecific reward
Grain foods	2.2 Feedback on behavior
	4.1 Instruction on how to perform the behavior
	5.1 Information about health consequences
	10.3 Nonspecific reward
Vegetables and fruits	2.2 Feedback on behavior
	3.1 Social support (unspecified)
	4.1 Instruction on how to perform the behavior
	4.2 Information about antecedents
	5.1 Information about health consequences
	10.3 Nonspecific reward
Plant protein foods	2.2 Feedback on behavior
	4.1 Instruction on how to perform the behavior
	5.1 Information about health consequences
	6.1 Demonstration of behavior
	10.3 Nonspecific reward
Animal protein foods	2.2 Feedback on behavior
	4.1 Instruction on how to perform the behavior
	5.1 Information about health consequences
	10.3 Nonspecific reward

Phase Ib: Refine

User Testing Session 1: Drinks and Grain Foods

The first user testing session assessing the drinks and grain foods modules occurred with student participants in grade 5 (8/12, 67% male) and grade 6 (10/18, 56% male). It was observed that grade 6 students progressed through the app at a faster rate than grade 5 students and seemed less engaged and interested in the app. Overall, 83% (25/30) agreed that the app was fun and 83% (25/30) agreed that the goals of the app were clear. Confusion and frustration in navigating the gameplay portion of the modules was also observed. Only 29% (8/30) of students agreed that the robot was easy to control. Compared with male students, female students appeared to be more interested in the healthy eating content and were more likely to want to play the app again. On the basis of this feedback, the aims for the next iteration of the app were to include more explicit tutorial instructions, improve the game controls and the movement of the robot, increase the use of gamification to enhance engagement, and clearly state the goals of the modules.

User Testing Session 2: Drinks

The next iteration of *Foodbot Factory* was tested with grade 4 student participants (10/20, 50% male). A direct observation revealed that students were more engaged with the dialogue and most students (14/20, 70%) demonstrated a high interest in reading the healthy eating content. Some students (4/20, 21%) showed visual frustration or confusion during game play and asked for help. Those who were confused were also observed skipping quickly through dialogue, indicating that they felt the module had too much reading. On basis of the observation, the students' ability to move the robot improved from the previous iteration. An improvement in the clarity of the app goals and the student's perceived enjoyment of the app was observed with this iteration of the app (Table 3). When students were asked about their favorite part of the game, 30% (6/20) identified the app characters and dialogue humor, whereas 25% (5/20) favored

the *food quiz* gamified component and 15% (3/20) identified learning about nutrition. Students demonstrated learning through using the app: 90% (18/20) of students correctly identified water as the best beverage choice and responded as follows: "Water is needed to stay hydrated."; "Water, because if you're thirsty you can stay hydrated." Some students provided extended commentary, stating "60% of your body is water." and "I learned even more about having pop less." The interviews demonstrated that students understood the key healthy eating messages embedded in the app as they recalled and explained why water is the best beverage choice for health and hydration.

Table 3. Summary of quantitative data for the drinks module (sessions 1-4; session 1 occurred with students in grades 5 to 6. Sessions 2 and 3 occurred with students in grade 4. Session 4 occurred with students in grades 4 to 7.)

User testing session	Strongly disagree and disagree, n (%)	Neutral, n (%)	Strongly agree and agree, n (%)
The app was fun			
Session 1 (N=30)	3 (10)	2 (7)	25 (83)
Session 2 (N=20)	1 (5)	0 (0)	18 (95)
Session 3 (N=14)	0 (0)	1 (7)	13 (93)
Session 4 (N=21)	0 (0)	2 (10)	19 (90)
The goals of the app were clear			
Session 1	4 (13)	1 (3)	25 (83)
Session 2	1 (5)	0 (0)	19 (95)
Session 3	1 (7)	1 (7)	12 (86)
Session 4	0 (0)	1 (5)	20 (95)

User Testing Session 3: Drinks

The next iteration of the drinks module was tested with grade 4 student participants (12/14, 86% male). Most students continued to agree or strongly agree with the statements that the app was fun and the goals of the app were clear (Table 3). The major modification to this iteration was the addition of two gamified *food catches* to introduce more gameplay, which the team hypothesized would increase engagement and interest in the app. However, more students skipped most of the dialogue compared with session 2 (Table 4). A total of 6 students demonstrated visual frustration or confusion, which was related to the *food catch*, as students had skipped through the added tutorial and did not understand what to catch to obtain points. Fewer students provided the correct answer when prompted to choose which drink is the best choice for health, with 57% (8/14) stating water. This may suggest that gamified components need to be carefully incorporated so that educational messages are not lost. Conversely, among the 36% (5/14) of students who incorrectly stated that milk was the best choice, many were able to express their knowledge of water during the interview. The high positive response of milk may reflect students' taste preference for milk over water or baseline knowledge from parents and care providers rather than a lack of knowledge on the importance of water for health and hydration.

Table 4. Summary of observational data for sessions 2 to 5.

User testing session	High reading interest[a], n (%)	Skipped most of the dialogue, n (%)
Session 2 (N=20)	14 (70)	2 (10)
Session 3 (N=14)	7 (51)	2 (14)
Session 4 (N=21)	13 (60)	2 (10)
Session 5 (N=17)	13 (75)	2 (13)

[a]High reading interest was defined as an observational rating of 4 or 5 on a 5-point Likert scale assessing engagement with dialogue content. Skipping dialogue was assessed via an observational note with options of none, some, and most. Students who skipped most of the dialogue were observed flipping quickly through the dialogue and not reading the majority of the content.

User Testing Session 4: Drinks and Grain Foods

The next testing session occurred with summer camp student participants at Ontario Tech University (14/21, 66% male; grades 4-6). The drinks module was unmodified from session 3. The grain foods module was modified to include a new interactive component to teach about different parts of the grain kernel and play an additional *food sort* minigame to increase the variety between the gamified portions of each module. It was hypothesized that engagement would increase by including a variety of interactive components. Observations during gameplay were similar to previous sessions; however, some students appeared to disengage, ie, appearing bored, after playing more than one module. This may explain why observational data showed a slight decline in engagement compared with sessions 2 and 3 where only one app module was tested (Table 4). As in previous sessions, students connected with the game characters and humorous dialogue, with 2 students commenting the following: "That was really fun!" In

this testing session, no students showed visual frustration or confusion, potentially indicating that the usability of the game and clarity of instructions had improved.

Knowledge gained by using *Foodbot Factory* improved from previous sessions. For the drinks module, all students (21/21) replied that water is the best for quenching thirst, with students recognizing that "You need it [water] to survive and stay hydrated" and "60% of the body is made of water." Students also conveyed their knowledge of milk and sugary drinks. Students' comments included the following: "Milk has calcium and Vitamin D and makes bones grow bigger and stronger" and "Pop is a treat like chocolate. You get cavities from the sugar." For the grain foods module, 70% (15/21) of students recalled that whole grain bread has more fiber than white bread, with students stating "Whole grain bread has more fiber and is better for you than white bread because of fiber" and "White bread has less fiber because they remove the bran and the germ." On average, students were able to fulfill the learning objectives for both the drinks and grain foods modules. Importantly, there were considerable improvements on two key metrics of concern. Overall, 90% (19/21) of students agreed with the statement "the app was fun" and 95% (20/21) agreed "the goals of the app were clear" (Table 3). Although observational data suggest that students may disengage when playing multiple modules in one sitting, the self-reported questionnaire data confirm that students found the game engaging and usable.

User Testing Session 5: Vegetables and Fruits and Protein Foods

In this last user testing session among grade 4 student participants (9/17, 55% male), the first iterations of the vegetables and fruits module or the protein foods module were tested. On the basis of the iterative feedback and lessons learned from previous testing sessions, these modules included multiple interactive components. For example, interactive components allowed students to choose the vegetables and fruits to learn about or to cook pasta, and there was added variation between the *food sort* and *food catch* minigames.

Similarly, high levels of engagement as seen in previous user testing sessions were observed (Table 4). Among those who played the vegetables and fruits module (N=8), 75% (6/8) of students recalled a key healthy eating message, with one student stating "They [vegetables & fruits] have lots of vitamins and minerals" and another stating "Most of the stuff [nutrients] helps your body and when you eat it, it helps your body work." By contrast, only 55% (5/9) of students who played the protein foods module (N=9) were able to recall a key healthy eating message. The team concluded that this first iteration of the protein foods module had an ambitious number of key messages and learning outcomes. Therefore, the protein foods module was divided into animal protein foods and plant protein foods to enhance the focus and clarity of healthy eating messages and improve students' ability to meet the learning objectives. Overall, the majority of users still found the modules to be fun, with clear goals, easy to use, and easy to read (Table 5). A total of 71% (12/17) of students agreed with the statement "I want to keep playing," indicating that *Foodbot Factory* is an engaging and acceptable way for students to learn about food, nutrition, and healthy eating (Table 5).

Table 5. Summary of quantitative data for user testing session 5 (N=17)

Question	Strongly disagree and disagree, n (%)	Neutral, n (%)	Strongly agree and agree, n (%)
The goals of the app are clear	1 (6)	0 (0)	16 (94)
The app was easy to use	1 (6)	4 (24)	12 (71)
The words in the app are easy to read	1 (6)	4 (24)	12 (71)
I want to keep playing	0 (0)	5 (29)	12 (71)

Subgroup Analyses

Differences in engagement between male and female student participants were observed from session 1, and it was hypothesized that female students were more engaged with the app than male students. The subgroup analysis (52/87, 60% male) indicated that female students were significantly more likely than male students to agree that they wanted to keep playing (29/35, 83% vs 37/52, 71%; *P*=.03). Trends indicate that female students tended to agree that the app was fun (33/35, 97% vs 45/52, 86%; *P*=.11) and tended to disagree that they were bored (28/35, 82% vs 39/52, 75%; *P*=.09) compared with male students; however, these differences were not statistically significant.

To address the differences in engagement between male and female students, the development team made significant efforts to enhance overall engagement with each iteration of *Foodbot Factory* by further incorporating games and humor as a strategy to better appeal to male students. A voiceover was also added to the app to enhance overall engagement and make the content more accessible for those who may have difficulty reading. Although female students tended to show higher levels of engagement compared with male students, an overall high proportion of male students still found the app fun, and they demonstrated an interest and ability to learn through gameplay, as evidenced by qualitative feedback and observation.

Discussion

Principal Findings

The iterative development and evaluation of *Foodbot Factory* has led to the creation of an evidence-based and engaging mHealth app to help Canadian children learn about healthy eating and nutrition. The interdisciplinary team of experts in nutrition, education, and game design, in collaboration with policymakers (Health Canada), ensured that the app is consistent

with the latest nutrition recommendations, incorporated current advances in gaming and education technology, and can be easily implemented in the classroom. The multimethod user testing data demonstrate that the majority of *Foodbot Factory's* target audience find the app engaging and usable, providing evidence that this public health intervention can help Canadian children learn about nutrition. *Foodbot Factory* has undergone evaluation consistent with *Phase I: Design* in the ORBIT model and is currently being evaluated in a randomized controlled trial as a part of *Phase 2: Preliminary Evaluation* [19]. A strength of this research is the use of an interdisciplinary team of experts to facilitate and participate in development. An interdisciplinary approach is not always implemented when designing electronic health and mHealth interventions [48], but it should be encouraged to ensure that the multiple disciplines involved in developing high-quality digital interventions are represented throughout the development process. Our research and development team worked effectively in a collaborative fashion through weekly scheduled meetings until the app development was fully complete. We further leveraged our strong relationships with partners across disciplines, including governments, nongovernmental organizations, and school boards, to obtain diverse expert opinions and conduct the research presented in this study.

Strengths and Limitations

In comparison with similar studies that have conducted iterative user testing methodologies to develop mHealth interventions, the research presented here has implemented significantly more iterative testing sessions with the target audience. The majority of other app development studies that report user testing often implement various stepwise iterative processes throughout development; however, many only utilize one testing session with their target audience [49-53]. The use of multiple testing sessions for *Foodbot Factory* is a strength of this study as it allowed the research team to improve engagement and usability of the app. Multiple testing sessions also enabled app evaluation with a larger sample size and tailored the content of the app to the diverse needs of *Foodbot Factory's* target audience. Finally, other studies that report on the development of mHealth interventions have not implemented multiple methods in their user testing sessions [51,53]. The use of multiple methods in this study provided the research team with a set of rich and comprehensive data on the user experience, increasing the efficiency and effectiveness of each successive iteration of *Foodbot Factory*. The development of future mHealth apps and interventions would be best informed by iterative testing to ensure that changes implemented to each iteration meet the needs of the target audience and advance the goals of the intervention. Testing should also utilize mixed method data collection, as recommended by the Iterative Convergent Design for Mobile Health Usability Testing [23].

Although there were positive outcomes to this project, we acknowledge that there were limitations. Observation was relied on as a way to assess user engagement with app dialogue content and the rate of progression; however, these measures of user engagement could have been enhanced by collecting quantitative interapp analytic data. Factors that influence the dietary intakes of children are often beyond the control of a young child (eg, socioeconomic status, parental influence, and neighborhood food environment) [54]. However, the knowledge gained from playing *Foodbot Factory* can assist children in making healthier dietary decisions where possible, and the foundational knowledge may stay with the child over the long term. Testing results suggest that student participants may disengage after playing more than one module during a session and that female students may be more engaged than male students. However, these findings relay important information for implementation into the classroom setting, suggesting that learning will be optimized by using the app for short periods over several days. *Foodbot Factory* can be played in a flexible format, providing users with a choice in the number of modules played in a given session. The app has also been designed with multiple gamified components that vary through each module to appeal to a variety of learning and gameplay preferences. Future research with *Foodbot Factory* includes a randomized trial measuring its impact on nutrition knowledge and development of a knowledge translation strategy to ensure maximal reach and adoption among teachers, health professionals, and parents.

In summary, *Foodbot Factory* is a viable and an engaging app that has the potential to improve the nutrition knowledge of Canadian children. The app aligns with the latest dietary recommendations for Canadians and curriculum expectations for most Canadian provinces [40,45]. *Foodbot Factory* utilizes increasingly available technologies, providing a contemporary means for children to learn about nutrition at home and in the classroom. The knowledge gained during childhood may be sustained into the adolescent years, with the intended benefit of improving dietary intake and reducing the risk of chronic diseases associated with poor diet quality.

Acknowledgments

This study was supported by grant funding from the Ontario Ministry of Research and Innovation. The authors would like to acknowledge the hard work of the app developers, Campbell Hamilton, Rebecca Joly, and Alain Sangalang, and those who assisted with conducting our user testing sessions, Hannah Froome, Tayyib Khan, Amina Mahmood, and Amina Siddiqi. They would also like to acknowledge and thank their partners at Health Canada's Office of Nutrition Policy and Promotion, including Dr Alfred Aziz, Emily Bell, and Mabel Pang, who have supported the development of *Foodbot Factory*. This research would not have been possible without the generosity and support of administrators and teachers at the Durham Catholic District School Board and Dr Shelia Rhodes at the Ontario Tech Summer Camp.

Authors' Contributions

JA and JH conceived this work. JB, RS, BK, and JA led the development of the app and established the research design. All authors participated in the iterative development process. JB led the writing of this manuscript and completed data collection and data analysis. AK created and distributed stakeholder surveys. JB, AK, RS, GW, AL, BK, and JA developed data collection tools and the user testing strategy. AK, RS, and GW assisted with data collection. All authors contributed to revising the manuscript. All authors provided the final approval for the manuscript and agreed to be accountable for its contents.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of the Foodbot Factory content.
[DOCX File , 11939 KB - formative_v4i4e15534_app1.docx]

Multimedia Appendix 2
A Foodbot Factory gameplay video.
[MP4 File (MP4 Video), 95747 KB - formative_v4i4e15534_app2.mp4]

Multimedia Appendix 3
Foodbot Factory gamification characteristics.
[DOCX File , 15 KB - formative_v4i4e15534_app3.docx]

References

1. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 May 11;393(10184):1958-1972 [FREE Full text] [doi: 10.1016/S0140-6736(19)30041-8] [Medline: 30954305]
2. Health Canada. Government of Canada. Ottawa, Ontario, Canada: Government of Canada; 2018. Sodium Intake of Canadians in 2017 URL: https://www.canada.ca/en/health-canada/services/publications/food-nutrition/sodium-intake-canadians-2017.html [accessed 2020-02-14]
3. Tugault-Lafleur CN, Black JL. Differences in the quantity and types of foods and beverages consumed by Canadians between 2004 and 2015. Nutrients 2019 Feb 28;11(3):pii: E526 [FREE Full text] [doi: 10.3390/nu11030526] [Medline: 30823448]
4. Basch CE. Healthier students are better learners: high-quality, strategically planned, and effectively coordinated school health programs must be a fundamental mission of schools to help close the achievement gap. J Sch Health 2011 Oct;81(10):650-662. [doi: 10.1111/j.1746-1561.2011.00640.x] [Medline: 21923878]
5. Craigie AM, Lake AA, Kelly SA, Adamson AJ, Mathers JC. Tracking of obesity-related behaviours from childhood to adulthood: a systematic review. Maturitas 2011 Nov;70(3):266-284. [doi: 10.1016/j.maturitas.2011.08.005] [Medline: 21920682]
6. World Health Organization. WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening. Geneva: World Health Organization; 2019.
7. Rose T, Barker M, Maria Jacob C, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. J Adolesc Health 2017 Dec;61(6):669-677 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.05.024] [Medline: 28822682]
8. Brannon EE, Cushing CC. A systematic review: is there an app for that? Translational science of pediatric behavior change for physical activity and dietary interventions. J Pediatr Psychol 2015 May;40(4):373-384. [doi: 10.1093/jpepsy/jsu108] [Medline: 25502745]
9. Schoeppe S, Alley S, van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 7;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
10. Baghaei N, Nandigam D, Casey J, Direito A, Maddison R. Diabetic Mario: designing and evaluating mobile games for diabetes education. Games Health J 2016 Aug;5(4):270-278. [doi: 10.1089/g4h.2015.0038] [Medline: 27304882]
11. de Niet J, Timman R, Bauer S, van den Akker E, Buijks H, de Klerk C, et al. The effect of a short message service maintenance treatment on body mass index and psychological well-being in overweight and obese children: a randomized controlled trial. Pediatr Obes 2012 Jun;7(3):205-219. [doi: 10.1111/j.2047-6310.2012.00048.x] [Medline: 22492669]
12. Thompson D, Bhatt R, Vazquez I, Cullen KW, Baranowski J, Baranowski T, et al. Creating action plans in a serious video game increases and maintains child fruit-vegetable intake: a randomized controlled trial. Int J Behav Nutr Phys Act 2015 Mar 18;12:39 [FREE Full text] [doi: 10.1186/s12966-015-0199-z] [Medline: 25890060]

13. Ali SH, Luo R, Li Y, Liu X, Tang C, Zhang P. Application of mobile health technologies aimed at salt reduction: systematic review. JMIR Mhealth Uhealth 2019 Apr 17;7(4):e13250 [FREE Full text] [doi: 10.2196/13250] [Medline: 30994467]
14. Champion K, Parmenter B, McGowan C, Spring B, Wafford Q, Gardner L, et al. Effectiveness of school-based eHealth interventions to prevent multiple lifestyle risk behaviours among adolescents: a systematic review and meta-analysis. Lancet Digit Health 2019 Sep;1(5):e206-e221 [FREE Full text] [doi: 10.1016/s2589-7500(19)30088-3]
15. eMarketer. 2016 Apr 13. In Canada, Children and Teens Have an Affinity for Digital Tools URL: https://www.emarketer.com/Article/Canada-Children-Teens-Have-Affinity-Digital-Tools/1013816 [accessed 2019-07-01]
16. Nejatinamini S, Debenham B, Prado C, Baracos VE, Chu Q, Wismer WW, et al. Canadian Nutrition Society: Scientific Abstracts from the 10th Annual Scientific Meeting / Société canadienne de nutrition: Résumés scientifiques de la 10 réunion scientifique annuelle. Appl Physiol Nutr Metab 2019 Apr;44(4 (Suppl 1)):S1-S54. [doi: 10.1139/apnm-2019-0152] [Medline: 30983410]
17. Dudley DA, Cotton WG, Peralta LR. Teaching approaches and strategies that promote healthy eating in primary school children: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2015 Feb 25;12:28 [FREE Full text] [doi: 10.1186/s12966-015-0182-8] [Medline: 25889098]
18. Brown JM, Savaglio R, Watson G, LeSage A, Kapralos B, Arcand J. The making of Foodbot Factory: An evidence-based mHealth nutrition education app for children. Canadian Nutrition Society: Scientific Abstracts from the 10th Annual Scientific Meeting / Société canadienne de nutrition: Résumés scientifiques de la 10 réunion scientifique annuelle 2019 Apr;44(4 (Suppl. 1)):S8. [doi: 10.1139/apnm-2019-0152] [Medline: 30983410]
19. Czajkowski SM, Powell LH, Adler N, Naar-King S, Reynolds KD, Hunter CM, et al. From ideas to efficacy: The ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015 Oct;34(10):971-982 [FREE Full text] [doi: 10.1037/hea0000161] [Medline: 25642841]
20. Ferdig RE. Handbook of Research on Effective Electronic Gaming in Education. Hershey, PA: Information Science Reference; 2009.
21. Zichermann G, Cunningham C. Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. First Edition. Sebastopol, Calif: O'Reilly Media; 2011.
22. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
23. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: mixed methods approach. JMIR Mhealth Uhealth 2019 Apr 26;7(4):e11656 [FREE Full text] [doi: 10.2196/11656] [Medline: 31025951]
24. Brockmyer JH, Fox CM, Curtiss KA, McBroom E, Burkhart KM, Pidruzny JN. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. J Exp Soc Psychol 2009;45(4):624-634. [doi: 10.1016/j.jesp.2009.02.016]
25. Health Canada. Government of Canada. Ottawa, Canada: Health Canada; 2012. Do Canadian Adolescents Meet Their Nutrient Requirements Through Food Intake Alone? Ottawa, CA URL: https://www.canada.ca/en/health-canada/services/food-nutrition/food-nutrition-surveillance/health-nutrition-surveys/canadian-community-health-survey-cchs/canadian-adolescents-meet-their-nutrient-requirements-through-food-intake-alone-health-canada-2012.html [accessed 2020-02-14]
26. Langlois K, Garriguet D, Gonzalez A, Sinclair S, Colapinto C. Change in total sugars consumption among Canadian children and adults. Health Rep 2019 Jan 16;30(1):10-19 [FREE Full text] [Medline: 30649778]
27. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: A systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
28. Dias JD, Domingues AN, Tibes CM, Zem-Mascarenhas SH, Fonseca LM. Serious games as an educational strategy to control childhood obesity: a systematic literature review1. Rev Lat Am Enfermagem 2018 Sep 3;26:e3036 [FREE Full text] [doi: 10.1590/1518-8345.2509.3036] [Medline: 30183872]
29. Aflague TF, Boushey CJ, Guerrero RT, Ahmad Z, Kerr DA, Delp EJ. Feasibility and use of the mobile food record for capturing eating occasions among children ages 3-10 years in Guam. Nutrients 2015 Jun 2;7(6):4403-4415 [FREE Full text] [doi: 10.3390/nu7064403] [Medline: 26043037]
30. Nollen NL, Hutcheson T, Carlson S, Rapoff M, Goggin K, Mayfield C, et al. Development and functionality of a handheld computer program to improve fruit and vegetable intake among low-income youth. Health Educ Res 2013 Apr;28(2):249-264 [FREE Full text] [doi: 10.1093/her/cys099] [Medline: 22949499]
31. Nollen NL, Mayo MS, Carlson SE, Rapoff MA, Goggin KJ, Ellerbeck EF. Mobile technology for obesity prevention: a randomized pilot study in racial- and ethnic-minority girls. Am J Prev Med 2014 Apr;46(4):404-408 [FREE Full text] [doi: 10.1016/j.amepre.2013.12.011] [Medline: 24650843]
32. Panizza CE, Boushey CJ, Delp EJ, Kerr DA, Lim E, Gandhi K, et al. Characterizing early adolescent plate waste using the mobile food record. Nutrients 2017 Jan 26;9(2):pii: E93 [FREE Full text] [doi: 10.3390/nu9020093] [Medline: 28134757]

33. Pedersen S, Grønhøj A, Thøgersen J. Texting your way to healthier eating? Effects of participating in a feedback intervention using text messaging on adolescents' fruit and vegetable intake. Health Educ Res 2016 Apr;31(2):171-184. [doi: 10.1093/her/cyv104] [Medline: 26850061]
34. Schiel R, Kaps A, Bieber G. Electronic health technology for the assessment of physical activity and eating habits in children and adolescents with overweight and obesity IDA. Appetite 2012 Apr;58(2):432-437. [doi: 10.1016/j.appet.2011.11.021] [Medline: 22155072]
35. Woolford SJ, Barr KL, Derry HA, Jepson CM, Clark SJ, Strecher VJ, et al. OMG do not say LOL: obese adolescents' perspectives on the content of text messages to enhance weight loss efforts. Obesity (Silver Spring) 2011 Dec;19(12):2382-2387 [FREE Full text] [doi: 10.1038/oby.2011.266] [Medline: 21869762]
36. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
37. Brown JM, Siddiqi A, Froome H, Arcand J. What's cooking? A content and quality analysis of food preparation mobile applications (P16-050-19). Curr Dev Nutr 2019 Jun;3(Supplement_1). [doi: 10.1093/cdn/nzz050.p16-050-19]
38. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
39. Tate EB, Spruijt-Metz D, O'Reilly G, Jordan-Marsh M, Gotsis M, Pentz MA, et al. mHealth approaches to child obesity prevention: successes, unique challenges, and next directions. Transl Behav Med 2013 Dec;3(4):406-415 [FREE Full text] [doi: 10.1007/s13142-013-0222-3] [Medline: 24294329]
40. Ontario Ministry of Education. The Ontario Curriculum: Health and Physical Education. Toronto, Ontario: Queen's Printer for Ontario; 2019.
41. British Columbia Ministry of Education. BC's New Curriculum. Victoria, BC: Government of British Columbia; 2016. Physical and Health Education 4 URL: https://curriculum.gov.bc.ca/curriculum/physical-health-education/4 [accessed 2020-02-02]
42. Saskatchewan Ministry of Education. Edonline. 2010. Saskatchewan Curriculum URL: https://www.edonline.sk.ca/webapps/moe-curriculum-BBLEARN/CurriculumOutcomeContent?id=49&oc=76303 [accessed 2020-02-14]
43. Ontario FSG. Ontario Food and Nutrition Strategy - A Comprehensive Evidence-Informed Plan for Healthy Food and Food Systems in Ontario. Toronto, Ontario 2017 [FREE Full text]
44. Vaitkeviciute R, Ball LE, Harris N. The relationship between food literacy and dietary intake in adolescents: a systematic review. Public Health Nutr 2015 Mar;18(4):649-658 [FREE Full text] [doi: 10.1017/S1368980014000962] [Medline: 24844778]
45. Health Canada. Canada's Dietary Guidelines for Health Professionals and Policy Makers. Ottawa, Canada: Health Canada; 2019.
46. Ley P, Florio T. The use of readability formulas in health care. Psychol Health Med 1996;1(1):7-28. [doi: 10.1080/13548509608400003]
47. Schmidt-Kraepelin M, Thiebes S, Tran M, Sunyaev A. What's in the Game? Developing a Taxonomy of Gamification Concepts for Health Apps. In: Proceedings of the 51st Hawaii International Conference on System Sciences |. 2018 Presented at: HICSS'18; January 3-6, 2018; Waikoloa Village, Hawaii. [doi: 10.24251/hicss.2018.150]
48. Tamim SR, Grant MM. Exploring how health professionals create eHealth and mHealth education interventions. Educ Technol Res Dev 2016;64(6):1053-1081. [doi: 10.1007/s11423-016-9447-4]
49. Mummah SA, King AC, Gardner CD, Sutton S. Iterative development of Vegethon: a theory-based mobile app intervention to increase vegetable consumption. Int J Behav Nutr Phys Act 2016 Aug 8;13:90 [FREE Full text] [doi: 10.1186/s12966-016-0400-z] [Medline: 27501724]
50. Simons D, de Bourdeaudhuij I, Clarys P, de Cocker K, Vandelanotte C, Deforche B. A smartphone app to promote an active lifestyle in lower-educated working young adults: development, usability, acceptability, and feasibility study. JMIR Mhealth Uhealth 2018 Feb 20;6(2):e44 [FREE Full text] [doi: 10.2196/mhealth.8287] [Medline: 29463491]
51. Gabrielli S, Dianti M, Maimone R, Betta M, Filippi L, Ghezzi M, et al. Design of a mobile app for nutrition education (TreC-LifeStyle) and formative evaluation with families of overweight children. JMIR Mhealth Uhealth 2017 Apr 13;5(4):e48 [FREE Full text] [doi: 10.2196/mhealth.7080] [Medline: 28408361]
52. Tonkin E, Jeffs L, Wycherley TP, Maher C, Smith R, Hart J, et al. A smartphone app to reduce sugar-sweetened beverage consumption among young adults in Australian remote indigenous communities: design, formative evaluation and user-testing. JMIR Mhealth Uhealth 2017 Dec 12;5(12):e192 [FREE Full text] [doi: 10.2196/mhealth.8651] [Medline: 29233803]
53. Duff O, Walsh D, Malone S, McDermott L, Furlong B, O'Connor N, et al. MedFit app, a behavior-changing, theoretically informed mobile app for patient self-management of cardiovascular disease: user-centered development. JMIR Form Res 2018 Apr 27;2(1):e8 [FREE Full text] [doi: 10.2196/formative.9550] [Medline: 30684426]
54. Story M, Kaphingst KM, Robinson-O'Brien R, Glanz K. Creating healthy food and eating environments: policy and environmental approaches. Annu Rev Public Health 2008;29:253-272. [doi: 10.1146/annurev.publhealth.29.020907.090926] [Medline: 18031223]

Abbreviations

BCT: behavior change technique
CFG: Canada's Food Guide
mHealth: mobile health
ORBIT: Obesity-Related Behavioral Intervention Trials

Edited by G Eysenbach; submitted 19.07.19; peer-reviewed by K Usop, S Smith, L Kramer, K Mateo; comments to author 30.09.19; revised version received 25.11.19; accepted 27.01.20; published 17.04.20.

Please cite as:
Brown JM, Savaglio R, Watson G, Kaplansky A, LeSage A, Hughes J, Kapralos B, Arcand J
Optimizing Child Nutrition Education With the Foodbot Factory Mobile Health App: Formative Evaluation and Analysis
JMIR Form Res 2020;4(4):e15534
URL: http://formative.jmir.org/2020/4/e15534/
doi:10.2196/15534
PMID:32301743

Original Paper

Using Crowdsourcing to Develop a Peer-Led Intervention for Safer Dating App Use: Pilot Study

William Chi Wai Wong[1*], MD, MPH, FRCGP; Lin Song[1*], PhD; Christopher See[2], PGCE, PhD; Stephanie Tze Hei Lau[1], BA, MPH; Wai Han Sun[1], BSocSc, MPhil; Kitty Wai Ying Choi[1], BSocSc (Hons); Joseph Tucker[3], MD, PhD

[1]Department of Family Medicine and Primary Care, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong

[2]School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong

[3]Division of Infectious Disease, Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[*]these authors contributed equally

Corresponding Author:
Lin Song, PhD
Department of Family Medicine and Primary Care
Li Ka Shing Faculty of Medicine
University of Hong Kong
2/F, Ap Lei Chau Clinic
161 Ap Lei Chau Main Street, Ap Lei Chau
Hong Kong
Phone: 852 67456763
Email: songlinhk@gmail.com

Abstract

Background: Smartphone-based dating apps are rapidly transforming how people seek potential sexual and romantic partners. However, they can also increase the risk of unsafe sexual behaviors, harassment, and infringement of personal privacy. Current research on interventions for safer dating app use remains insufficient.

Objective: The goal of this study was to describe the development of an intervention for safer dating app usage using crowdsourcing and peer-led approaches.

Methods: This paper describes the development of an intervention program designed to promote safer dating app use among college students. Crowdsourcing and peer-led approaches were adopted during key stages of the development process. Focus group discussions were held to assess the experience and needs of dating app users. A crowdsourcing contest then solicited ideas for performance objectives for the intervention. These objectives were grouped to further identify practical strategies. A one-day intensive workshop was subsequently held with peer mentors to brainstorm ideas for the production of creative interventional materials. The intervention programs were produced and tested in a pilot study. The app's effectiveness will be evaluated in a cluster randomized controlled trial.

Results: The intervention program consists of a risk assessment tool, a first-person scenario game, and four short videos. The risk assessment tool, comprised of 14 questions, will give the participant a score to determine their level of risk of adverse events when using dating apps. The scenario game is a first-person simulation game where the players are presented with choices when faced with different scenarios. The short videos each last 2-4 minutes, with points of discussion aimed at addressing the risks of using dating apps. The programs were piloted and were found to be relatable and helpful when further modifications were made.

Conclusions: Potential challenges identified during the development process included data management and analysis, sustaining peer mentors' interests and participation, and balancing between providing more information and perpetuating social stigma around dating app use. By integrating new approaches, such as crowdsourcing and the peer-led approach, in developing an intervention for safer dating app use, our development process provides a viable model for developing future interventions to address the risks associated with dating app use.

(JMIR Form Res 2020;4(4):e12098) doi:10.2196/12098

KEYWORDS

dating apps; intervention mapping; crowdsourcing; peer-led approach; sexual health

Introduction

Background and Motivation

The popularization of smartphone-based dating apps has drastically changed dating patterns and behaviors among young adults across the globe. Researchers from various countries have reported high percentages of dating app use among young adults of different genders and sexual orientations. In the United States, for instance, approximately 40% of sampled young heterosexual men and women used dating apps in 2017 [1], and the percentage for men who have sex with men is significantly higher at around 69% [2]. In Hong Kong, a 2016 study found an equally high proportion (52.9%) of dating app users among both heterosexual and homosexual college students [3].

By facilitating convenient interpersonal exchanges, dating apps broaden one's personal and romantic networks and bring new freedoms, opportunities, and pleasures [4]. However, they also engender a range of potential risks in personal safety, privacy, and sexual health. A high correlation between dating app use and multiple sexual partners has been identified among lesbian, gay, transgender, and bisexual populations, which is an alarming indicator for sexual health–related risks, such as HIV or hepatitis C infection [5]. Some studies have also raised concerns about the lack of awareness around the protection of personal information on dating apps [6,7], highlighting potential privacy and safety risks.

Currently, resources are insufficient for dating app users in Hong Kong or elsewhere to understand and manage these risks. A review of a large number of smartphone-based dating apps found that the majority contained no information about sexual health or sexual assault, so users had low exposure to sexual health and personal safety content [8]. To address these issues, our research team engaged dating app users in Hong Kong to develop intervention programs for promoting safer dating app usage among college students. We implemented innovative approaches, including crowdsourcing and the peer-led approach, in various stages of the development of the intervention.

Study Aims

This study offers a close, step-by-step investigation of using crowdsourcing and peer-led approaches to develop an intervention for safer dating app use. It aims to propose a viable model for developing similar dating app use interventions in the future. Potential challenges in the intervention development have been identified, and the effectiveness of adopting hybridized approaches was explored. This study also reports on the findings of a pilot study for the intervention.

Methods

Needs Assessment

The needs assessment is a systematic study of the discrepancy between what is and what should be in a group and situation of interest [9]. Data for the assessment can be collected through various channels, including a literature review, brainstorming, and focus groups conducted with the target population. Findings from the needs assessment are used to specify intended changes in individual behaviors and environmental conditions.

Since the popularity of dating apps is a relatively new phenomenon, literature on its risks remains limited. We, therefore, decided to conduct focus group discussions and engage app users in discussing their safety strategies to elicit insights into their behaviors, motivations, and perceptions when using dating apps. We recruited participants through a Hong Kong–based organization, "StickyRiceLove," which specializes in peer-led, online, sexual health education. Inclusion criteria for participants included: experience with using a dating app, being between the ages of 18-35 years old, and able to speak Cantonese.

Each focus group was purposively constructed to include people of different genders, ages, and sexual orientations to facilitate heterogeneity in each focus group discussion and provide potential contrasts in discussions [10]. Focus group discussions of 6-8 people were conducted until data saturation was achieved, resulting in four groups (N=25 participants) in total. More demographic information can be found in Table 1.

Table 1. Needs assessment participants' demographics (N=25).

Demographics	n
Gender	
Male	15
Female	10
Age group (years)	
18-20	4
21-23	4
24-26	7
27-29	4
Over 30	6
Sexual Orientation	
Heterosexual	17
Homosexual	6
Queer	1
Bisexual	1
App use experience (years)	
0-1	7
2-3	8
4+	10
Instances of meeting friends via the app, n	
0-1	3
2-5	7
6-10	5
>10	9

We used a semistructured interview guide developed from the literature to conduct the discussions (see Multimedia Appendix 1). Open-ended questions on dating app user experience, motivation, beliefs, and behaviors regarding intimacy and risk were used to facilitate discussion. Audio recordings were transcribed and translated from Cantonese to English.

These data were analyzed using thematic analysis, commencing with an open-coding approach of repeated line-by-line reading of the text and condensing of concepts, events, and meanings into representative codes. Two members of the research team (CS, STHL) independently undertook the coding to increase the robustness. Discrepancies between the initial coding were discussed and clarified with input from other members of the research team. Upon further cycles of analysis, themes were elicited from the data by constantly comparing the interview transcripts (see Multimedia Appendix 2).

Notably, users reported uncertainty and worry about potentially risky sexual health behaviors, which spanned from casual sex, multiple partners, or open relationships. They also expressed concerns about data privacy, including fear of blackmail using intimate data from dates or private photo exchanges from apps. Furthermore, we found that dating app users experienced several financial scams, some of which resulted in monetary loss. These findings posed key questions to consider for the intervention: How do we help dating app users in improving their sexual health knowledge, skills, and behaviors? How can we encourage dating app users to practice data safety and privacy while building relationships? How do we alert dating app users to the risks of financial scams and empower them by providing practical strategies? Apart from these questions, the participants also suggested that the experience of stigma and stereotyping remains a problem in the local dating app user community and that the intervention should provide reliable information and promote behavioral skills while avoiding further stigmatization.

Based on the findings from the focus group, we determined four prominent domains for intervention: sexual health, the security of personal privacy, sexual harassment and assault, and monetary scams. These domains are generally in line with existing literature on dating apps, although some of them are more specific to the Hong Kong context. The information gathered from the focus groups substantially contributed to the needs assessment. The research team then reviewed relevant theoretical and empirical literature and produced five behavioral outcome goals, as outlined in Table 2. We then further identified the determinants for each goal to pave the way for a more systematic approach in developing practical strategies in the next step.

Table 2. Behavioral outcomes and determinants.

Behavioral Outcomes	Examples (Determinants)
Users will understand the general benefits and risks associated with dating apps.	• Name benefits and risks of dating apps (Awareness).
Users will evaluate relationship development and follow safe dating practices. Users will know the importance of safe sex and use condoms correctly and consistently when having sex.	• Recognize risks of dating with people met on apps (Risk Perception). • Select safe dating venues (Skills). • Confidence in communicating personal limits and intentions on dates (Self-efficacy). • Be aware of sexual health risks (Risk perception). • Acknowledge condom use as effective protection against STIs[a] (Attitude). • Negotiate using condoms (Skills). • Use condoms correctly (Skills). • Confidence in using condoms (Self-efficacy). • Confidence in choosing not to have sex without condom use (Self-efficacy).
Users will use practical skills to protect personal information and privacy.	• Understand what should be categorized as sensitive personal information (Awareness). • Acknowledge personal privacy risks in dating app use (Risk perception).
Users will be more aware of and better protect themselves against financial scams on dating apps.	• Identify and avoid potential financial scams (Skills).

[a]STI: sexually transmitted infection.

Crowdsourcing Contest

The next task was to select evidence-based methods to translate them into practical strategies effective for the target population. We combined crowdsourcing with the peer-led approach to tap into the collective knowledge of dating app users by hosting a crowdsourcing contest. The "'Hi, Stranger!' Dating Apps Education Design Contest" was held between November 2017 and February 2018, and dating app users were encouraged to submit online multimedia content (eg, texts, images, and videos) that addressed the risks and coping strategies used in dating app use. The contest was promoted through three channels: social network outlets such as Facebook and Instagram, local nongovernmental organizations focusing on sex education, and a large general education course on sexuality and culture at a local university, where we invited students to submit a website about dating app use in Hong Kong as part of their final projects. Our recruitment of participants ensured that while they were from diverse ethnic, cultural, educational, and class backgrounds, they were of similar age to the target population and had experience with dating apps.

We received a total of 24 group submissions in the form of websites from 127 participants. The websites covered a wide range of topics, from relationship-building to sexual harassment and dating violence. Apart from reviewing media reports and online resources, participants also drew on surveys and in-depth interviews that they conducted with their peers. More data about the crowdsourcing process can be found in Textboxes 1 and 2.

Textbox 1. Demographics of the crowd for the crowdsourcing contests.

Size of crowd

- 127 participants (24 group submissions)

Age

- 17-22 years old

Gender

- Unspecified

Education

- College-educated

Racial/ethnic background

- Mixed, but mostly Asian.

Relationship to the research question

- Dating app users themselves.

Referral source

- Social media and classroom promotion of the crowdsourcing contest

Textbox 2. Logistics of the crowdsourcing.

Length of time

- Four months

Complexity of the task

- Participants were asked to create multimedia and Web-based materials related to dating app use.

Web platform

- We recommended websites be created with the free Web-based platform wix.com.
- Submissions were made electronically in the form of website URLs.

Incentives offered

- A prize of HK $2000 (US $255) was offered for the best submission

Data collected

- A total of 24 submissions were received in the form of websites.

Peer-Led Intervention Development

Upon collecting data from the crowdsourcing contest, we held a peer-led workshop to select practical strategies and produce intervention materials. We cooperated with StickyRiceLove to recruit peer educators from their pool of volunteers. A total of 12 peer educators were recruited who were between the ages of 20-30 years old and who had extensive experience in working on sex education projects. A one-day intensive workshop was held to select practical strategies and prepare for the production of intervention materials.

We briefed the peer educators about perceived benefits and risks associated with dating app use in Hong Kong, the purpose of the intervention, and the performance objectives. The peer educators were then divided into three groups to look into the websites from the crowdsourcing contest. They were given sufficient time to examine and discuss the different elements on the website both within and between groups. Through this process, the peer educators were able to identify the most effective and practical strategies based on the outcome goals, which included a risk assessment tool, a Web-based first-person simulation game that communicates benefits and risks of dating app use and coping strategies, and four short videos promoting risk perception, awareness of and skills for personal safety protection, and self-efficacy.

After the components of the intervention were determined, we adopted the Peer-Vetted Creative Production (PVCP) approach in producing intervention materials. Theorized in Brabham's

four-type typology for crowdsourcing, the PVCP approach is recommended for combining crowdsourcing and peer-led approaches to create and select creative ideas. It is ideal for ideation problems where solutions are highly sensitive to the taste of the target group [11]. We allocated time for peer educators to conduct brainstorming sessions and group discussions, through which they produced the format and scenarios for the game and themes and frameworks for the videos. These frameworks were taken up by the research team, who collaborated with sexual health professionals, volunteer actors and actresses, and photographers to produce the intervention materials.

The final intervention programs were hosted online in the form of a website and are easily accessible by computer, smartphone, and tablet. Screenshots of the intervention programs can be found in Figures 1 and 2.

Figure 1. Screenshots of the simulation game.

Figure 2. Screenshots of the short videos.

Results

The Intervention

The final intervention was comprised of four (2-4 minutes) short videos with points of discussion, an interactive scenario game (20-30 minutes), and a risk assessment tool. The videos aim to address the risks and benefits of using dating apps and encourage viewers to reflect on their perception of the dating apps used in the points of discussion. The first video illustrates similarities between meeting people on dating apps and in real life, providing examples of being misled by profile characteristics and monetary scams, the second video explains privacy concerns, the third demonstrates risky sexual behaviors associated with dating apps, and the fourth explains the legal issues and risks of sexual assault, including precautionary steps and available resources.

The scenario game is a first-person simulation game where the players are presented with multiple choices when faced with real-life scenarios. Players have to create a profile with attributes such as sexual orientation, gender, and types of relationship. Their choices generate a dating app profile, and chat bubbles will appear that contain messages from a virtual partner. The game is designed with various algorithms that result in both positive and adverse outcomes. The final page displays a brief summary explaining the culmination of the choices made.

The risk assessment tool consists of 14 questions that gives the participant a score to determine their risk level for adverse events when using dating apps. It covers five domains: sexual harassment, privacy, monetary and legal issues, and mental wellbeing.

Pilot Study and Evaluation

Before full implementation, the interventional materials and processes were tested in a pilot study. Recruitment was done through university internal mass mail services for students, and criteria included prior experience with dating app use. We recruited 18 local college students aged between 18-25 years old, and then divided the participants into two small groups. We showed them the videos and invited them to play the game. They were then asked to fill out a postintervention questionnaire for general comments about the programs, which took about an hour altogether. We also held semistructured focus group discussions among the participants to collect further qualitative feedback about the intervention programs.

The feedback collected from the pilot was generally positive. Participants found the content of the intervention programs to be relatable and realistic for dating app users in Hong Kong, and also informative about the potential risks of dating app usage. Participants also offered useful suggestions about places for improvement, including enlarging the subtitles in the videos for a clearer read and modifying details in the app to make it more realistic. The intervention programs were revised based on these comments.

The effectiveness of the intervention will be further evaluated through a randomized control trial. The intervention will be disseminated among college students aged 18-25 years old in Hong Kong through engagements with public health officials, higher education institutions, and social media companies. The intervention delivery will be a combination of classroom and social media delivery.

Discussion

Principal Findings

Crowdsourcing was implemented as a core methodology and integrated into different parts of the development of the intervention. In the planning stage, crowdsourcing proved to be a useful method to get community input for public health issues. This is particularly important for emerging issues that have not been thoroughly researched in the existing literature. In the preparation of intervention materials, combining crowdsourcing with the peer-led approach also contributed significantly to designing programs that were more relatable

and thus were more effective on the target population. Furthermore, crowdsourcing can also be a cost-effective tool since volunteers can do a considerable portion of the work.

Two key issues in adopting the crowdsourcing and peer-led approach are motivation and the extent of participation [12,13]. Brabham finds that a collection of motivators exists for crowdsourcing, including a love of community, the opportunity to develop one's skills, and building a portfolio for future employment [13]. Apart from these factors, we observed that interest in the topic and relatability to personal experiences also played important roles in motivating the crowd to contribute to intervention development. We found that crowdsourcing is effective in empowering community members to take a more active part in developing the intervention, thus avoiding situations of nonparticipation like tokenism, decoration, and manipulation. The level of participation varies in different stages of the intervention development according to the aims of that particular stage. For stages involving brainstorming, assessment of needs and tastes of the target population, and design of creative content, we encouraged a high level of participation so the projects were peer-initiated, directed, and decided. For other stages that required professional knowledge, we also drew heavily on peer-generated data and ideas. This method has helped us to use crowdsourcing effectively to develop a peer-led intervention.

Potential Challenges

During the intervention development process, the research team identified several challenges in applying crowdsourcing and the peer-led approach. One major challenge of the crowdsourcing approach is data management and analysis. Crowdsourcing could generate a considerable amount of loosely structured data, especially in the earlier stage of intervention development. This requires the research team to adopt effective strategies in data collection and interpretation. When eliciting ideas from the crowd, the research team should communicate the nature and objectives of the study, as well as the requirements for the content, to facilitate more relevant responses. To better interpret data, the research team also needs to develop a systematic analytical strategy that combines relevant theories with the existing literature, especially since the ultimate drafting of intervention outcomes and objectives will be the responsibility of the research team. One way to meet the challenge of data management is to combine large-scale, extensive crowdsourcing with small focus group discussions. This method can help ensure both the scope and depth of the data collected. The two sets of data can also complement each other and be compared for better accuracy.

Furthermore, while the PVCP approach was efficient in intensive face-to-face workshops, we found it challenging to sustain peer mentors' interest and participation in the production over time. The production of creative content usually involves various stages and therefore lasts for weeks or even months. Because our peer mentors were all volunteers who had full-time occupations, it was impractical to ask for their full participation throughout the production process. As a result, parts of the production, especially the scriptwriting, were undertaken by the research team. Better engagement of peer mentors is needed in this step. Methods may include providing incentives or scheduling several workshops in key stages of the production.

Finally, and more generally, developing an intervention for dating app use in Hong Kong poses challenges in balancing between providing more information and perpetuating social stigma around dating app users. Due to conservative attitudes toward sex in the local context, dating app users have been stigmatized as casual, promiscuous, and immoral. Consequently, there is a fine line between raising awareness about risks in dating app use and further stigmatizing dating app users as immoral and irresponsible. To avoid the latter outcome, we adopted various methods, such as engaging dating app users in setting the intervention goals and addressing both benefits and risks in the intervention programs. However, it remains unclear whether these methods will be sufficient. Intervention developers should always consider cultural specificities and make their best efforts to eliminate any possible negative impacts brought by the intervention.

Limitations

Several limitations should be considered in this study. First, the recruiting and promotion processes during the needs assessment, crowdsourcing contest, and peer-led intervention development were done through the research team's network of nongovernmental organizations and a large general education course at a local university. While this allows the research team to target select groups of people who may have an interest in this topic, it may overlook other individuals who belong to other organizations or universities. Second, since the intervention development is heavily based on local contexts in Hong Kong, its results may not be generalizable in other places.

Conclusion

This paper describes the evidence-based development of an intervention for safer dating app use using crowdsourcing and the peer-led approach. Our findings suggest that the advantages of crowdsourcing lay in its cost-effectiveness, its ability to collect extensive data, and its engagement of community members. We recommend that crowdsourcing be strategically used in various stages of intervention development.

This paper has also identified potential challenges in adopting the crowdsourcing approach, including data management and analysis and peer mentors' participation in intervention material production. We suggest that intervention developers should pay special attention to the communication of program objectives during crowdsourcing and should work out a well-structured schedule to better sustain peer mentors' interests throughout material production. Moreover, we have highlighted local cultural specificities as an important issue that should be taken into consideration during intervention development. The crowdsourced, peer-led model of intervention development harnesses both new technologies and collective intelligence. Therefore, we believe that it is particularly suitable to tackling new public health issues that have not been thoroughly researched in the existing literature.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview schedule (bilingual).
[DOCX File , 18 KB - formative_v4i4e12098_app1.docx]

Multimedia Appendix 2
Themes from focus group discussions.
[DOCX File , 18 KB - formative_v4i4e12098_app2.docx]

References

1. Sawyer AN, Smith ER, Benotsch EG. Dating Application Use and Sexual Risk Behavior Among Young Adults. Sex Res Soc Policy 2017 Sep 6;15(2):183-191. [doi: 10.1007/s13178-017-0297-6]
2. Goedel WC, Duncan DT. Geosocial-Networking App Usage Patterns of Gay, Bisexual, and Other Men Who Have Sex With Men: Survey Among Users of Grindr, A Mobile Dating App. JMIR Public Health Surveill 2015 May 08;1(1):e4 [FREE Full text] [doi: 10.2196/publichealth.4353] [Medline: 27227127]
3. Choi EP, Wong JY, Lo HH, Wong W, Chio JH, Fong DY. The association between smartphone dating applications and college students' casual sex encounters and condom use. Sex Reprod Healthc 2016 Oct;9:38-41. [doi: 10.1016/j.srhc.2016.07.001] [Medline: 27634663]
4. Hobbs M, Owen S, Gerber L. Liquid love? Dating apps, sex, relationships and the digital transformation of intimacy. Journal of Sociology 2016 Sep 05;53(2):271-284. [doi: 10.1177/1440783316662718]
5. Choi EPH, Wong JYH, Fong DYT. The use of social networking applications of smartphone and associated sexual risks in lesbian, gay, bisexual, and transgender populations: a systematic review. AIDS Care 2017 Feb 25;29(2):145-155. [doi: 10.1080/09540121.2016.1211606] [Medline: 27454158]
6. Hoang N, Asano Y, Yoshikawa M. Your Neighbors Are My Spies: Location and Other Privacy Concerns in Dating Apps. 2017 Presented at: 2017 19th International Conference on Advanced Communication Technology (ICACT); 19-22 February; Bongpyeong, South Korea. [doi: 10.23919/icact.2017.7890236]
7. Farnden J, Martini B, Choo KKR. arXiv.org.: Cornell University; 2015. Privacy Risks in Mobile Dating Apps URL: https://arxiv.org/abs/1505.02906 [accessed 2020-01-21]
8. Huang ET, Williams H, Hocking JS, Lim MS. Safe Sex Messages Within Dating and Entertainment Smartphone Apps: A Review. JMIR Mhealth Uhealth 2016 Nov 08;4(4):e124 [FREE Full text] [doi: 10.2196/mhealth.5760] [Medline: 27826133]
9. Gilmore G, Campbell M. Needs assessment strategies for health education and health promotion. 2nd ed. Madison, Wisconsin, United States: Brown and Benchmark Publishers; 1996.
10. Powell RA, Single HM. Focus groups. Int J Qual Health Care 1996 Oct;8(5):499-504. [doi: 10.1093/intqhc/8.5.499] [Medline: 9117204]
11. Brabham DC, Ribisl KM, Kirchner TR, Bernhardt JM. Crowdsourcing applications for public health. Am J Prev Med 2014 Feb;46(2):179-187. [doi: 10.1016/j.amepre.2013.10.016] [Medline: 24439353]
12. Sun WH, Miu HYH, Wong CKH, Tucker JD, Wong WCW. Assessing Participation and Effectiveness of the Peer-Led Approach in Youth Sexual Health Education: Systematic Review and Meta-Analysis in More Developed Countries. J Sex Res 2018 Jan 29;55(1):31-44. [doi: 10.1080/00224499.2016.1247779] [Medline: 27898248]
13. Brabham DC. Moving the Crowd at Threadless: Motivations for participation in a crowdsourcing application. Information, Communication & Society 2010 Dec;13(8):1122-1145. [doi: 10.1080/13691181003624090]

Abbreviations

PVCP: Peer-Vetted Creative Production

Edited by G Eysenbach; submitted 03.09.18; peer-reviewed by YA Hong, B Ranard, D Swendeman; comments to author 01.04.19; revised version received 15.08.19; accepted 16.12.19; published 21.04.20.

Please cite as:
Wong WCW, Song L, See C, Lau STH, Sun WH, Choi KWY, Tucker J
Using Crowdsourcing to Develop a Peer-Led Intervention for Safer Dating App Use: Pilot Study
JMIR Form Res 2020;4(4):e12098
URL: http://formative.jmir.org/2020/4/e12098/
doi:10.2196/12098
PMID:32314975

Original Paper

An Electronic Medication Module to Improve Health Literacy in Patients With Type 2 Diabetes Mellitus: Pilot Randomized Controlled Trial

Hanna Marita Seidling[1,2*], Dr sc hum; Cornelia Mahler[3,4*], Dr sc hum; Beate Strauß[3], Pharmacist; Aline Weis[3], MSc; Marion Stützle[1,2], Dr rer nat; Johannes Krisam[5], Dr sc hum; INFOPAT P4/P5 Study Team[6*]; Joachim Szecsenyi[3], MD; Walter Emil Haefeli[1], MD

[1]Department of Clinical Pharmacology and Pharmacoepidemiology, Heidelberg University Hospital, Heidelberg, Germany

[2]Cooperation Unit Clinical Pharmacy, Heidelberg University Hospital, Heidelberg, Germany

[3]Department of General Practice and Health Services Research, Heidelberg University Hospital, Heidelberg, Germany

[4]Department of Nursing Science, University Hospital Tuebingen, Tuebingen, Germany

[5]Institute of Medical Biometry and Informatics, Heidelberg University Hospital, Heidelberg, Germany

[6]See Acknowledgments

[*]these authors contributed equally

Corresponding Author:
Walter Emil Haefeli, MD
Department of Clinical Pharmacology and Pharmacoepidemiology
Heidelberg University Hospital
Im Neuenheimer Feld 410
Heidelberg, 69120
Germany
Phone: 49 6221 56 8740
Email: walter.emil.haefeli@med.uni-heidelberg.de

Abstract

Background: In primary care, patients play a crucial role in managing care processes and handling drug treatment. A decisive factor for success is their health literacy, and several interventions have been introduced to support patients in fulfilling their responsibility.

Objective: The aim of this study is to assess the influence of such an intervention (ie, a medication module) within a patient-led electronic health record on patients' health literacy.

Methods: We conducted a randomized controlled study among community-dwelling patients with type 2 diabetes mellitus. Patients were recruited from primary care practices. After randomization, patients either had access to an internet-based medication module allowing them to store their medication information, look up drug information, and print a medication schedule (intervention group), or they received an information brochure on the importance of medication schedules (control group). After 4-8 weeks, all patients were invited to attend a structured medication review (ie, follow-up visit). Data were collected via questionnaires before the start of the intervention and during the follow-up visit. The main outcome measure was the mean difference in health literacy between baseline and follow-up assessments of patients in the control and intervention groups.

Results: Of 116 recruited patients, 107 (92.2%) completed the follow-up assessment and were eligible for intention-to-treat analyses. Only 73 patients, of which 29 were in the intervention group, followed the study protocol and were eligible for per-protocol analysis. No differences in overall health literacy were observed in either the intention-to-treat or in the per-protocol cohorts. Reasons for a null effect might be that the cohort was not particularly enriched with participants with low health literacy, thus precluding measurable improvement (ie, ceiling effect). Moreover, the success of implementation was considered poor because both the correct application of the study procedure (ie, randomization according to the protocol and dropout of 29 patients) and the actual interaction with the medication module was modest (ie, dropout of 9 patients).

Conclusions: The conduct of this randomized controlled study was challenging, leaving it open whether inadequate implementation, too short of a duration, or insufficient efficacy of the intervention, as such, contributed to the null effect of this study. This clearly outlines the value of piloting complex interventions and the accompanying process evaluations.

(JMIR Form Res 2020;4(4):e13746) doi:10.2196/13746

KEYWORDS

medication self-management; patient empowerment; health literacy; chronic diseases; type 2 diabetes mellitus; electronic health record; PEPA; electronic medication module; structured medication review

Introduction

In the course of their disease, patients with type 2 diabetes mellitus (T2DM) depend on the lifelong intake of drugs, which requires continuous unremitting efforts to self-manage the medication process. Thereby, medication self-management refers not only to active drug administration but also to filling and picking up prescriptions, understanding the medication regimen, integrating it in a daily schedule, monitoring drug effects, and, finally, sustaining it over the long term [1].

Besides the complexity of the disease and the medication regimen, medication self-management capacity is also influenced by patient characteristics, such as patients' health literacy [1,2]. According to Sørensen and coworkers, health literacy "... entails people's knowledge, motivation, and competencies to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention, and health promotion to maintain or improve quality of life during the life course" [3]. All four dimensions of health literacy are relevant within medication self-management: (1) access and obtain information relevant to health, (2) understand information relevant to health, (3) process and appraise information relevant to health, and (4) apply and use information relevant to health. Health literacy, therefore, plays an important role in medication self-management because limited health literacy can be a barrier to medication reconciliation [3] and can be associated with medication beliefs [4,5] and medication nonadherence [4,6].

Improving health literacy has been targeted by a number of single and mixed interventions aimed at improving comprehension [7] or provision of programs supporting patients to actively fulfill their roles in medication self-management as do patient portals, for instance [8]. Patient portals are often interlinked with electronic health records (EHRs), which, until today, have the predominant aim to store all disease-related documentation, facilitating interchange and transfer of information between health care professionals. Current evidence on patients' attitudes toward patient portals suggest that patients might not use these portals on a regular basis [9], for instance, because they are not familiar with the functionalities they offer [10]. Even if the functionalities are known, the handling might still be complicated for the target population or they might be reluctant to use them, as has been shown for additional functionalities, such as secure messaging. Hence, these functionalities should always be assessed for potential barriers in usage [11]. However, if patients are individually informed and trained according to their needs on how to use the portal, and a particular effort is put on system usability from the beginning, patient portals might be useful tools to strengthen the patient's role [10,12,13].

In 2012, a patient-controlled PEPA (personal EHR) was launched in the region of Rhine-Neckar in the southwest of Germany, within the INFOPAT (INFOrmation Technology for PATient-oriented Healthcare in the Rhine-Neckar metropolitan region) project [14]. The PEPA is "owned" by the patient so that he or she gains control over all disease-associated and medication-related data shared in the PEPA [15,16]. The PEPA offers the function of documentation but also actively involves patients in his or her process of care. In contrast to previous EHRs, patients can customize their PEPA themselves and give health care professionals access to its content [17]. To further support patients, the PEPA was equipped with an interactive medication module where patients can access information on different drugs, check drug-drug interactions, get evidence-based information on diabetes and its drug treatment, and compose and update their personal medication schedule in order to support medication self-management.

In total, the PEPA with the medication module forms a complex intervention that might support patients in fulfilling their active roles in the medication process and during drug handling. To assess the medication module's effects, a complex and well-piloted study design will be required. Thereby, it will be particularly challenging to analyze causal relationships for success or reasons for failure of such an intervention, which may both result from inadequate implementation as well as inadequate effects.

The aim of this study was to pilot the feasibility of the study procedure (ie, patient recruitment, randomization, data assessment, and documentation in the primary care practices) and of the intervention (ie, conduction of training episodes and accessibility of the medication module), while at the same time assessing the influence of the medication module on health literacy.

Methods

Overview

We performed an unblinded, exploratory, prospective, randomized controlled study with two data collection points among community-dwelling patients attending primary care practices. The study protocol was approved by the Ethics Committees of the Medical Faculty of Heidelberg University (S-540/2015) and of the State Medical Board of the state of Baden-Württemberg, Germany.

Recruitment

Primary care practices were recruited via a practice network of the Department of General Practice and Health Services Research at Heidelberg University Hospital, Germany. In each practice, the responsible general practitioner (GP) agreed to participate and entrusted a medical assistant with the coordination of the study in his or her practice and the conduct

of a structured medication review with all included patients. Because the medical assistants were asked to evaluate each structured medication review, they were also included as study participants and signed informed consent forms.

Patients were eligible if they were 18 years of age or older, were diagnosed with T2DM and treated with oral antidiabetic drugs and/or insulin, were considered mentally and physically capable of participating in the study, were generally familiar with computer use, had access to a computer or mobile device with internet access and an email account, and consented in written form to participate in the study. Patients not fulfilling the inclusion criteria could not participate in the study; apart from that, there were no formal exclusion criteria.

Each participating primary care practice received €100 for study-related expenditures, all participating medical assistants received compensation of €200 plus €40 per patient, and patients received compensation of €50 for their efforts.

Intervention

The intervention comprised the personal use of the medication module that was embedded in the PEPA. The medication module was developed based on previous qualitative focus group discussions with T2DM patients, GPs, and pharmacists, who defined user requirements [18]. The user requirements were transferred to an initial prototype of the medication module that was then pretested with 6 patients and subsequently optimized considering the patients' feedback. Finally, the module consisted of a documentation submodule for the patients' medication regimen and a general information portal on drug administration. The documentation submodule facilitated searching all drugs available on the German drug market, putting together a medication regimen, and entering additional information, such as drug dosage, indication, and administration advice. Subsequently, the patients could store the medication regimen to access and modify in the future. Moreover, they could print out a medication schedule containing all previously entered information. For each drug, the patient had access to basic drug information, such as the package information leaflet and more detailed information on drug administration (eg, on whether a tablet could be split or crushed). Moreover, the entire medication regimen was checked for drug-drug interactions involving over-the-counter drugs as well as contraindicated combinations.

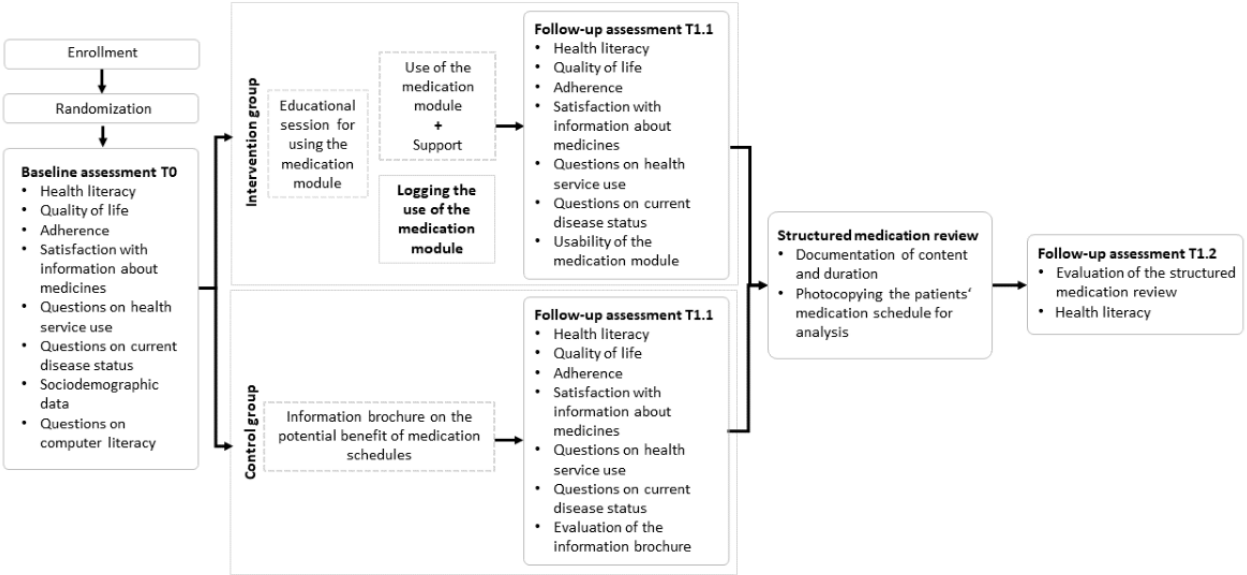
Before using the medication module for the first time, each patient completed a 45-minute, face-to-face, hands-on structured training session by a member of the study team and printed education material on how to use the module. Moreover, patients could approach a first-level support person in case they had problems using the medication module of the PEPA. However, participants were not contacted proactively to ensure that they used the tool.

Participants of the control group received a brochure about the importance and content of a medication schedule.

Study Procedure and Outcome Measures

The study (see Figure 1) was conducted between May and October 2016. Each practice invited patients who met the inclusion criteria to participate in the study. Patients who agreed to participate were randomized by the medical assistant to the intervention group and the control group at a 1:1 ratio, following a previously created randomization list.

Figure 1. Study procedure of the prospective randomized controlled study. Data assessments at T0, T1.1, and T1.2 were done via questionnaires filled in by the patient during his or her visits to the primary care practice.



The study consisted of two patient visits in the primary care practice before (ie, baseline visit, T0) and after the intervention (ie, follow-up visit, T1). During these visits, there were three data-assessment points: one at T0 and two at T1, with T1.1 before the structured medication review and T1.2 after the medication review. During these time points, both the baseline information, such as sociodemographic information, and the patient- and process-oriented outcomes were documented using paper-based questionnaires, typically in the GP practice. All patient-oriented outcomes were measured using German versions of validated survey instruments.

During the baseline assessment, the patient filled in a questionnaire documenting health literacy using the Health Literacy Questionnaire (HLQ), which was developed by Osborne and colleagues [19,20]. Out of the eight subscales of the questionnaire, we selected the four subscales that were of particular interest for health literacy with regard to drug treatment: scales 5 (*appraisal of health information*), 6 (*ability to actively engage with health care providers*), 8 (*ability to find good health information*), and 9 (*understanding health information well enough to know what to do*). Moreover, patients filled in questionnaires regarding quality of life (ie, the World Health Organization Quality of Life instrument, brief version [WHOQOL-BREF]-global items [21]), self-reported adherence (ie, the Medication Adherence Report Scale, German version [MARS-D] [22]), satisfaction with drug information (ie, the Satisfaction with Information about Medicines Scale, German version [SIMS-D] [23,24]), and utilization of medical services (ie, the Mannheimer Modul Ressourcenverbrauch [MMRV] [25]). Patients also provided specified information regarding their sociodemographic and disease status (ie, blood pressure, hypoglycemia, hemoglobin A1c value, fasting blood glucose levels, and drug treatment) as well as current internet and computer use.

Directly after the baseline assessment, the patients in the control group received an information brochure on the potential benefit of medication schedules. In the intervention group, patients were scheduled for a training date to be educated about the use of the electronic medication module. The training session was performed by four study team members to ensure high reliability and consistent quality for all training sessions.

A total of 4-8 weeks after randomization, the medical assistant invited the patients to the primary care practice for a follow-up visit (T1). During this visit, patients received the same questionnaire that was administered during the baseline assessment, except that it contained only scales 5 and 8 from the HLQ and no sociodemographic data or questions about computer literacy. In the intervention group, the questionnaire additionally contained questions regarding the usability of the medication module (ie, the System Usability Scale [SUS] [26]), while in the control group, patients were asked about the comprehensibility of the information brochure. Subsequently, the medical assistant performed a structured medication review with each patient to document which drugs the patient was actually taking. During the medication review, potential problems with the actual medication were assessed and documented to be discussed later on with the GP [27]. If the patient possessed a printed medication schedule, the medical assistant would photocopy it but was instructed not to use this schedule during the medication review. The duration of the structured medication review was documented. After the structured medication review, both patient and medical assistant filled in a questionnaire evaluating the conduct of the medication review. The patient questionnaire additionally contained scales 6 and 9 from the HLQ. Throughout the intervention period, the use of the medication module (ie, number of log-ins and performed actions) was logged.

Primary and Secondary Endpoints

As primary endpoint, the change in health literacy between T0 and T1 was assessed. As secondary endpoints, the following were evaluated: differences in quality of life, adherence, and satisfaction with drug information; differences in hemoglobin A1c, fasting blood glucose levels, and hypoglycemia; the course, time consumption, satisfaction, and evaluation of the structured medication review; the prevalence and quality of the patients' medication schedules; and the use of, and satisfaction with, the medication module and the information brochure. This paper will focus on the general feasibility of the study design and will report the primary endpoint. Secondary endpoints will be aggregated and reported elsewhere.

Statistical Analysis

We planned to enroll 120 patients (n=60 per group). This sample size was primarily based on matters of feasibility and allowed us to detect a standardized treatment effect of Cohen *d*=0.6 with a power of 1–beta=.9 when applying a two-sided *t* test with a two-sided significance level of alpha=.05.

We conducted two different types of analyses on two populations to assess primary and secondary outcomes. The intention-to-treat population comprised all randomized patients, while the per-protocol population comprised all randomized patients without protocol deviations. The primary endpoint was the difference between T1 and T0 in the sum score of scales 5, 6, 8, and 9 of the HLQ [19,20]. Each of the four scales is composed of five items with either four (scale 5) or five (scales 6, 8, and 9) Likert-scale values. Scale scores were determined by calculating the mean of all answered items, if at least three items per scale were answered. In case all four scale scores could be determined, the HLQ sum score was defined as the mean of all four sum scores and was otherwise set to missing.

Due to the hierarchical data structure, a multilevel analysis was conducted, with patients at level 1 and practice at level 2. The primary model was a linear mixed model with the HLQ score difference (T1–T0) as the dependent variable; treatment group and baseline HLQ score as fixed factor and covariate, respectively; and practice as a random factor, using the restricted maximum-likelihood method to fit the model. The primary analysis was conducted using the multilevel approach, assuming that missing values in the primary outcome can either be explained by the baseline HLQ score or the treatment group. Thus, no additional imputation of missing values was conducted.

Secondary endpoints such as scale scores were analyzed using the same multilevel approach. Effect estimates were calculated alongside 95% CIs and *P* values. Due to the exploratory character of the study, all *P* values were only of a descriptive nature, thus, no adjustment for multiple testing was performed. *P* values less than .05 were regarded as statistically significant. All analyses were conducted using SAS 9.4 (SAS Institute).

Results

Participants

Overall, 13 GPs agreed to participate and recruit patients, with one GP entrusting three medical assistants to coordinate the

study and conduct the structured medication reviews. Hence, 15 medical assistants in 13 primary care practices were involved in randomization and data assessment.

Overall, 116 patients agreed to participate in the study and were allocated to either the intervention or control group (see Figure 2). Of those, 113 (97.4%) participated in the baseline assessment and 107 (92.2%) also participated in the follow-up assessment. These participants were included in the intention-to-treat analysis. In total, 75.0% of included participants (87/116) were randomized according to the protocol. Of those, 73 participants (84%) followed the intervention as intended, completed the follow-up assessment, and were consequently eligible for per-protocol analysis. Table 1 presents the sociodemographic information as well as information about diagnosis and medication of all included participants.

Figure 2. Patient flow during the study procedure. All patients from the intervention group are displayed on the right side of the flowchart and all patients from the control group are displayed on the left side. As not all patients were correctly enrolled in the study, there are three branches displayed on each side. Dropouts according to loss of follow-up are displayed vertically.

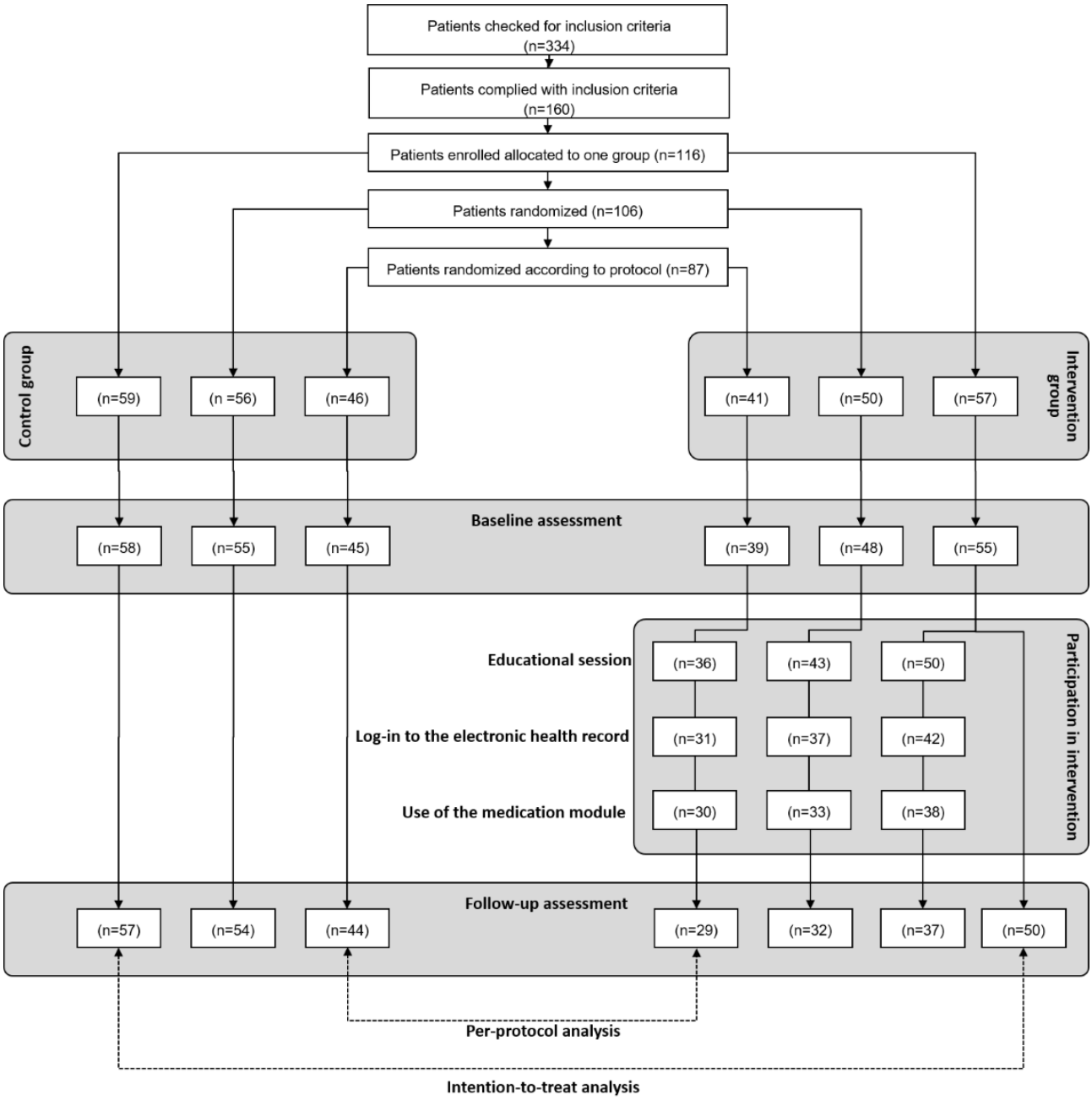


Table 1. Characteristics of participants.

Characteristic	Intention-to-treat analysis			Per-protocol analysis		
	Intervention group (n=55)	Control group (n=58)	Total (n=113)	Intervention group (n=29)	Control group (n=44)	Total (n=73)
Age (years), mean (SD)	57.5 (11.2)	60.1 (10.2)	58.9 (10.8)	55.3 (10.5)	60.3 (10.8)	58.3 (10.9)
Gender, n (%)						
Male	30 (56)	37 (64)	67 (59.8)	21 (75)	27 (61)	48 (67)
Missing values	1 (2)	0 (0)	1 (0.9)	1 (4)	0 (0)	1 (1)
Current employment, n (%)						
Employed	27 (50)	33 (57)	60 (53.6)	17 (61)	26 (59)	43 (60)
Missing values	1 (2)	0 (0)	1 (0.9)	1 (4)	0 (0)	1 (1)
Living alone, n (%)	12 (22)	13 (22)	25 (22.1)	8 (28)	10 (23)	18 (25)
Comorbidities (yes), n (%)	50 (91)	48 (83)	98 (86.7)	27 (93)	37 (84)	64 (88)
Diabetes diagnosis						
Time since diagnosis (years), mean (SD)	9.6 (8.8)	9.7 (7.4)	9.7 (8.1)	9.8 (7.9)	8.9 (7.4)	9.3 (7.6)
Missing values, n (%)	6 (11)	11 (19)	17 (15.0)	4 (14)	10 (23)	14 (19)
Drugs						
Number of drugs per day, mean (SD)	5.3 (3.3)	4.6 (2.7)	4.9 (3.0)	5.0 (3.1)	4.5 (3.0)	4.7 (3.0)
Missing values, n (%)	3 (5)	3 (5)	6 (5.3)	1 (3)	3 (7)	4 (5)
Diabetes medication, n (%)						
Tablets	34 (64)	40 (71)	74 (67.9)	17 (61)	29 (69)	46 (66)
Tablets and injection	11 (21)	12 (21)	23 (21.1)	6 (21)	9 (21)	15 (21)
Injection	8 (15)	3 (5)	11 (10.1)	5 (18)	3 (7)	8 (11)
Other	0 (0)	1 (2)	1 (0.9)	0 (0)	1 (2)	1 (1)
Missing values	2 (4)	2 (4)	4 (3.7)	1 (4)	2 (5)	3 (4)

Health Literacy

Having access to the electronic medication platform did not have any significant effect on health literacy when comparing the two treatment groups, but it appeared to influence some aspects of health literacy when only assessing the effect in the intervention group (see Table 2). Results of the per-protocol analysis tended to be similar to those of the intention-to-treat analysis, albeit intervention group effects were slightly higher (see Table 3).

Table 2. Differences between Health Literacy Questionnaire (HLQ) sum scores and scores from scales 5, 6, 8, and 9 by group at baseline (T0) and follow-up visits (T1), including effect estimates by group adjusted for baseline value and two-level structure .

Measures	Intention-to-treat analysis, T1–T0 scores			Per-protocol analysis, T1–T0 scores		
	Intervention group	Control group	Total	Intervention group	Control group	Total
HLQ sum score (primary endpoint)						
Complete observations, n	50	55	105	29	43	72
Mean (SD)	0.18 (0.34)	0.15 (0.39)	0.17 (0.37)	0.24 (0.36)	0.12 (0.34)	0.17 (0.35)
Median	0.16	0.15	0.15	0.20	0.14	0.15
Minimum, maximum	–0.35, 1.25	–1.35, 1.25	–1.35, 1.25	–0.31, 1.25	–1.35, 0.80	–1.35, 1.25
Effect estimate	0.18	0.15		0.22	0.11	
SE	0.05	0.05		0.07	0.06	
P value	.001	.005		.002	.07	
95% CI	0.07 to 0.28	0.05 to 0.25		0.09 to 0.36	–0.01 to 0.23	
Scale 5: Appraisal of health information (four scale values)						
Complete observations, n	50	56	106	29	44	73
Mean (SD)	0.11 (0.44)	0.16 (0.55)	0.13 (0.50)	0.09 (0.47)	0.08 (0.53)	0.09 (0.50)
Median	0.00	0.20	0.00	0.00	0.00	0.00
Minimum, maximum	–1.00, 1.60	–1.20, 1.80	–1.20, 1.80	–1.00, 1.60	–1.20, 1.80	–1.20, 1.80
Effect estimate	0.16	0.11		0.12	0.06	
SE	0.06	0.06		0.08	0.06	
P value	.01	.05		.13	.32	
95% CI	0.04 to 0.28	–0.00 to 0.23		–0.04 to 0.28	–0.06 to 0.19	
Scale 6: Ability to actively engage with health care providers (five scale values)						
Complete observations, n	50	57	107	29	44	73
Mean (SD)	0.86 (0.93)	0.93 (0.98)	0.90 (0.95)	0.99 (0.76)	0.88 (0.96)	0.92 (0.88)
Median	0.88	0.80	0.80	1.00	0.70	0.80
Minimum, maximum	–2.40, 3.00	–0.80, 3.20	–2.40, 3.20	–0.60, 3.00	–0.80, 3.20	–0.80, 3.20
Effect estimate	0.81	0.95		0.95	0.86	
SE	0.11	0.11		0.13	0.11	
P value	<.001	<.001		<.001	<.001	
95% CI	0.58 to 1.03	0.73 to 1.17		0.69 to 1.21	0.63 to 1.09	
Scale 8: Ability to find good health information (five scale values)						
Complete observations, n	50	56	106	29	44	73
Mean (SD)	–0.41 (0.70)	–0.36 (0.89)	–0.38 (0.81)	–0.30 (0.51)	–0.34 (0.84)	–0.33 (0.72)
Median	–0.30	–0.20	–0.20	–0.20	–0.20	–0.20
Minimum, maximum	–2.40, 1.00	–3.40, 2.20	–3.40, 2.20	–1.20, 1.00	–3.40, 1.20	–3.40, 1.20
Effect estimate	–0.40	–0.41		–0.33	–0.39	
SE	0.13	0.13		0.17	0.14	
P value	.003	.002		.05	.009	
95% CI	–0.66 to –0.14	–0.66 to –0.15		–0.66 to 0.00	–0.68 to –0.10	
Scale 9: Understand health information well enough to know what to do (five scale values)						
Complete observations, n	50	56	106	29	43	72
Mean (SD)	0.18 (0.62)	–0.06 (0.70)	0.05 (0.67)	0.18 (0.63)	–0.08 (0.68)	0.02 (0.67)
Median	0.20	–0.20	0.00	0.20	0.00	0.00

Measures	Intention-to-treat analysis, T1–T0 scores			Per-protocol analysis, T1–T0 scores		
	Intervention group	Control group	Total	Intervention group	Control group	Total
Minimum, maximum	–1.20, 2.00	–2.80, 1.60	–2.80, 2.00	–0.80, 2.00	–2.80, 1.40	–2.80, 2.00
Effect estimate	0.12	–0.02		0.13	–0.09	
SE	0.09	0.08		0.12	0.10	
P value	.15	.83		.29	.39	
95% CI	–0.05 to 0.30	–0.18 to 0.14		–0.11 to 0.37	–0.30 to 0.12	

Table 3. Intervention effects for the intervention group compared to the control group adjusted for baseline value and two-level structure: difference between baseline visit (T0) and follow-up visit (T1).

Scales	Effect[a]	SE	*P* value	95% CI	Cohen *d*[b]
Intention-to-treat analysis					
Sum score	0.031	0.06	.62	–0.09 to 0.16	0.08
Scale 5: Appraisal of health information	0.043	0.08	.60	–0.12 to 0.20	0.09
Scale 6: Ability to actively engage with health care providers	–0.141	0.11	.21	–0.36 to 0.08	–0.15
Scale 8: Ability to find good health information	0.004	0.14	.98	–0.27 to 0.28	0.01
Scale 9: Understand health information well enough to know what to do	0.142	0.10	.17	–0.06 to 0.35	0.21
Per-protocol analysis					
Sum score	0.116	0.08	.12	–0.03 to 0.27	0.33
Scale 5: Appraisal of health information	0.058	0.10	.57	–0.14 to 0.26	0.12
Scale 6: Ability to actively engage with health care providers	0.093	0.12	.44	–0.15 to 0.33	0.11
Scale 8: Ability to find good health information	0.056	0.16	.73	–0.26 to 0.38	0.07
Scale 9: Understand health information well enough to know what to do	0.219	0.13	.09	–0.04 to 0.48	0.33

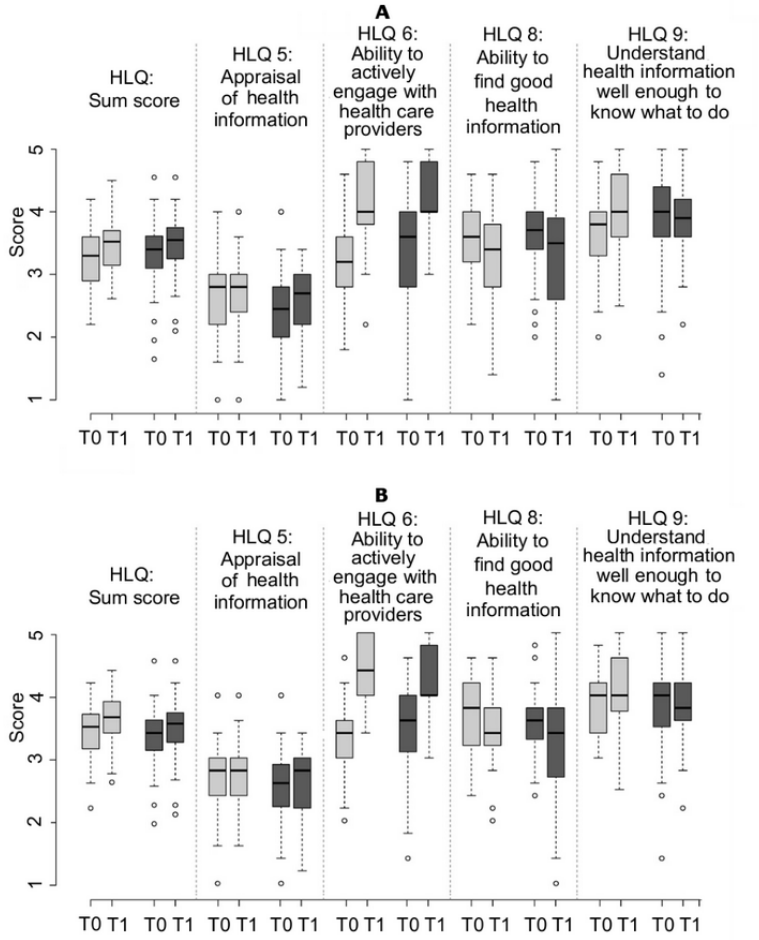
[a]Treatment effect estimate: difference between the intervention group and control group. Positive effect estimates indicate an advantage of the intervention group over the control group.

[b]Standardized effect estimate adjusted for standard deviation.

Most included participants declared to have appraised the health information (HLQ scale 5) and to have been able to actively engage with health care providers (HLQ scale 6) similarly well both before and after taking part in the study (see Figure 3). Nevertheless, both the appraisal of health information and the ability to engage with health care providers were ranked slightly higher after participating, whereas the effect on the ability to engage with health care providers was greater and statistically significant (*P*<.001; see Table 2). However, this effect occurred in both groups (see Table 3).

Intervention and control groups differed with regard to HLQ scale 9, and participants of the intervention group were better able to understand health information well enough to know what to do than were the participants of the control group. The understanding slightly rose in the intervention group and slightly decreased in the control group at the follow-up assessment (see Figure 3), but the difference between treatment groups was not statistically significant (see Table 3). The ability to find good health information (HLQ scale 8) was estimated as slightly worse after participating than before (see Figure 3). Even here, no statistically significant difference between intervention and control groups could be observed (see Table 3).

Figure 3. Results for health literacy. Graph A displays the results of the intention-to-treat analysis and Graph B displays the results of the per-protocol analysis. Patients from the intervention group are displayed in light gray and patients from the control group are displayed in dark gray. Health Literacy Questionnaire (HLQ) subscales are indicated by their numbers (5, 6, 8, and 9). T0: baseline visit; T1: follow-up visit.



Discussion

Principal Findings

In this controlled pilot study among T2DM patients, we assessed the impact of providing access to a PEPA with an interactive medication module on health literacy and found no change. This result should be critically discussed on several levels, as done in the following sections. This approach also highlights the limitations of the study.

Selection of Assessment Method of Health Literacy

We assessed health literacy with a validated tool evaluating four of its key dimensions (see Tables 2 and 3). The instrument has been successfully used in previous studies and showed good efficacy in distinguishing different health literacy levels [19,20]. Moreover, it has been applied in intervention studies in a wide range of countries [28] and showed reliable results, while, at the same time, being shorter than other tools [28], suggesting that the questionnaire might be easier to use as an outcome measure tool. Indeed, the included patients dealt well with the provided questionnaires, as shown in the small number of missing values. The fact that we decided to use only four dimensions of the tool was owed to the fact that these dimensions explicitly focused on medication-related issues (ie, the main focus of our study). Obviously, we might have received different results if we had applied all eight dimensions or used other tools assessing health literacy, such as the Health Literacy Survey, German version (HLS-Ger) [29], which consists of 47 items and has shown a correlation between low health literacy and adherence among the German population. However, the following issues kept us from using this tool in our study: the length of the questionnaire, the fact that it assesses a broad concept of health literacy, and the fact that it was primarily developed to obtain and compare epidemiological data on health literacy in various populations, rather than to be used in intervention studies [30].

Hence, we believe that both the tool as well as the assessment method used in this study were appropriate and should have detected potential differences in health literacy, if they had occurred.

Duration of the Intervention

Previous intervention studies that successfully addressed health literacy often had longer durations, typically lasting 9 months to 2 years [31-33]. Therefore, our study might have been too short to produce its full effects. This is also highlighted by the fact that not all patients in the intervention group actually accessed the PEPA.

Health Literacy Level of Included Patients and Expected Effects

At the beginning of the study, our study participants already showed comparable or higher health literacy levels compared to previous cohorts that were included in intervention studies to improve health literacy, suggesting that ceiling effects might have occurred, making the detection of small changes more difficult.

A population-based Danish study assessing health literacy in subscales 6 and 9 of the HLQ among people with long-term chronic conditions revealed lower subscale means than in our study [34]. The results on the HLQ subscales in a study among an Australian population of older patients with multiple conditions showed higher mean values on *appraisal of health information* (mean 2.78) than in our study. The *ability to actively engage with health care professionals* was lower in our study compared to the Australian population (mean 3.97); however, this value increased and was higher after the intervention (mean 4.17), demonstrating the benefit of the contact with the medical assistants. The *ability to find good health information* showed lower mean values compared to the Australian population (mean 3.65), whereas in the subscale *understand health information enough to know what to do*, our values stayed within the 95% CI reported in the Australian population (95% CI 3.81-3.91) [35]. The comparability with this sample is restricted, as patients in our sample were younger; however, the results may indicate that the medication module was a support to patients regarding information finding and action.

Moreover, during the sample size calculation, we assumed that with the given sample size, only a relatively high standardized treatment effect of Cohen *d*=0.6 could be detected with a sufficiently high power. Since we only observed small-to-moderate effects, the highest being a Cohen *d* of 0.33 for the primary endpoint and scale 9 in the per-protocol population, our study sample size was too small to yield statistically significant treatment group differences but high enough to now thoroughly plan a prospective intervention study. Another potential reason for an insignificant study outcome might be that we did not enrich our study population with patients with low health literacy.

Success of the Implementation

Strikingly high dropout rates among the intervention and control groups, which left only 73 instead of an initial 116 enrolled patients for per-protocol analysis, clearly suggest that insufficient power could be a reason for nonefficacy. However, the high dropout rates appeared in several stages of the study procedure; this is worth a detailed discussion to identify weaknesses and strengths of the study design and to derive lessons learned for future studies that can guide future complex interventions in this field.

After inclusion in the study, group allocation already appeared difficult for medical assistants, and of 116 patients, only 87 were randomized according to the protocol. Thereby, randomization according to the protocol either failed because patients were deliberately allocated to one group (n=10) or because the correct randomization scheme was not understood (n=19). Hence, incorrect randomization already diminished study power by about 10 percent-points as compared to the power that was originally aspired to of 90%, assuming a treatment effect of Cohen *d*=0.6.

Subsequently, dropouts occurred at the level of the baseline assessment, with one patient in the control group and two patients in the intervention group withdrawing consent or failing to keep their appointment for the baseline assessment. While these dropouts are typical for any type of intervention study, the dropout rate during the intervention phase, in particular, was unexpectedly large. This was either because the patients failed to attend the educational sessions where they were trained in the use of the medication platform or, even more often, because they never used the medication platform during the intervention phase. Indeed, after training, eight patients of the intervention group never even logged in to the PEPA and four more logged in at least once but never used the medication platform. Hence, almost one in three patients did not take part in the intervention, even though they agreed to take part in the study, filled in the baseline assessment, and participated in a training course. This high dropout rate during that stage was unexpected and certainly not considered during the sample size calculation.

Reasons for nonacceptance of the intervention were only qualitatively and sporadically assessed but included the fact that the intervention time was considered too short (eg, patients were on holidays during the entire intervention time) or that patients were not as comfortable in using computers as they had stated during the baseline screening.

Overall, the high dropout rates suggest that the study might have been underpowered. However, more importantly, it also highlights crucial pitfalls of health services interventions indicating that an intervention, as such, might be effective but must be carefully implemented and accepted by the people using it in order to result in effectiveness.

Conclusions

The feasibility of this randomized controlled study was challenging, giving no indication whether the inadequate implementation or insufficient efficacy of the intervention, as such, contributed to the null effect of this study. This has important implications for the proper monitoring of the study quality and the many different steps of a complex intervention; this also stresses the need for meticulously planned and conducted pilot studies testing unrecognized sources of variability before pivotal intervention studies are implemented.

Acknowledgments

CM was affiliated with the Department of General Practice and Health Services Research at Heidelberg University Hospital for most of the duration of the study and is currently affiliated with the Department of Nursing Science at University Hospital Tuebingen.

The study was conducted within the INFOPAT project [14], which was supported by grants from the German Federal Ministry of Education and Research (funding code 01KQ1003B). The authors thank all participating GPs, medical assistants, and patients for participating in the trial. Moreover, we thank all colleagues from the Center for Information Technology and Medical Engineering, University Hospital Heidelberg, who developed the PEPA and were involved in the study conduct as first-level support.

The following are the members of the INFOPAT P4/P5 (Projects 4 and 5) study team: Gerda Bernhard, Department of General Practice and Health Services Research, University of Heidelberg; Christian Jacke, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University; Bärbel Handlos, Heidelberger Selbsthilfebüro; Jens Kaltschmidt, Department of Clinical Pharmacology and Pharmacoepidemiology, University of Heidelberg; Gökce Karakas, Patientenberatung Rhein-Neckar; Christian Kasperk, Department of Internal Medicine I and Clinical Chemistry, University Hospital Heidelberg; Marion Kiel, Department of General Practice and Health Services Research, University of Heidelberg; Jieli Li, Department of Clinical Pharmacology and Pharmacoepidemiology, University of Heidelberg; Michael Metzner, Department of Clinical Pharmacology and Pharmacoepidemiology, University of Heidelberg; Peter Nawroth, Department of Internal Medicine I and Clinical Chemistry, University Hospital Heidelberg; and Hans Joachim Salize, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University.

Authors' Contributions

WEH, CM, JS, and HMS initiated the project and the corresponding research. All authors wrote and reviewed the study protocol. AW, CM, HMS, BS, and MS were involved in the conduct of the study, including contacting and training of primary care practices and training of patients in the intervention group. JK, AW, CM, HMS, and BS handled and analyzed the study data. All authors provided valuable feedback during the study conceptualization, conduct of the study, and analysis of data. All authors had full access to all data, discussed the results, and wrote, reviewed, and approved the final draft of the manuscript.

Conflicts of Interest

HMS, MS, and WEH are involved in the development of AiD*Klinik*, a drug information system developed at Heidelberg University Hospital, serving as the basis for the medication module. WEH is also a shareholder of Dosing GmbH, a spinoff company distributing AiD*Klinik*, and MS is a part-time employee of the same company.

This randomized study was not registered because the authors deemed it unnecessary. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, because the risk of bias appears low and the study was considered formative. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
CONSORT - EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 103 KB - formative_v4i4e13746_app1.pdf]

References

1. Bailey SC, Oramasionwu CU, Wolf MS. Rethinking adherence: A health literacy-informed model of medication self-management. J Health Commun 2013;18 Suppl 1:20-30 [FREE Full text] [doi: 10.1080/10810730.2013.825672] [Medline: 24093342]
2. Persell SD, Osborn CY, Richard R, Skripkauskas S, Wolf MS. Limited health literacy is a barrier to medication reconciliation in ambulatory care. J Gen Intern Med 2007 Nov;22(11):1523-1526 [FREE Full text] [doi: 10.1007/s11606-007-0334-x] [Medline: 17786521]
3. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: A systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
4. Shiyanbola OO, Unni E, Huang Y, Lanier C. The association of health literacy with illness perceptions, medication beliefs, and medication adherence among individuals with type 2 diabetes. Res Social Adm Pharm 2018 Sep;14(9):824-830. [doi: 10.1016/j.sapharm.2017.12.005] [Medline: 29317189]

5. Kale MS, Federman AD, Krauskopf K, Wolf M, O'Conor R, Martynenko M, et al. The association of health literacy with illness and medication beliefs among patients with chronic obstructive pulmonary disease. PLoS One 2015;10(4):e0123937 [FREE Full text] [doi: 10.1371/journal.pone.0123937] [Medline: 25915420]
6. Zhang NJ, Terry A, McHorney CA. Impact of health literacy on medication adherence: A systematic review and meta-analysis. Ann Pharmacother 2014 Jun;48(6):741-751. [doi: 10.1177/1060028014526562] [Medline: 24619949]
7. 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]
8. Ose D, Baudendistel I, Pohlmann S, Winkler EC, Kunz A, Szecsenyi J. Personal health records on the Internet. A narrative review of attitudes, expectations, utilization and effects on health outcomes [Article in German]. Z Evid Fortbild Qual Gesundhwes 2017 May;122:9-21. [doi: 10.1016/j.zefq.2017.04.006] [Medline: 28499764]
9. Sun R, Korytkowski MT, Sereika SM, Saul MI, Li D, Burke LE. Patient portal use in diabetes management: Literature review. JMIR Diabetes 2018 Nov 06;3(4):e11199 [FREE Full text] [doi: 10.2196/11199] [Medline: 30401665]
10. Powell KR. Patient-perceived facilitators of and barriers to electronic portal use: A systematic review. Comput Inform Nurs 2017 Nov;35(11):565-573. [doi: 10.1097/CIN.0000000000000377] [Medline: 28723832]
11. Wade-Vuturo AE, Mayberry LS, Osborn CY. Secure messaging and diabetes management: Experiences and perspectives of patient portal users. J Am Med Inform Assoc 2013 May 01;20(3):519-525 [FREE Full text] [doi: 10.1136/amiajnl-2012-001253] [Medline: 23242764]
12. Irizarry T, Shoemake J, Nilsen ML, Czaja S, Beach S, DeVito Dabbs A. Patient portals as a tool for health care engagement: A mixed-method study of older adults with varying levels of health literacy and prior patient portal use. J Med Internet Res 2017 Mar 30;19(3):e99 [FREE Full text] [doi: 10.2196/jmir.7099] [Medline: 28360022]
13. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: A systematic review. J Med Internet Res 2015 Feb 10;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
14. INFOPAT. URL: https://www.infopat.eu/ [accessed 2020-03-08]
15. Baudendistel I, Winkler E, Kamradt M, Längst G, Eckrich F, Heinze O, et al. Personal electronic health records: Understanding user requirements and needs in chronic cancer care. J Med Internet Res 2015 May 21;17(5):e121 [FREE Full text] [doi: 10.2196/jmir.3884] [Medline: 25998006]
16. Baudendistel I, Winkler E, Kamradt M, Brophy S, Längst G, Eckrich F, et al. The patients' active role in managing a personal electronic health record: A qualitative analysis. Support Care Cancer 2015 Sep;23(9):2613-2621. [doi: 10.1007/s00520-015-2620-1] [Medline: 25652149]
17. Bernhard G, Ose D, Baudendistel I, Seidling HM, Stützle M, Szecsenyi J, et al. Understanding challenges, strategies, and the role of support networks in medication self-management among patients with type 2 diabetes. Diabetes Educ 2017 Apr;43(2):190-205. [doi: 10.1177/0145721717697243] [Medline: 28340544]
18. 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]
19. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health 2013 Jul 16;13:658 [FREE Full text] [doi: 10.1186/1471-2458-13-658] [Medline: 23855504]
20. Nolte S, Osborne RH, Dwinger S, Elsworth GR, Conrad ML, Rose M, et al. German translation, cultural adaptation, and validation of the Health Literacy Questionnaire (HLQ). PLoS One 2017;12(2):e0172340 [FREE Full text] [doi: 10.1371/journal.pone.0172340] [Medline: 28234987]
21. World Health Organization. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
22. Mahler C, Hermann K, Horne R, Ludt S, Haefeli WE, Szecsenyi J, et al. Assessing reported adherence to pharmacological treatment recommendations. Translation and evaluation of the Medication Adherence Report Scale (MARS) in Germany. J Eval Clin Pract 2010 Jun;16(3):574-579. [doi: 10.1111/j.1365-2753.2009.01169.x] [Medline: 20210821]
23. Mahler C, Jank S, Hermann K, Horne R, Ludt S, Haefeli WE, et al. Psychometric properties of a German version of the "Satisfaction with Information about Medicines Scale" (SIMS-D). Value Health 2009;12(8):1176-1179 [FREE Full text] [doi: 10.1111/j.1524-4733.2009.00575.x] [Medline: 19659705]
24. Horne R, Hankins M, Jenkins R. The Satisfaction with Information about Medicines Scale (SIMS): A new measurement tool for audit and research. Qual Health Care 2001 Sep;10(3):135-140 [FREE Full text] [doi: 10.1136/qhc.0100135] [Medline: 11533420]
25. Salize H, Kilian R. Gesundheitsökonomische Analysen in der Psychiatrie. Stuttgart, Germany: Kohlhammer; 2010.
26. Brooke J. SUS: A retrospective. J Usability Stud 2013 Feb;8(2):29-40 [FREE Full text]
27. Mahler C, Freund T, Baldauf A, Jank S, Ludt S, Peters-Klimm F, et al. Structured medication management in primary care: A tool to promote medication safety [Article in German]. Z Evid Fortbild Qual Gesundhwes 2014;108(5-6):258-269. [doi: 10.1016/j.zefq.2013.07.011] [Medline: 25066344]

28. Dodson S, Good S, Osborne RH. Health Literacy Toolkit for Low- and Middle-Income Countries: A Series of Information Sheets to Empower Communities and Strengthen Health Systems. New Delhi, India: World Health Organization, Regional Office for South-East Asia; 2015. URL: http://apps.searo.who.int/PDS_DOCS/B5148.pdf [accessed 2020-03-17]
29. Schaeffer D, Vogt D, Berens EM, Hurrelmann K. Gesundheitskompetenz der Bevölkerung in Deutschland. Bielefeld, Germany: Universität Bielefeld, Fakultät für Gesundheitswissenschaften; 2016 Dec. URL: https://pub.uni-bielefeld.de/download/2908845/2908882 [accessed 2020-03-17]
30. Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G, HLS-EU Consortium. Health literacy in Europe: Comparative results of the European Health Literacy Survey (HLS-EU). Eur J Public Health 2015 Dec;25(6):1053-1058 [FREE Full text] [doi: 10.1093/eurpub/ckv043] [Medline: 25843827]
31. Rudd RE, Blanch DC, Gall V, Chibnik LB, Wright EA, Reichmann W, et al. A randomized controlled trial of an intervention to reduce low literacy barriers in inflammatory arthritis management. Patient Educ Couns 2009 Jun;75(3):334-339 [FREE Full text] [doi: 10.1016/j.pec.2009.03.001] [Medline: 19345053]
32. Bosworth HB, Olsen MK, Gentry P, Orr M, Dudley T, McCant F, et al. Nurse administered telephone intervention for blood pressure control: A patient-tailored multifactorial intervention. Patient Educ Couns 2005 Apr;57(1):5-14. [doi: 10.1016/j.pec.2004.03.011] [Medline: 15797147]
33. Murray MD, Young J, Hoke S, Tu W, Weiner M, Morrow D, et al. Pharmacist intervention to improve medication adherence in heart failure: A randomized trial. Ann Intern Med 2007 May 15;146(10):714-725. [doi: 10.7326/0003-4819-146-10-200705150-00005] [Medline: 17502632]
34. Friis K, Vind BD, Simmons RK, Maindal HT. The relationship between health literacy and health behaviour in people with diabetes: A Danish population-based study. J Diabetes Res 2016;2016:7823130 [FREE Full text] [doi: 10.1155/2016/7823130] [Medline: 27761473]
35. Beauchamp A, Buchbinder R, Dodson S, Batterham RW, Elsworth GR, McPhee C, et al. Distribution of health literacy strengths and weaknesses across socio-demographic groups: A cross-sectional survey using the Health Literacy Questionnaire (HLQ). BMC Public Health 2015 Jul 21;15:678 [FREE Full text] [doi: 10.1186/s12889-015-2056-z] [Medline: 26194350]

Abbreviations

EHR: electronic health record
GP: general practitioner
HLQ: Health Literacy Questionnaire
HLS-Ger: Health Literacy Survey, German version
INFOPAT: INFOrmation Technology for PATient-oriented Healthcare in the Rhine-Neckar metropolitan region
MARS-D: Medication Adherence Report Scale, German version
MMRV: Mannheimer Modul Ressourcenverbrauch
P4/P5: Projects 4 and 5
PEPA: personal EHR
SIMS-D: Satisfaction with Information about Medicines Scale, German version
SUS: System Usability Scale
T0: baseline visit
T1: follow-up visit
T2DM: type 2 diabetes mellitus
WHOQOL-BREF: World Health Organization Quality of Life instrument, brief version

Edited by G Eysenbach; submitted 19.02.19; peer-reviewed by K Powell, S Molinos; comments to author 14.06.19; revised version received 19.07.19; accepted 02.09.19; published 28.04.20.

Please cite as:
Seidling HM, Mahler C, Strauß B, Weis A, Stützle M, Krisam J, INFOPAT P4/P5 Study Team, Szecsenyi J, Haefeli WE
An Electronic Medication Module to Improve Health Literacy in Patients With Type 2 Diabetes Mellitus: Pilot Randomized Controlled Trial
JMIR Form Res 2020;4(4):e13746
URL: http://formative.jmir.org/2020/4/e13746/
doi:10.2196/13746
PMID:32343246

Original Paper

Usability of Wearable Devices to Remotely Monitor Sleep Patterns Among Patients With Ischemic Heart Disease: Observational Study

Michael Fortunato[1], BA; Srinath Adusumalli[1], MD; Neel Chokshi[1], MBA, MD; Joseph Harrison[1], BSc; Charles Rareshide[1], MSc; Mitesh Patel[1,2], MBA, MD

[1]Crescenz Veterans Affairs Medical Center, Philadelphia, PA, United States

[2]University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Mitesh Patel, MBA, MD
Crescenz Veterans Affairs Medical Center
3800 Woodland Ave
South Pavilion 14-176
Philadelphia, PA,
United States
Phone: 1 215 823 5800
Email: mpatel@pennmedicine.upenn.edu

Abstract

Background: There is growing interest in using wearable devices to remotely monitor patient behaviors. However, there has been little evaluation of how often these technologies are used to monitor sleep patterns over longer term periods, particularly among more high-risk patients.

Objective: The goal of the research was to evaluate the proportion of time that patients with ischemic heart disease used wearable devices to monitor their sleep and identify differences in characteristics of patients with higher versus lower use.

Methods: We evaluated wearable device data from a previously conducted clinical trial testing the use of wearable devices with personalized goal-setting and financial incentives. Patients with ischemic heart disease established a sleep baseline and were then followed for 24 weeks. The proportion of days that sleep data was collected was compared over the 24 weeks and by study arm. Characteristics of patients were compared to groups with high, low, or no sleep data.

Results: The sample comprised 99 patients with ischemic heart disease, among which 79% (78/99) used the wearable device to track their sleep. During the 6-month trial, sleep data were collected on 60% (10,024/16,632) of patient-days. These rates declined over time from 77% (4292/5544) in months 1 and 2 to 58% (3188/5544) in months 3 and 4 to 46% (2544/5544) in months 5 and 6. Sleep data were collected at higher rates among the intervention group compared with control (67% vs 55%, $P<.001$). In the main intervention period (months 3 and 4), patients with higher rates of sleep data were on average older ($P=.03$), had a history of smoking ($P=.007$), and had higher rates of commercial health insurance ($P=.03$).

Conclusions: Among patients with ischemic heart disease in a physical activity trial, a high proportion used wearable devices to track their sleep; however, rates declined over time. Future research should consider larger evaluations coupled with behavioral interventions.

Trial Registration: ClinicalTrials.gov NCT02531022; https://clinicaltrials.gov/ct2/show/NCT02531022

(JMIR Form Res 2020;4(4):e14508) doi:10.2196/14508

KEYWORDS

sleep; wearable devices; ischemic heart disease

Introduction

Shorter sleep duration and poor sleep quality have been demonstrated to be associated with higher rates of all-cause mortality, cardiovascular disease, hypertension, and obesity [1,2]. Most of these evaluations have relied on patient self-report of sleep patterns, which requires effort from the patient and can be subject to reporter bias. There has been growing interest in

using wearable devices to passively collect data on patient behaviors [3,4]. However, there has been little evaluation of how often these technologies are used to remotely monitor sleep patterns, particularly among more high-risk patients and over longer term periods.

There are more than 50 different wearable devices that promote the ability to track sleep patterns [5]. In a recent review article, 43 articles identified studied how these wearables were used to track sleep patterns [6]. More than half were focused on validating sleep accuracy; few were focused on monitoring sleep during behavioral interventions. Several studies have evaluated the use of these devices over 24-hour periods [7,8]. Others have evaluated tracking sleep during behavioral interventions focused on physical activity [6,9], but these have typically been for periods of 3 months or less.

In this study, we used data from a behavioral intervention focused on increasing physical activity among ischemic heart disease patients over 24 weeks. The objective was to evaluate the proportion of time that patients with ischemic heart disease used wearable devices to monitor their sleep and identify differences in characteristics of patients with higher versus lower use.

Methods

The sample comprised patients with ischemic heart disease who used wearable devices to establish baseline sleep levels during the ACTIVE REWARD (A Clinical Trial Investigating Effects of a Randomized Evaluation of Wearable Activity Trackers with Financial Rewards) trial, a previously conducted 24-week randomized clinical trial focused on increasing physical activity [10]. All patients established baseline daily step counts and were then randomly assigned to passive monitoring or an intervention that used personalized goal-setting and financial incentives to increase physical activity levels. In this study, we included patients who established baseline levels of sleep during the run-in period (99/105 patients). Financial incentives for meeting step goals were offered to the intervention group for 16 weeks followed by 8 weeks without incentives. Patients were asked to use the wearable devices (Shine, Misfit) during the day and night. The wearable did not require charging (with a battery life longer than 6 months), was waterproof, and displayed progress toward the step goal (rather than the actual number of steps) on its display. Patients could use the mobile app to obtain the actual number of steps. The Shine has been found to be reliable for monitoring sleep duration when compared with polysomnography [11].

To evaluate use of wearable devices for tracking sleep data, we analyzed the proportion of patient-days data that were collected overall and during 8 week increments throughout the three trial phases. Intervention patients had 3 phases: ramp-up phase with incentives (valued at $2 per day) and gradually increasing step goals, maintenance phase with incentives and static step goals, and a follow-up phase with static step goals but no incentives. We compared patient characteristics for groups of patients with different levels of overall data collection above and below half of the study period duration (more than 50% of days with data, less than 50% of days with data, no data).

This study was approved by the University of Pennsylvania institutional review board, and patients provided informed consent. The study was registered with ClinicalTrials.gov [NCT02531022]. Analyses were conducted in SAS version 9.4 (SAS Institute Inc).

Results

The sample comprised 99 patients with ischemic heart disease, with 79% (78/99) using the wearable device to track their sleep. During the 6-month trial, sleep data were collected on 60% of patient-days (Table 1). These rates declined over time from 77% (10,024/16,632) in months 1 and 2 to 58% (3188/5544) in months 3 and 4 to 46% (2544/5544) in months 5 and 6. Sleep data were collected at higher rates among the intervention group compared with control (67% vs 55%; $P<.001$). In the main intervention period (months 3 and 4), patients with higher rates of sleep data were on average older, were less likely to be actively smoking, and had higher rates of private health insurance (Table 2).

Table 1. Proportion of patient-days that sleep data was collected by period and arm.

Trial phase	Control (n=2912), n (%)	Intervention (n=2632), n (%)
Ramp-up period: weeks 1-8	2170 (75.52)	2122 (80.62)
Maintenance period: weeks 9-16	1444 (49.59)	1744 (66.26)
Follow-up period: weeks 17-24	1153 (39.59)	1391 (52.85)

Table 2. Patient characteristics by use of wearable devices to track sleep. Sleep data are based on the main intervention period (weeks 9 to 16) of the trial.

Characteristics	≥50% sleep data collected (n=60)	<50% sleep data collected (n=18)	No sleep data collected (n=21)	*P* value
Sociodemographics				
Age in years, mean (SD)	62 (9.2)	55.1 (12.4)	55.8 (11.5)	.03
Male, n (%)	40 (67)	13 (72)	15 (71)	.86
Race/ethnicity, n (%)				.32
White non-Hispanic	49 (82)	11 (61)	14 (67)	
Black non-Hispanic	8 (13)	5 (28)	6 (29)	
Other	3 (5)	2 (11)	1 (5)	
Education, n (%)				.77
Some high school	3 (5)	2 (11)	1 (5)	
High school graduate	12 (20)	5 (28)	4 (19)	
Some college or specialized training	13 (22)	3 (17)	8 (38)	
College graduate	31 (52)	8 (44)	8 (38)	
Missing	1 (2)	0 (0)	0 (0)	
Marital status, n (%)				.37
Single	12 (20)	5 (28)	6 (29)	
Married	40 (67)	8 (44)	13 (62)	
Other	8 (13)	5 (28)	2 (10)	
Insurance, n (%)				.03
Private	35 (58)	5 (28)	9 (43)	
Medicare	23 (38)	9 (50)	11 (52)	
Medicaid	1 (2)	4 (22)	1 (5)	
Military	1 (2)	0 (0)	0 (0)	
Annual household income, n (%)				.74
Less than $50,000	20 (33)	9 (50)	7 (33)	
$50,000 to $100,000	12 (20)	4 (22)	6 (29)	
Greater than $100,000	18 (30)	4 (22)	6 (29)	
Missing	10 (17)	1 (6)	2 (10)	
Baseline measures				
Baseline step count, mean (SD)	7214.5 (3618.2)	6617.7 (2584.1)	5481.2 (1808.4)	.13
Body mass index, mean (SD)	30.1 (5.9)	29.6 (5.7)	32 (6.2)	.35
Diabetes, n (%)	16 (27)	4 (22)	11 (52)	.06
Hypertension, n (%)	49 (82)	16 (89)	17 (81)	.75
Hyperlipidemia, n (%)	51 (85)	14 (78)	16 (76)	.59
Smoking history, n (%)				.007
Nonsmoker	30 (50)	10 (56)	5 (24)	
History of smoking	29 (48)	5 (28)	11 (52)	
Actively smoking	1 (2)	3 (17)	5 (24)	

Discussion

Principal Findings

There is growing evidence that our sleep patterns influence our longer term health, with poor sleep associated with higher risk for cardiovascular disease [1,2]. Therefore, new ways to collect data on an individual's sleep could be important to help inform the design of future interventions. To our knowledge, this is one of the first studies to evaluate sleep data collected by wearable devices in a longer 24-week clinical trial.

Our findings reveal several important insights. First, a high proportion of patients used wearable devices to track their sleep; however, rates declined over time. Yet similar to a previous study [9], it is important to recognize that patients were enrolled in a trial focusing on physical activity, not sleep. Therefore, while rates of sleep data were high, they could be increased with a greater intervention emphasis on sleep. Second, we found that sleep data were collected at higher rates among patients in the intervention arm relative to control. This may be because engagement was more related to behavioral reasons (ie, motivation vs loss of interest over time) rather than technical issues (eg, device failure). Taken together, this indicates that use of wearable devices may increase by combined provision of the technology with a behavior change strategy. A recent review suggested that more research is needed on using wearables to monitor sleep in behavioral interventions [6]. Third, use varied across several patient characteristics. Notably, use was significantly higher among older patients, those with a history of smoking, and those with commercial health insurance. This indicates that sleep data may be obtainable from high-risk, older patient groups. Similar results were seen in a shorter, 24-hour study [7], and our findings demonstrate similar insights over a 24-week period. These may need to be investigated in future trials with larger and more diverse populations.

Strengths and Limitations

This study had several strengths. First, most recent studies have focused on comparing the accuracy of wearable devices in tracking sleep patterns rather than their successful implementation in new behavioral intervention strategies [6]. Second, most clinical studies assessing sleep quantity and quality to date have relied on either self-reported measures of sleep duration or on polysomnography measurements. Polysomnography remains the gold standard for sleep measurement, but its use is currently only feasible in a laboratory setting, and studies typically use only 1 or 2 nights of measurement. Self-reported sleep measurements limit the generalizability of any results because of errors due to reporter bias.

This study has limitations. First, it was conducted within a single health system in a clinical trial, and we limited the sample to those patients who obtained sleep baselines (approximately 94% of patients). Second, the trial focused more on physical activity than on sleep. Third, we were not adequately powered to perform evaluations of differences in sleep patterns. Fourth, we did not have qualitative data on patient perspectives of barriers and facilitators to using these devices to monitor sleep. Fifth, while the wearable device used has been found to be accurate for monitoring sleep duration [11], we did not validate it in this study. Sixth, further work is needed to understand what proportion of days devices need to be used to provide robust assessments of sleep patterns.

Conclusion

In conclusion, a high proportion of patients with ischemic heart disease used wearable devices to track sleep patterns over a 24-week period. Use declined over time but varied based on patient characteristics and was greater in the intervention group than in control. Future research should consider larger evaluations combining wearable devices with behavioral interventions to test ways to risk stratify patients and improve sleep patterns.

Acknowledgments

This study was supported in part by grant number UL1TR000003 from the National Center for Advancing Translational Science. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Advancing Translational Science or the National Institutes of Health. The study was also supported in part by the Institute for Translational Medicine and Therapeutics and the University of Pennsylvania Health System through the Penn Medicine Nudge Unit. Misfit donated the wearable devices used in this study. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. MF and MP had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

MP is supported by career development awards from the Department of Veterans Affairs Health Services Research and Development and the Doris Duke Charitable Foundation. MP is founder of Catalyst Health, a technology and behavior change consulting firm. MP also has received research funding from Deloitte, which is not related to the work described in this manuscript. The remaining authors declare no conflict of interest.

References

1. Itani O, Jike M, Watanabe N, Kaneita Y. Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. Sleep Med 2017 Apr;32:246-256. [doi: 10.1016/j.sleep.2016.08.006] [Medline: 27743803]
2. Domínguez F, Fuster V, Fernández-Alvira J, Fernández-Friera L, López-Melgar B, Blanco-Rojo R, et al. Association of sleep duration and quality with subclinical atherosclerosis. J Am Coll Cardiol 2019 Jan 22;73(2):134-144 [FREE Full text] [doi: 10.1016/j.jacc.2018.10.060] [Medline: 30654884]
3. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 03;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
4. Patel MS, Foschini L, Kurtzman GW, Zhu J, Wang W, Rareshide CA, et al. Using wearable devices and smartphones to track physical activity: initial activation, sustained use, and step counts across sociodemographic characteristics in a national sample. Ann Intern Med 2017 Nov 21;167(10):755-757. [doi: 10.7326/M17-1495] [Medline: 28973116]
5. Shelgikar AV, Anderson PF, Stephens MR. Sleep tracking, wearable technology, and opportunities for research and clinical care. Chest 2016 Sep;150(3):732-743. [doi: 10.1016/j.chest.2016.04.016] [Medline: 27132701]
6. Baron KG, Duffecy J, Berendsen MA, Cheung Mason I, Lattie EG, Manalo NC. Feeling validated yet? A scoping review of the use of consumer-targeted wearable and mobile technology to measure and improve sleep. Sleep Med Rev 2018 Aug;40:151-159 [FREE Full text] [doi: 10.1016/j.smrv.2017.12.002] [Medline: 29395985]
7. Kroll RR, McKenzie ED, Boyd JG, Sheth P, Howes D, Wood M, WEARable Information Technology for hospital INpatients (WEARIT-IN) study group. Use of wearable devices for post-discharge monitoring of ICU patients: a feasibility study. J Intensive Care 2017;5:64 [FREE Full text] [doi: 10.1186/s40560-017-0261-9] [Medline: 29201377]
8. Rosenberger M, Buman M, Haskell W, McConnell M, Carstensen L. Twenty-four hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Med Sci Sports Exerc 2016 Mar;48(3):457-465 [FREE Full text] [doi: 10.1249/MSS.0000000000000778] [Medline: 26484953]
9. Melton BF, Buman MP, Vogel RL, Harris BS, Bigham LE. Wearable devices to improve physical activity and sleep. J Black Studies 2016 Jul 27;47(6):610-625. [doi: 10.1177/0021934716653349]
10. Chokshi NP, Adusumalli S, Small DS, Morris A, Feingold J, Ha YP, et al. Loss-framed financial incentives and personalized goal-setting to increase physical activity among ischemic heart disease patients using wearable devices: the ACTIVE REWARD randomized trial. J Am Heart Assoc 2018 Jun 13;7(12) [FREE Full text] [doi: 10.1161/JAHA.118.009173] [Medline: 29899015]
11. Mantua J, Gravel N, Spencer R. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. Sensors (Basel) 2016 May 05;16(5) [FREE Full text] [doi: 10.3390/s16050646] [Medline: 27164110]

Abbreviations

ACTIVE REWARD: A Clinical Trial Investigating Effects of a Randomized Evaluation of Wearable Activity Trackers with Financial Rewards

Edited by G Eysenbach; submitted 28.04.19; peer-reviewed by J Sasaki, D Maslove, C Poon; comments to author 14.12.19; revised version received 20.01.20; accepted 09.02.20; published 07.04.20.

Please cite as:
Fortunato M, Adusumalli S, Chokshi N, Harrison J, Rareshide C, Patel M
Usability of Wearable Devices to Remotely Monitor Sleep Patterns Among Patients With Ischemic Heart Disease: Observational Study
JMIR Form Res 2020;4(4):e14508
URL: https://formative.jmir.org/2020/4/e14508
doi:10.2196/14508
PMID:32254044

Original Paper

Development and Modification of a Mobile Health Program to Promote Postpartum Weight Loss in Women at Elevated Risk for Cardiometabolic Disease: Single-Arm Pilot Study

Jacinda M Nicklas[1], MA, MD, MPH; Jenn A Leiferman[2], PhD; Steven Lockhart[3], MPH; Kristen M Daly[2], MPH, MSc, PhD; Sheana S Bull[2], MPH, PhD; Linda A Barbour[4,5], MD, MSPH

[1]Division of General Internal Medicine, University of Colorado School of Medicine, Aurora, CO, United States

[2]Community and Behavioral Health, Colorado School of Public Health, Aurora, CO, United States

[3]Adult and Child Consortium for Health Outcomes Research and Delivery Science, University of Colorado School of Medicine, Aurora, CO, United States

[4]Division of Endocrinology, Metabolism and Diabetes, University of Colorado School of Medicine, Aurora, CO, United States

[5]Division of Maternal Fetal Medicine, University of Colorado School of Medicine, Aurora, CO, United States

Corresponding Author:
Jacinda M Nicklas, MA, MD, MPH
Division of General Internal Medicine
University of Colorado School of Medicine
12348 E Montview Blvd
Aurora, CO, 80045
United States
Phone: 1 3037249028
Email: Jacinda.Nicklas@cuanschutz.edu

Abstract

Background: Pregnancy complications in combination with postpartum weight retention lead to significant risks of cardiometabolic disease and obesity. The majority of traditional face-to-face interventions have not been effective in postpartum women. Mobile technology enables the active engagement of postpartum women to promote lifestyle changes to prevent chronic diseases.

Objective: We sought to employ an interactive, user-centered, and participatory method of development, evaluation, and iteration to design and optimize the mobile health (mHealth) *Fit After Baby* program.

Methods: For the initial development, a multidisciplinary team integrated evidence-based approaches for health behavior, diet and physical activity, and user-centered design and engagement. We implemented an iterative feedback and design process via 3 month-long beta pilots in which postpartum women with cardiometabolic risk factors participated in the program and provided weekly and ongoing feedback. We also conducted two group interviews using a structured interview guide to gather additional feedback. Qualitative data were recorded, transcribed, and analyzed using established qualitative methods. Modifications based on feedback were integrated into successive versions of the app.

Results: We conducted three pilot testing rounds with a total of 26 women. Feedback from each pilot cohort informed changes to the functionality and content of the app, and then a subsequent pilot group participated in the program. We optimized the program in response to feedback through three iterations leading to a final version.

Conclusions: This study demonstrates the feasibility of using an interactive, user-centered, participatory method of rapid, iterative design and evaluation to develop and optimize a mHealth intervention program for postpartum women.

Trial Registration: ClinicalTrials.gov NCT02384226; https://www.clinicaltrials.gov/ct2/show/NCT02384226

(JMIR Form Res 2020;4(4):e16151) doi:10.2196/16151

KEYWORDS

mobile health; postpartum; chronic disease; prevention; weight loss

Introduction

Pregnancy as a Stress Test to Expose Predilection to Chronic Disease

Certain pregnancy complications provide an early warning of future cardiometabolic risk [1]. Women with pregnancies complicated by gestational diabetes mellitus (GDM) have an approximately 50% increased risk for developing type 2 diabetes mellitus within 10 years, are likely to develop atherosclerosis earlier [2], and have an increased risk for hypertension [3] and cardiovascular disease (CVD) [4,5]. Preeclampsia, preterm delivery, delivery of a small-for-gestational age (SGA) neonate, hypertensive disorders in pregnancy, and GDM are independently associated with a 50%-300% increased risk for CVD [5]. About 30% of US women will have at least one of these predictive conditions during pregnancy [4].

The Postpartum Period Is a Critical Window of Opportunity for Primary Prevention

Previous studies demonstrate that pregnancy weight retained beyond 6 to 12 months postpartum is usually retained long term and is a powerful independent risk factor for future obesity [6]. Given the significance of postpartum weight retention, the postpartum year is considered a critical window of opportunity to make lifestyle changes to decrease future risk of obesity and chronic disease [7-9]. Lifestyle changes, including weight loss, smoking cessation, improved diet, and physical activity can decrease the risk of diabetes and CVD [10-14]. Reducing postpartum weight retention also decreases the risk of weight-related complications in future pregnancies [15,16]. Unfortunately, most women do not return to their prepregnancy weight postpartum, and with each subsequent pregnancy, their risk of these obesity-related complications amplifies. However, previous studies demonstrate that postpartum women may be receptive to making lifestyle changes given their new awareness of risk factors that were unmasked during pregnancy as well as their motivation to create a healthy home for their offspring [17].

A Lack of Available Treatment Options Tailored to Postpartum Mothers

Despite the importance and critical timing of the postpartum period, there are currently no clinically available evidence-based programs designed for overweight and obese postpartum women with recent pregnancy complications. Few lifestyle intervention studies have been conducted in postpartum women at elevated cardiometabolic risk. Studies attempting intensive face-to-face methodologies for weight loss and reduction in diabetes incidence similar to the successful Diabetes Prevention Program (DPP) demonstrate limited efficacy and poor retention in postpartum women [18-20]. This is due, at least in part, to multiple barriers to face-to-face participation described by postpartum women, including time constraints, infant and breastfeeding demands, older childcare responsibilities, and reluctance to spend time away from family [21,22]. Given these barriers, there is increasing interest in using technology to improve the efficacy of lifestyle interventions for this high-risk population [20,23].

Use of Mobile Health

As of 2015, 82% of the US population aged 18 to 49 years owned an app-enabled mobile phone [24]. Women of childbearing age are one of the fastest growing user groups for mobile phones, across race and socioeconomic class. Mobile technology facilitates tracking of behavior and weight, allowing for real-time recording, feedback, and accountability. In light of escalating health care costs and rapid increases in the incidence of cardiometabolic disease, extending the reach of health promotion into daily life is an innovative approach for high-risk women with multiple and intensive family/work demands. However, despite the potential of this technology, the vast majority of available apps do not adequately include evidence-based strategies [25-27] or use behavioral theory [28,29], and almost none have been rigorously tested [30].

This study describes the iterative development process designed to optimize the mobile health (mHealth) *Fit After Baby* program. We employed an interactive user-centered and participatory method of rapid development, evaluation, and iteration [31-33]. We designed the *Fit After Baby* program to decrease obesity and risk factors for chronic disease by increasing postpartum weight loss, improving diet, and increasing physical activity.

Methods

Development of the Fit After Baby Program

We developed the mHealth *Fit After Baby* program using evidence-based strategies for weight loss, cardiometabolic disease prevention, and behavior change based on current evidence and best practices [34]. Guided by the Integrated Theory of mHealth [35], we incorporated a theoretical framework that included traditional health communication and behavioral theories including the *elaboration likelihood* model [36], the theory of planned behavior [37], the life-course approach [38,39], and Bandura's model of social cognitive theory for behavioral change [40], as well as engagement strategies from design-thinking, user-centered design, and mobile technology in health promotion [41-45]. This inclusion of multiple and diverse theoretical perspectives in app design is critical to the creation of a mobile solution that is engaging, effective, and scalable. The content and structure were adapted from the DPP specifically for a postpartum population, and we incorporated features and techniques that have proven efficacy for weight loss such as self-monitoring, goal setting, remote coaching with tailored feedback, and social support [46-49]. We designed the *Fit After Baby* app to integrate with the commercial apps Fitbit and MyFitnessPal. Gamification through points and badges was used to motivate behaviors through accountability and reinforcement [50]. The use of game-like components in health and nutrition apps is widespread and some initial studies demonstrate that women in particular may derive social benefit from gamification [51-53]. The *Fit After Baby* program also included the engagement strategies of push notifications to remind or trigger the user to interact with the app, as well as *favoriting*, in which users could curate their own health library for future use and thereby personalize the app.

A registered dietitian with experience in both the DPP and motivational interviewing served as the lifestyle coach for the

Fit After Baby program. She based her conversations on the weekly content contained in the app but was encouraged to follow the participants' lead if they wanted to discuss other topics. The lifestyle coach kept detailed notes on her conversations and reviewed any issues weekly with the principal investigator.

Iterative Feedback and Design Process

Rapid iterative design is a process in which technology is tested and repeatedly refined with small groups of users to optimize functionality and usability [54,55]. We refined the content and delivery of the mHealth *Fit After Baby* program over three rounds of beta-testing through an iterative design process, and then we developed a final version of the program.

Recruitment

Women between 18 and 40 years of age with a postpartum BMI of 24 to 45 kg/m^2 who were within 6 months of a recent singleton or twin delivery complicated by gestational hypertension, preeclampsia, preterm delivery (32-37 weeks), an SGA neonate (weight <10th percentile for gestational age), and/or GDM were recruited for the study from women delivering at the University of Colorado Hospital on the Anschutz Medical Campus in Aurora, Colorado. Women were required to have access to an iPhone or iPod (Apple Inc, California) touch ≥5S, and women with a history of preexisting diabetes, cancer, or CVD were excluded. The institutional review board at the University of Colorado approved the study, and all patients gave written informed consent.

Pilot Testing

We conducted three 4-week beta-testing pilots with unique participants. The participants used the *Fit After Baby* app for 4 weeks, received support from the lifestyle coach, and provided ongoing feedback. Participants were required to open the app and provide online feedback in week 1 to receive a Fitbit by mail to use for the remainder of the study. Participants were asked to log on to a Web-based asynchronous user group platform each week and answer a set of questions in threaded discussions. Interactions with the app and threaded discussions were tracked by researchers. The three 4-week pilots were conducted during the time period June 2015 through February 2016.

Feedback and Iterative Design

After each round, we analyzed feedback using content analysis, and changes were made to the mobile app and program before conducting subsequent rounds. Consequently, each subsequent round had a new iteration of the app and program. We held two in-person group structured interviews lasting 90 min after all three rounds were completed. Participants from each round were included. Group interviews were digitally recorded and transcribed verbatim.

Analysis

We used an iterative and team-based process guided by qualitative content analysis [56,57]. A qualitatively trained analyst and principal investigator both inductively and deductively developed the code book. Initial codes were based on the interview guide domains, and the code book was expanded based on codes that emerged from the data. The analyst and investigator jointly reviewed and coded the threaded discussions and group interviews until no new codes were identified and there was strong code assignment agreement. All transcripts were independently read, double coded, and then merged before analysis. Any discrepancies in coding were addressed through discussion and consensus among the coders. Throughout the analytic process, the analyst and principal investigator met regularly to check new findings, discuss emergent new codes and themes, and assess the preliminary and final results. ATLAS.ti version 8.0 was used for data organization and management. Data from all rounds and group interviews were used to design a final version of the mobile app and program.

Results

Demographic, Tracking, and Online Feedback Data for Beta Testing

Table 1 demonstrates the characteristics of the participants who participated in beta testing. The median age for beta testing was 33 years, with a median BMI in the obese range of 31.8 kg/m^2. The majority of participants were white, and preterm birth, pre-eclampsia, and gestational hypertension were the most common pregnancy complications.

Table 2 shows the user data for the mobile app. The beta-testers opened the app an average of 21 out of 28 days. Participants tracked a median of 14 of 21 days of suggested diet tracking on MyFitnessPal, and tracked physical activity using Fitbit or the exercise tracker within the app for a median of 16 of 21 suggested days. We recommended that participants check in with their coach and enter their weight once per week. On average, participants checked in three times over the course of the program and weighed in five times. The overall mean weight loss was 4.3 (SD 2.3) lbs over the 4-week period. The first beta pilot round included 4 women, and they were actively engaged in the program and app. Many changes were made to the app between rounds 1 and 2 in response to feedback and user data. In round 2, there was more variable participation, with some women engaging much less in the app. After we conducted the next round of iterative changes, the engagement increased again in round 3.

Table 1. Demographic data of participants in three rounds of beta-testing of the mobile health Fit After Baby program.

Characteristics of beta-testers	Round 1 (n=4)	Round 2 (n=13)	Round 3 (n=9)	All groups (n=26)
Age (years), median (IQR)	34 (33.3-37)	33 (31-34.5)	30 (28-33.5)	33 (30.75-34)
Race, n (%)				
White	4 (100)	12 (92)	8 (89)	24 (92)
Asian	0 (0)	1 (8)	1 (11)	2 (8)
Hispanic or Latino	0 (0)	0 (0)	1 (11)	1(4)
BMI (kg/m^2), median (IQR)	35.9 (30.2-41.1)	31.2 (29.1-40.1)	32.3 (28.6-35.9)	31.8 (30.1-40.4)
Pregnancy complication, n (%)				
Gestational diabetes	1 (25)	4 (31)	2 (22)	7 (27)
Pre-eclampsia	1 (25)	6 (46)	2 (22)	9 (35)
Gestational hypertension	1 (25)	2 (15)	5 (56)	8 (31)
Small-for-gestational age	0 (0)	2 (15)	2 (22)	4 (15)
Preterm birth	1 (25)	5 (38)	2 (22)	8 (31)

Table 2. User data for three rounds of beta-testing of the mobile health Fit After Baby program.

Beta-testers	Round 1 (n=4)	Round 2 (n=13)	Round 3 (n=9)	All groups (n=26)
Days app launched (out of 28), median (IQR)	24 (16-27)	16 (12-24)	21 (16-25)	20.5 (13-25)
Days diet tracked (out of 21), median (IQR)	17 (12-21)	14 (0-18)	10 (0-18)	14 (0-18)
Days steps or exercise tracked, median (IQR)	25 (19-27)	12 (5-17)	19 (7-22)	16 (9-22)
Number of surveys answered, median (IQR)	11 (9-13)	10 (2-12)	9 (5-12)	10 (4-12)
Coach check ins, median, (IQR)	2 (1-3)	1 (0-4)	4 (2.5-4)	3 (1-4)
Weigh-ins, median (IQR)	4 (4-6)	6 (4-11)	5 (4-7)	5 (4-7)
Weight loss in lbs, mean (SD)	4.4 (1.6)	4.3 (2.8)	4.1 (2.0)	4.3 (2.3)

Feedback and Iterative Development

Feedback addressed usability, navigability, content, and function, and certain themes arose which we used to develop changes in subsequent iterations of the app and program. Multimedia Appendix 1 details the iterative changes made throughout the rounds, including representative quotes associated with the changes.

Navigation

One major theme addressed through successive iterations was navigation. The round 1 version opened to a content screen tailored for that particular day, and participants expressed that they did not know where they were in the app and they were not clear about what they needed to do each day. In response, we developed a home screen consisting of a task list. Tapping on tasks in the list navigated the participant to the screen for that task, and a checkmark automatically appeared once the participant had completed a particular task. New participants in subsequent rounds were enthusiastic about checkmarks and the improved navigation. The final navigation strategy developed after round 3 employed rotating task cards that linked with tasks in the apps and retained the function of automatically checking off completed tasks.

Diet Tracking

We asked participants to use MyFitnessPal for diet tracking. Participants expressed some frustration with tracking of dietary intake. Participants had to download the MyFitnessPal app and open the app to record food and drink, and they complained that it was difficult to switch back and forth between apps. As has been shown previously, tracking of dietary intake is difficult and cumbersome, and can be difficult to maintain for long periods of time [58]. Several participants expressed how difficult it was to track dietary intake while taking care of an infant. In the ultimate version of the program, we elected to ask patients to track diet for only certain days coordinated with a particular dietary theme (ie, tracking saturated fat during week 6 which focuses on dietary fat). However, we retained the option to track every day and to earn points for this. In the final version, we also decided to ask participants to use the Fitbit app for both diet and exercise tracking, so they would not have to download or navigate to the MyFitnessPal app, and therefore there would be less back and forth between apps.

Tracking and Syncing of Physical Activity

Participants were asked to use Fitbit for step tracking, and there was also an exercise tracker built into the *Fit After Baby* app. Participants were sent a Fitbit Zip at the beginning of week 2, if they had been using the app. They needed to download the Fitbit app and connect it with their Fitbit device. Although

participants did not actually need to open the Fitbit app once it was downloaded, they complained that their steps were not always syncing qith the Fit AFter Baby app. The majority of participants stated that they would have preferred a wrist band Fitbit, so we transitioned to this for the final version of the program. The wrist Fitbit (Fitbit Flex 2) is also waterproof, so we felt that this would encourage adherence since some participants told us they forgot to remove the clip Fitbits from their clothes before washing them. We retained the ability to log workouts in the *Fit After Baby* app specifically to give participants the ability to keep track and earn points if they forgot to wear their Fitbit.

Content

Participants in all three rounds requested more content, particularly content tailored specifically to postpartum women. In round 1, the most common requests for additional content included more information about breastfeeding, working, and postpartum mental health. This content was added before round 2, and group 2 also requested additional content about breastfeeding, as well as healthy recipes and information about exercising with an infant. Although we added this content, round 3 participants requested even more exercise information as well as recipes and more detailed diet information. For the final version of the program, we added additional exercise content, multiple exercises with instructions and pictures demonstrating ways to exercise with an infant, a yoga curriculum, and a more extensive recipe collection.

Photos/Graphics

Participants from all three rounds requested more diversity in photos to better reflect their own postpartum experiences. They expressed that women who look like thin models did not reflect what it was like to be a postpartum mom struggling to lose weight. However, some women mentioned that the photos were appropriate, so at each round, we worked to diversify some images but kept others. A few women also commented that the women in the photos looked too calm and relaxed, and they wanted to see some images that would reflect the stress of parenting a newborn. In response, we added additional images to reflect the broad range of emotional experiences postpartum.

Points and Rewards

Participants in round 1 felt that the points and badges had little meaning and it was not clear how points were earned. Before round 2, we simplified the badges and designed an explanatory function that would detail what activities earned points and how many points were earned. However, participants continued to express confusion about point accumulation and also felt that the point and reward system was not motivating. When asked, participants were enthusiastic about the idea of tangible rewards like gift cards. In our final version, we improved the points and rewards system by simplifying our point system to a single health warrior badge with 4 levels (bronze, silver, gold, and platinum). When a participant reaches each health warrior level, they earn gift cards that we send to their email.

Coaching

In round 1, some of the participants mentioned that although they knew that they were supposed to check in with their coach, it was not clear what they were supposed to talk about. For the subsequent iterations, we added topic suggestions to each coach check-in prompt. Some participants also felt the coach check-in should not be the responsibility of the participant but should rather be initiated by the coach. Others mentioned that the information in the coaching sessions should be more individualized. Participants also felt that when the coach was more aware of exactly what participants were doing with exercise, diet, and weight, they were more motivating. To help the coach provide more targeted coaching, we designed a coach app with a dashboard showing individual activities by participants who were using the *Fit After Baby* app, dietary intake, weight, steps, and points earned. This enabled the coach to follow each participant's activities in real time and provide targeted feedback and coaching.

Website

We designed a companion website to the *Fit After Baby* program. The website was designed to contain additional content that was not included in the app, including recipes. In addition, the website also provided a portal for participants to view personal data and rewards in greater depth and at a larger visual size. However, the website was rarely used, and most participants stated that the website was not useful or that they did not use the website because it was not easy to use on their mobile phones. We decided to remove the companion website for the final iteration of the program.

Additional Suggested Features

Several participants requested that mind-body and meditation techniques be included in the app. In the final iteration, we added a substantial amount of content including mind-body techniques and a yoga curriculum introducing two new poses per week. Many participants also expressed interest in sharing their information socially with the cohort. Further discussion revealed that they preferred the idea of sharing within the group simultaneously participating in the *Fit After Baby* program and going through similar challenges to sharing with friends or family members. There were mixed reviews on the idea of competition within the cohort, and on the idea of working together to earn points toward a common goal. Other participants requested more stories from women who had succeeded at achieving their postpartum goals. These features were not added to the final version but may be considered in the future.

Overall Impression of the Program

Overall, most participants responded that they were satisfied with the *Fit After Baby* program. Interestingly, we found that although the participants responded favorably to the content in general, they consistently requested more at each round. This feedback substantially modified the content of subsequent versions and was valuable in informing the final iteration of the program. Our improvements in the navigation of the app resulted in fewer complaints about navigational issues in subsequent iterations. All participants from the third iteration who responded in week 4 about whether they would recommend the *Fit After Baby* program to others said that they would. Participants in the group interviews responded favorably to the program overall and affirmed that it was helpful. They felt that the program

provided motivation and accountability to make changes in the postpartum period. Both participants in the second group interview expressed that they would have liked to continue to use the program. Overall, participants felt that the most useful components were the reminders and tracking components, and almost everyone seemed to enjoy using the Fitbit. In addition, most seemed to appreciate the information on healthy eating, quick and easy recipes, and exercises to do with baby.

Discussion

Principal Findings

We successfully employed an interactive, user-centered, participatory method of rapid, iterative design and evaluation to optimize the mHealth *Fit After Baby* program for postpartum women. Our multidisciplinary team of researchers and designers substantially improved the content and functionality of the app at each successive iteration by integrating user feedback. Qualitative data collected in the *Fit After Baby* development process provide valuable insight into the use of mobile technology–based weight loss apps in the target population of postpartum women at elevated cardiometabolic risk.

The most common and consistent theme through the iterative development process was a request for added content, and particularly added content tailored to postpartum mothers. Participants requested more postpartum-focused diet, exercise, and weight loss information, and also information on breastfeeding and postpartum mental health, including mind-body techniques for coping with stress. A consistent theme was the need to improve navigation because of difficulties in navigating to the appropriate screen. We responded to suggestions about navigation by creating a home screen with checkboxes so that daily and weekly tasks were more clear. The feedback from our participants suggested that the navigation was improved significantly by the final iteration. We were somewhat surprised to learn that women did not find the companion website to be useful, and furthermore, they found it difficult to use on a mobile phone. We eliminated the companion website for the final iteration. Diet and exercise tracking turned out to be challenging to most of our participants. Participants were clearly frustrated with the amount of time it took to track their diet and also disliked switching between apps, so in our final iteration, we limited tracking to particular days and decreased app switching by using the Fitbit app to track both diet and exercise. Many women said that they would prefer a wrist Fitbit, and we moved to this for the final version. We learned that the gamification component of points and badges was confusing and not very motivating. Many women said that they would prefer a tangible reward, so we added in the opportunity to earn gift cards for the final version. The participants also expressed a need for content that we did not originally include which led to the addition of a week of content focusing on mental health, anxiety, and stress in the final version.

Limitations

There are several limitations to this study. We developed and tested the *Fit After Baby* app on an iPhone Operating System platform. This affected the demographics and economic status of the recruited participants, which may thereby affect generalizability. We are currently developing an Android version of the *Fit After Baby* app and will be able to use both versions for future studies. The group sizes for the iterations were small, and therefore may not be adequately representative. Consistent with the demands of women during this very challenging postpartum period, all participants did not contribute to the online focus group every week. Although we attempted to conduct 2 in-person focus groups, we had several women who were unable to attend at the last minute and had to convert the 2 focus groups to 2-person group interviews. For this iterative design process, we did not include a control group, and therefore are unable to assess the significance of postpartum weight loss in this study.

Conclusions

Fit After Baby is a theory-based app using mHealth weight loss best practices and developed using the principles of rapid iterative design. To our knowledge, *Fit After Baby* is one of only a few weight loss apps designed specifically for postpartum women and the only one specifically focused on postpartum women at increased cardiometabolic risk. The design process involved a multidisciplinary team of researchers in collaboration with a technology team. We found the process of iterative development to substantially change the content and function of the Fit After Baby program which may be helpful for the development of similar programs in the future. Our first iteration compared with our final version has undergone substantial modification owing to this iterative process, and we are currently testing the final version of the *Fit After Baby* program in a pilot randomized trial with a primary goal of postpartum weight loss.

Acknowledgments

The authors thank the women who participated in the *Fit After Baby* pilot testing. We thank Dean Hovey, Susan Gilbert, Sue Arment, Digifit, and Comprendo for their outstanding collaboration and contribution to this work. The authors also thank the mHealth Impact Lab at the University of Colorado and our research assistant Danielle Cook. This study was supported by National Institutes of Health (NIH) BIRWCH K12HD057022, National Heart, Lung, and Blood Institute 1K23HL133604, the Center for Women's Health Research at the University of Colorado, the Nutrition and Obesity Research Center at the University of Colorado (NIDDK P30DK048520-23), and NIH/National Center for Advancing Translational Sciences Colorado CTSAUL1TR001082.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Representative quotes and iterative changes made throughout the three rounds of pilot testing and for the final version.
[DOCX File , 1261 KB - formative_v4i4e16151_app1.docx]

References

1. Mosca L, Benjamin EJ, Berra K, Bezanson JL, Dolor RJ, Lloyd-Jones DM, et al. Effectiveness-based guidelines for the prevention of cardiovascular disease in women--2011 update: a guideline from The American Heart Association. Circulation 2011 Mar 22;123(11):1243-1262 [FREE Full text] [doi: 10.1161/CIR.0b013e31820faaf8] [Medline: 21325087]
2. Gunderson EP, Chiang V, Pletcher MJ, Jacobs DR, Quesenberry CP, Sidney S, et al. History of gestational diabetes mellitus and future risk of atherosclerosis in mid-life: the Coronary Artery Risk Development in Young Adults study. J Am Heart Assoc 2014 Mar 12;3(2):e000490 [FREE Full text] [doi: 10.1161/JAHA.113.000490] [Medline: 24622610]
3. Tobias DK, Hu FB, Forman JP, Chavarro J, Zhang C. Increased risk of hypertension after gestational diabetes mellitus: findings from a large prospective cohort study. Diabetes Care 2011 Jul;34(7):1582-1584 [FREE Full text] [doi: 10.2337/dc11-0268] [Medline: 21593289]
4. Rich-Edwards JW. The predictive pregnancy: what complicated pregnancies tell us about mother's future cardiovascular risk. Circulation 2012 Mar 20;125(11):1336-1338. [doi: 10.1161/CIRCULATIONAHA.112.093872] [Medline: 22354939]
5. Fraser A, Nelson SM, Macdonald-Wallis C, Cherry L, Butler E, Sattar N, et al. Associations of pregnancy complications with calculated cardiovascular disease risk and cardiovascular risk factors in middle age: the Avon Longitudinal Study of Parents and Children. Circulation 2012 Mar 20;125(11):1367-1380 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.044784] [Medline: 22344039]
6. Villamor E, Cnattingius S. Interpregnancy weight change and risk of adverse pregnancy outcomes: a population-based study. Lancet 2006 Sep 30;368(9542):1164-1170. [doi: 10.1016/S0140-6736(06)69473-7] [Medline: 17011943]
7. Rodríguez-González GL, Castro-Rodríguez DC, Zambrano E. Pregnancy and Lactation: A Window of Opportunity to Improve Individual Health. Methods Mol Biol 2018;1735:115-144. [doi: 10.1007/978-1-4939-7614-0_7] [Medline: 29380310]
8. Bentley-Lewis R, Levkoff S, Stuebe A, Seely EW. Gestational diabetes mellitus: postpartum opportunities for the diagnosis and prevention of type 2 diabetes mellitus. Nat Clin Pract Endocrinol Metab 2008 Oct;4(10):552-558 [FREE Full text] [doi: 10.1038/ncpendmet0965] [Medline: 18779843]
9. Smith GN. Development of preeclampsia provides a window of opportunity for early cardiovascular risk screening and intervention. Expert Rev Obstet and Gynecol 2009;4(4):355-357. [doi: 10.1586/eog.09.16]
10. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 7;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]
11. Eckel RH, Jakicic JM, Ard JD, de Jesus JM, Miller NH, Hubbard VS, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 AHA/ACC guideline on lifestyle management to reduce cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014 Jun 24;129(25 Suppl 2):S76-S99. [doi: 10.1161/01.cir.0000437740.48606.d1] [Medline: 24222015]
12. Bao W, Tobias DK, Bowers K, Chavarro J, Vaag A, Grunnet LG, et al. Physical activity and sedentary behaviors associated with risk of progression from gestational diabetes mellitus to type 2 diabetes mellitus: a prospective cohort study. JAMA Intern Med 2014 Jul;174(7):1047-1055 [FREE Full text] [doi: 10.1001/jamainternmed.2014.1795] [Medline: 24841449]
13. Hedderson M, Ferrara A. A call to increase physical activity among women of reproductive age: is it possible? JAMA Intern Med 2014 Jul;174(7):1056-1057. [doi: 10.1001/jamainternmed.2014.709] [Medline: 24841171]
14. Huang Z, Willett WC, Manson JE, Rosner B, Stampfer MJ, Speizer FE, et al. Body weight, weight change, and risk for hypertension in women. Ann Intern Med 1998 Jan 15;128(2):81-88. [doi: 10.7326/0003-4819-128-2-199801150-00001] [Medline: 9441586]
15. Catalano PM, Ehrenberg HM. The short- and long-term implications of maternal obesity on the mother and her offspring. Br J Obstet Gynaecol 2006 Oct;113(10):1126-1133 [FREE Full text] [doi: 10.1111/j.1471-0528.2006.00989.x] [Medline: 16827826]
16. Glazer NL, Hendrickson AF, Schellenbaum GD, Mueller BA. Weight change and the risk of gestational diabetes in obese women. Epidemiology 2004 Nov;15(6):733-737. [doi: 10.1097/01.ede.0000142151.16880.03] [Medline: 15475723]
17. Hoedjes M, Berks D, Vogel I, Franx A, Duvekot JJ, Oenema A, et al. Motivators and barriers to a healthy postpartum lifestyle in women at increased cardiovascular and metabolic risk: a focus-group study. Hypertens Pregnancy 2012;31(1):147-155. [doi: 10.3109/10641955.2010.544803] [Medline: 21250888]
18. Østbye T, Krause KM, Lovelady CA, Morey MC, Bastian LA, Peterson BL, et al. Active Mothers Postpartum: a randomized controlled weight-loss intervention trial. Am J Prev Med 2009 Sep;37(3):173-180 [FREE Full text] [doi: 10.1016/j.amepre.2009.05.016] [Medline: 19595557]
19. Sarwer DB, Allison KC, Gibbons LM, Markowitz JT, Nelson DB. Pregnancy and obesity: a review and agenda for future research. J Womens Health (Larchmt) 2006;15(6):720-733. [doi: 10.1089/jwh.2006.15.720] [Medline: 16910904]

20. Nicklas JM, Barbour LA. Optimizing weight for maternal and infant health - tenable, or too late? Expert Rev Endocrinol Metab 2015;10(2):227-242 [FREE Full text] [doi: 10.1586/17446651.2014.991102] [Medline: 26442123]
21. Jones EJ, Peercy M, Woods JC, Parker SP, Jackson T, Mata SA, et al. Identifying postpartum intervention approaches to reduce cardiometabolic risk among American Indian women with prior gestational diabetes, Oklahoma, 2012-2013. Prev Chronic Dis 2015 Apr 02;12:E45 [FREE Full text] [doi: 10.5888/pcd12.140566] [Medline: 25837258]
22. Dasgupta K, da Costa D, Pillay S, de Civita M, Gougeon R, Leong A, et al. Strategies to optimize participation in diabetes prevention programs following gestational diabetes: a focus group study. PLoS One 2013;8(7):e67878 [FREE Full text] [doi: 10.1371/journal.pone.0067878] [Medline: 23861824]
23. Hearn L, Miller M, Lester L. Reaching perinatal women online: the Healthy You, Healthy Baby website and app. J Obes 2014;2014:573928 [FREE Full text] [doi: 10.1155/2014/573928] [Medline: 24872891]
24. Smith A. Pew Research Center. Washington, DC: Pew Research Center; 2015 Apr 1. US Smartphone Use in 2015 URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2020-01-13]
25. Breton ER, Fuemmeler BF, Abroms LC. Weight loss-there is an app for that! But does it adhere to evidence-informed practices? Transl Behav Med 2011 Dec;1(4):523-529 [FREE Full text] [doi: 10.1007/s13142-011-0076-5] [Medline: 24073074]
26. Khaylis A, Yiaslas T, Bergstrom J, Gore-Felton C. A review of efficacious technology-based weight-loss interventions: five key components. Telemed J E Health 2010 Nov;16(9):931-938 [FREE Full text] [doi: 10.1089/tmj.2010.0065] [Medline: 21091286]
27. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]
28. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
29. Conroy DE, Yang C, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
30. Burke LE, Ma J, Azar KM, Bennett GG, Peterson ED, Zheng Y, American Heart Association Publications Committee of the Council on Epidemiology and Prevention, Behavior Change Committee of the Council on Cardiometabolic Health, Council on Cardiovascular and Stroke Nursing, Council on Functional Genomics and Translational Biology, Council on Quality of Care and Outcomes Research,Stroke Council. Current science on consumer use of mobile health for cardiovascular disease prevention: a scientific statement from The American Heart Association. Circulation 2015 Sep 22;132(12):1157-1213. [doi: 10.1161/CIR.0000000000000232] [Medline: 26271892]
31. Mummah SA, King AC, Gardner CD, Sutton S. Iterative development of Vegethon: a theory-based mobile app intervention to increase vegetable consumption. Int J Behav Nutr Phys Act 2016 Aug 8;13:90 [FREE Full text] [doi: 10.1186/s12966-016-0400-z] [Medline: 27501724]
32. Ybarra ML, Prescott TL, Philips GL, Bull SS, Parsons JT, Mustanski B. Iteratively developing an mHealth HIV prevention program for sexual minority adolescent men. AIDS Behav 2016 Jun;20(6):1157-1172 [FREE Full text] [doi: 10.1007/s10461-015-1146-3] [Medline: 26238038]
33. Lin PH, Intille S, Bennett G, Bosworth HB, Corsino L, Voils C, et al. Adaptive intervention design in mobile health: intervention design and development in the Cell Phone Intervention for You trial. Clin Trials 2015 Dec;12(6):634-645 [FREE Full text] [doi: 10.1177/1740774515597222] [Medline: 26229119]
34. DiFilippo KN, Huang W, Chapman-Novakofski KM. A New Tool for Nutrition App Quality Evaluation (AQEL): development, validation, and reliability testing. JMIR Mhealth Uhealth 2017 Oct 27;5(10):e163 [FREE Full text] [doi: 10.2196/mhealth.7441] [Medline: 29079554]
35. Bull S, Ezeanochie N. From Foucault to Freire through Facebook: toward an integrated theory of mHealth. Health Educ Behav 2016 Aug;43(4):399-411. [doi: 10.1177/1090198115605310] [Medline: 26384499]
36. Webb TL, Sniehotta FF, Michie S. Using theories of behaviour change to inform interventions for addictive behaviours. Addiction 2010 Nov;105(11):1879-1892. [doi: 10.1111/j.1360-0443.2010.03028.x] [Medline: 20670346]
37. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckmann J, editors. Action-Control: From Cognition to Behavior. Berlin, Heidelberg: Springer; 1985:11-39.
38. Lu MC, Kotelchuck M, Culhane JF, Hobel CJ, Klerman LV, Thorp JM. Preconception care between pregnancies: the content of internatal care. Matern Child Health J 2006 Sep;10(5 Suppl):S107-S122 [FREE Full text] [doi: 10.1007/s10995-006-0118-7] [Medline: 16817001]
39. Beckles GL, Thompson-Reid PE. CDC stacks.: Centers for Disease Control and Prevention; 2001. Diabetes & Women's Health Across the Life Stages: A Public Health Perspective URL: https://stacks.cdc.gov/view/cdc/6426 [accessed 2020-01-13]
40. Bandura A. Self-Efficacy: The Exercise of Control. New York: Freeman; 1997.
41. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 5;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]

42. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008 Aug 25;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
43. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
44. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17(Suppl 1):11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
45. Zhang J. Human-centered computing in health information systems. Part 1: analysis and design. J Biomed Inform 2005 Feb;38(1):1-3 [FREE Full text] [doi: 10.1016/j.jbi.2004.12.002] [Medline: 15694880]
46. Hingle M, Patrick H. There are thousands of apps for that: navigating mobile technology for nutrition education and behavior. J Nutr Educ Behav 2016 Mar;48(3):213-8.e1. [doi: 10.1016/j.jneb.2015.12.009] [Medline: 26965099]
47. Waterlander W, Whittaker R, McRobbie H, Dorey E, Ball K, Maddison R, et al. Development of an evidence-based mHealth weight management program using a formative research process. JMIR Mhealth Uhealth 2014 Jul 11;2(3):e18 [FREE Full text] [doi: 10.2196/mhealth.2850] [Medline: 25098337]
48. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
49. Stephens J, Allen JK, Himmelfarb CR. 'Smart' coaching to promote physical activity, diet change, and cardiovascular health. J Cardiovasc Nurs 2011;26(4):282-284 [FREE Full text] [doi: 10.1097/JCN.0b013e31821ddd76] [Medline: 21666418]
50. Hickey DT. The EvoLLLution. 2012 Nov 20. Recognizing, Supporting, and Attracting Adult Learners with Digital Badges URL: https://evolllution.com/opinions/recognizing-supporting-and-attracting-adult-learners-with-digital-badges/ [accessed 2020-01-13]
51. Koivisto J, Hamari J. Demographic differences in perceived benefits from gamification. Comput Human Behav 2014 Jun;35:179-188. [doi: 10.1016/j.chb.2014.03.007]
52. Miller AS, Cafazzo JA, Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2016 Jun;22(2):184-193. [doi: 10.1177/1460458214537511] [Medline: 24986104]
53. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. JMIR Serious Games 2014 Aug 4;2(2):e9 [FREE Full text] [doi: 10.2196/games.3413] [Medline: 25654660]
54. Gary K, Enquobahrie A, Ibanez L, Cheng P, Yaniv Z, Cleary K, et al. Agile methods for open source safety-critical software. Softw Pract Exp 2011 Aug 1;41(9):945-962 [FREE Full text] [doi: 10.1002/spe.1075] [Medline: 21799545]
55. Larman C, Basili VR. Iterative and incremental developments. A brief history. Comput 2003 Jun;36(6):47-56. [doi: 10.1109/mc.2003.1204375]
56. 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]
57. Stemler S. An overview of content analysis. Pract Assess Res Eval 2000;7(17) [FREE Full text]
58. Burke LE, Warziski M, Starrett T, Choo J, Music E, Sereika S, et al. Self-monitoring dietary intake: current and future practices. J Ren Nutr 2005 Jul;15(3):281-290. [doi: 10.1016/j.jrn.2005.04.002] [Medline: 16007557]

Abbreviations

CVD: cardiovascular disease
DPP: Diabetes Prevention Program
GDM: gestational diabetes mellitus
mHealth: mobile health
NIH: National Institutes of Health
SGA: small-for-gestational age

Edited by G Eysenbach; submitted 05.09.19; peer-reviewed by K Vesco, W Zhang; comments to author 09.11.19; revised version received 18.11.19; accepted 15.12.19; published 09.04.20.

Please cite as:
Nicklas JM, Leiferman JA, Lockhart S, Daly KM, Bull SS, Barbour LA
Development and Modification of a Mobile Health Program to Promote Postpartum Weight Loss in Women at Elevated Risk for Cardiometabolic Disease: Single-Arm Pilot Study
JMIR Form Res 2020;4(4):e16151
URL: https://formative.jmir.org/2020/4/e16151
doi:10.2196/16151
PMID:32271149

Original Paper

Interventions to Increase the Reachability of Migrants in Germany With Health Interview Surveys: Mixed-Mode Feasibility Study

Marie-Luise Zeisler[1], MA; Leman Bilgic[2], MA; Maria Schumann[1], MSc; Annelene Wengler[1], Dr Rer Pol; Johannes Lemcke[1], Dipl Soz; Antje Gößwald[1], MPH, Dr med; Thomas Lampert[1], Dr; Claudia Santos-Hövener[1], MPH, Dr; Patrick Schmich[1], Dipl Soz

[1]Robert Koch Institute, Berlin, Germany

[2]Charité Berlin, Berlin, Germany

Corresponding Author:
Marie-Luise Zeisler, MA
Robert Koch Institute
Berlin
Germany
Phone: 49 30 18754 3694
Email: ZeislerM@rki.de

Abstract

Background: Germany is a popular destination for immigrants, and migration has increased in recent years. It is therefore important to collect reliable data on migrants' health. The Robert Koch Institute, Berlin, Germany, has launched the Improving Health Monitoring in Migrant Populations (IMIRA) project to sustainably integrate migrant populations into health monitoring in Germany.

Objective: One of IMIRA's objectives is to implement a feasibility study (the IMIRA survey) that focuses on testing various interventions to increase the reachability of migrants with health interview surveys. Possible causes of nonresponse should be identified so as to increase participation in future surveys.

Methods: The survey target populations were Turkish, Polish, Romanian, Syrian, and Croatian migrants, who represent the biggest migrant groups living in Germany. We used probability sampling, using data from the registration offices in 2 states (Berlin and Brandenburg); we randomly selected 9068 persons by nationality in 7 sample points. We applied age (3 categories: 18-44, 45-64, and ≥65 years) and sex strata. Modes and methods used to test their usability were culturally sensitive materials, online questionnaires, telephone interviews, personal contact, and personal interviews, using multilingual materials and interviewers. To evaluate the effectiveness of the interventions, we used an intervention group (group A) and a control group (group B). There were also focus groups with the interviewers to get more information about the participants' motivation. We used the European Health Interview Survey, with additional instruments on religious affiliation, experience of discrimination, and subjective social status. We evaluated results according to their final contact result (disposition code).

Results: We collected data from January to May 2018 in Berlin and Brandenburg, Germany. The survey had an overall response rate of 15.88% (1190/7494). However, final disposition codes varied greatly with regard to citizenship. In addition to the quantitative results, interviewers reported in the focus groups a "feeling of connectedness" to the participants due to the multilingual interventions. The interviewers were particularly positive about the home visits, because "if you are standing at the front door, you will be let in for sure."

Conclusions: The IMIRA survey appraised the usability of mixed-mode or mixed-method approaches among migrant groups with a probability sample in 2 German states. When conducting the survey, we were confronted with issues regarding the translation of the questionnaire, as well as the validity of some instruments in the survey languages. A major result was that personal face-to-face contact was the most effective intervention to recruit our participants. We will implement the findings in the upcoming health monitoring study at the Robert Koch Institute.

(JMIR Form Res 2020;4(4):e14747) doi:10.2196/14747

KEYWORDS

transients and migrants; surveys and questionnaires; cross-sectional studies; feasibility studies; multilingualism and health monitoring

Introduction

Background

According to the Microcensus of German households, a person has a migrant background if he or she, or at least one parent, has no German citizenship at birth [1]. In Germany, this applied to 19.3 million people in 2017, corresponding to 23.6% of the total population. Of these, 9.8 million had German citizenship, and 9.4 million had another citizenship. Nearly two-thirds of persons with a migrant background (PMB) had migrated themselves (first generation), and one-third were born in Germany (subsequent generations) [1]. Migration has increased since the Second World War, for several reasons. Whereas the recruitment of migrant "guest workers" in the 1950s, and the resulting family reunifications in the 1970s, and European Union enlargement in the 2000s were primarily work oriented, numerous conflicts and wars have led to an increase of refugees coming to Germany, especially since the 2000s [2-4]. The group of PMB in Germany can be described as very heterogeneous, due to differences in the phases of influx, migration background, or specific life circumstances. PMB are underrepresented in nationwide health surveys, such as the health monitoring surveys of the Robert Koch Institute (RKI) in Berlin, Germany [5]. Lower response rates among PMB in population-based surveys have been described in many Western countries [6,7]. The reasons for this can be summarized inter alia as (1) inadequate sampling approaches and hence failing to include PMB in surveys; or (2) barriers to the recruitment of PMB due, for example, to language or cultural issues [8]. It is therefore necessary to evaluate various sampling strategies to include PMB adequately in surveys and to design appropriate recruitment efforts that can minimize nonresponse.

There is an increasing need for reliable data on migrants' health in order to give a more representative picture of the population. As the national public health institute in Germany, the RKI has the task of extending health monitoring by means of the preferably representative integration of PMB into its health and examination studies. In this context, the Improving Health Monitoring in Migrant Populations (IMIRA) project was initiated, of which the IMIRA survey described here is a part.

Sampling Strategies

Depending on the objectives of a survey and the associated generalizability of the results to the population, different sampling strategies can be applied. The health monitoring of the RKI aims to cover the German population as a whole; thus, only probability-based sampling approaches can be applied as a specific sampling strategy. In probability samples the selected population is sampled randomly; that means, from a statistical standpoint, the probability of inclusion in the probability sample is predictable and thus results can be generalized. This is not the case for sampling strategies where no inclusion probability is known, for example, snowball or convenience samples, in which participants are recruited by other participants [9,10]. Although nonprobability sampling strategies are considered effective in recruiting hard-to-reach populations [8], the results can hardly be generalized. Probability sampling approaches are frequently based on registers, for example, population registers.

In the RKI's monitoring studies, a 2-stage probability sampling design is applied [11-15]. In the first stage, primary sampling units are selected that are representative of German municipalities. In the second stage, target persons are randomly drawn, according to proportional age and sex strata, through the residents registries (*Einwohnermeldeämter*) within these municipalities. Since Germany has a federal administrative system, in contrast to other European countries, no central residents registry exists [4]. Every municipality therefore needs to be contacted separately.

The residents registry captures various characteristics with which to identify PMB, such as citizenship and place of birth [16]. Citizenship is commonly used for the oversampling of persons with another citizenship to compensate for the assumed lower response rates of PMB. This has been applied in the RKI studies Study of the German Health Interview and Examination Survey for Adults [15,17] and Study of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) [12-14]. Use of the citizenship characteristic excludes PMB who are naturalized or belong to subsequent generations without non-German citizenship. Place of birth would also be a way to identify PMB with German citizenship. Since it is provided as a string variable in the residents registries, it is prone to error and therefore is not suitable for a big sample [16].

Another important component mentioned in the literature is onomastic procedures, which try to identify naturalized PMB or persons of subsequent migrant generations. These procedures use an algorithm to allocate the origin of a person according to their name [18]. The success of the method might be biased; the name algorithm can lead to more wrong allocations in some PMB groups than in others [19]. Onomastic procedures were not used in the process of sampling at RKI, but to allocate bilingual study information [5,20].

The aforementioned sampling strategies can be considered migrant-sensitive sampling approaches for a better representation of PMB in health surveys. The subsequent steps include measures to increase the response rate, which we describe below.

Measures to Increase the Response Rate of Persons With a Migrant Background

The heterogeneity of PMB requires a highly differentiated approach for increasing their participation rates in surveys [21]. According to findings in contemporary research, and alongside the challenges and limitations in sampling described above, PMB also have a higher participation threshold than persons with no migrant background, which is among other things presumably due to language and cultural barriers [8,21]. In that respect, lacking the ability to understand the survey language, but moreover illiteracy, might prevent survey participation. A lack of interest in survey participation is also reported as a reason for lower participation rates [7,22]. Researchers have found the following possible reasons for the lower response rate of PMB in surveys: a lack of trust in research, a fear that the reported information might be misused, or cultural beliefs preventing participation in surveys that include intimate or sensitive topics [8]. Specific measures should be taken to increase response rates.

The most obvious way to overcome barriers related to language and literacy skills is to offer multilingual survey materials in plain but culturally sensitive language [23,24]. It is also important to consider the mode of questionnaire administration. Self-administered interview modes, such as online or paper-and-pencil questionnaires, can be differentiated from interviewer-administered modes, such as telephone interviews or face-to-face interviews. A self-administered mode such as an online questionnaire can be accessed easily with a mobile phone and might have a positive impact on the participation of younger and more Web-savvy persons [25]. A paper-and-pencil questionnaire can be filled out without further technical devices and may be associated with less effort for the participants. In general, self-administered interview modes are known to be less subject to effects of social desirability or interviewer influence [26,27].

Nonetheless, telephone or face-to-face interviews can facilitate the participation of persons with literacy problems or other issues in understanding surveys [8,26,28]. Current research emphasizes the importance of personal contact with PMB, such as home visits, in order to increase response rates [8,21]. Personal contact makes it possible to overcome barriers based on mistrust or anxiety, and provides more detailed information about the survey's objectives [8]. A similar cultural background between the interviewer and participant can increase the response rate [21]. Personal contact leads to a more heterogeneous sample composition according to characteristics of social status, health status, and educational level. Especially vulnerable persons, who would or could not participate in the survey without help or further interventions, can be reached at home. In KiGGS, home visits by specifically trained survey staff proved an effective measure for increasing willingness to participate in the survey and doubled the response rate [20].

Self-administered interview modes were preferred in the most recent RKI monitoring surveys because the sampling strategy involved residents registries, from which only the address of the person could be drawn. This means that, for example, telephone numbers were not available for telephone interviews.

In addition to interventions that directly target participants, some indirect measures are also discussed in the literature. One possibility is to involve the gatekeepers of migrant communities, namely persons who can influence survey promotion and attitudes toward survey participation within the communities, to enhance the participation rate [8]. Furthermore, the involvement of gatekeepers should be considered in order to develop specific public relations when addressing the concerns of PMB [24,29].

Objectives

Our main objective with the IMIRA project was to identify methodological procedures to further identify methods to better reach PMB in health surveys, and to thus to increase the response rate in future health surveys. The IMIRA survey is a feasibility survey. One focus of the IMIRA survey is to test various interventions that take into account possible cultural and language barriers through the use of a mixed-mode design and multilingual administration. The objectives of the feasibility study (the IMIRA survey) can be summarized as follows: (1) to develop and test the feasibility of an adapted survey design, which is the basis of a subsequent nationwide survey in Germany; (2) to improve the response rate of PMB who have been poorly reached so far and who belong to the biggest groups in Germany; and (3) to identify the causes of nonparticipation in order to increase motivation to participate in health surveys of PMB.

Methods

Selection of the Target Population

We applied 2 main criteria for the selection of the target population for the IMIRA survey: the size of the PMB group in the German population, which Table 1 [30] shows; and the experience in previous RKI surveys with the respective PMB regarding response and reachability [5,20]. On the basis of these 2 criteria, we included persons with Turkish, Syrian, Romanian, Croatian, and Polish citizenship in the sample for the IMIRA survey. We selected them according to these criteria regardless of whether they had additional German or other citizenship. We thus excluded from the sample naturalized persons, meaning persons who have acquired German citizenship in replacement of or in addition to another citizenship and persons of subsequent migrant generations with only German citizenship.

We received ethics approval for the IMIRA survey on October 30, 2017 from the ethics committee of the Medical University of Charité, Berlin, Germany (EA1/210/17). The study protocol was approved by the Federal Commissioner for Data Protection and Freedom of Information January 3, 2018 (13-401/008#0085).

Table 1. Foreign population in Germany according to citizenship in 2017 [30].

Citizenship	Count, n
Afghan	251,640
Bulgarian	310,415
Croatian	367,900
Greek	362,245
Italian	643,065
Polish	866,855
Romanian	622,780
Russian	249,205
Syrian	698,950
Turkish	1,483,515

Selection of Sample Points and Sample Size

Congruent with the health monitoring studies at the RKI, we used a register-based random sample in the IMIRA survey. The sampling was a 2-staged procedure: in the first stage, we applied a criteria-based selection of sample points in Berlin and Brandenburg; in the second stage, we randomly selected people according to their citizenship.

Step 1: Criteria-Based Selection of Sample Points

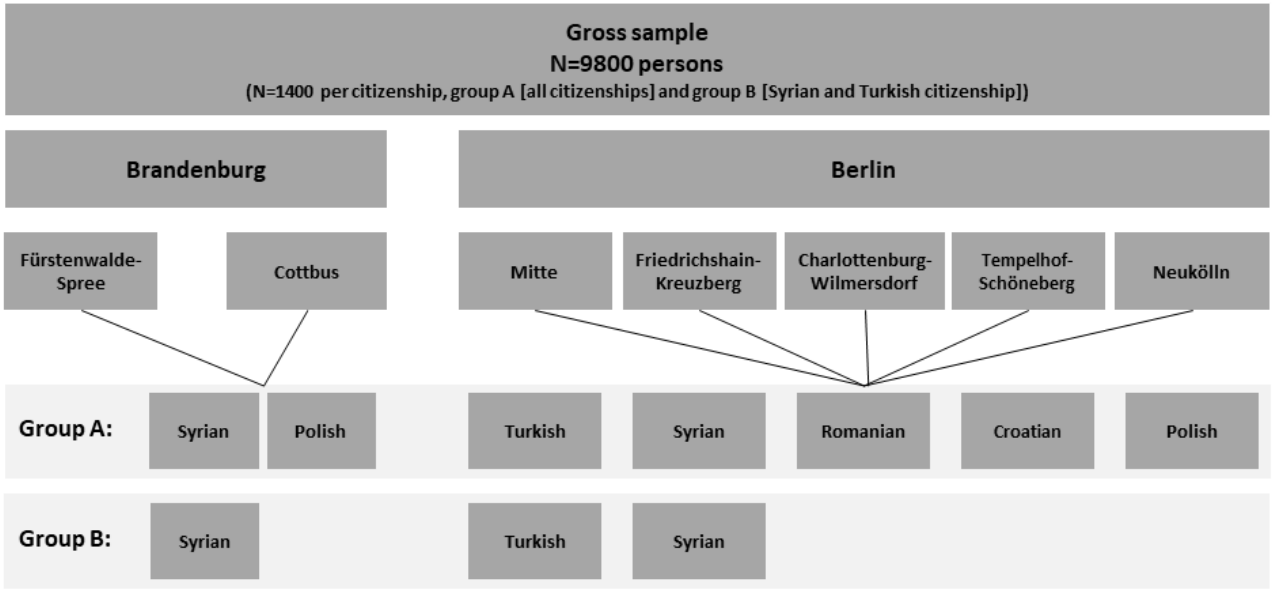
We decided to focus on 2 German federal states in the IMIRA survey, with Berlin representing urban regions and Brandenburg representing rural regions. We selected the sample points taking into account the highest proportions of persons without German citizenship in the respective federal states, using data from the Statistical Office in Berlin Brandenburg at the municipal level from 2015 [31]. Since the proportions of persons without German citizenship in Berlin and Brandenburg vary greatly (in 2016 Berlin was home to 16.7% non-Germans, and Brandenburg was home to 4.0% [32]), we focused the selection of our sample points on Berlin. We selected 5 urban sample points (Mitte, Neukölln, Charlottenburg-Wilmersdorf, Friedrichshain-Kreuzberg, and Tempelhof-Schöneberg districts) in Berlin and 2 relatively rural sample points (Cottbus and Fürstenwalde/Spree) in Brandenburg, which met the above criteria.

Step: 2 Selection of Persons in the Residents Registry

In the second stage of the sampling, we selected people through the residents registration offices of the chosen sample points. The sampling procedure differed in Berlin and Brandenburg, since not all 5 PMB populations were sufficiently represented in all sample points. In Berlin the target population consisted of people with Turkish, Syrian, Romanian, Croatian, and Polish citizenship, whereas only a few people with Turkish, Romanian, and Croatian citizenship lived in the 2 sample points in Brandenburg. We discovered this through a statistical enquiry in the residents registration offices in Cottbus and Fürstenwalde/Spree on May 24, 2017. Accordingly, we selected only people with Polish or Syrian citizenship in the 2 sample points in Brandenburg. We selected the whole sample with a proportional distribution strata of sexes and 3 age groups (18-44 years, 45-64 years, and ≥65 years): we selected exactly the same number of people in each sex and age category in each citizenship due to the experimental nature of the feasibility study. In this sample we tried to include enough participants in each sex and age category for a better interpretation of our results and the comparison of interventions.

The aim of this survey was to evaluate the effectiveness of specific interventions in increasing response rates. We thus established 2 groups (A and B) for participants with Turkish and Syrian citizenship (the process is detailed in the Data Collection section below). To achieve sufficiently high case numbers in groups A and B of the survey design, we doubled the number of Turkish and Syrian people sampled. We thus selected a total of 1400 people with Romanian, Croatian, and Polish citizenship, and a total of 2800 people of Turkish and Syrian citizenship, resulting in a gross sample of N=9800 for the IMIRA survey (see Figure 1).

Figure 1. Sample design.



Questionnaire

The survey instrument was based on the questionnaire of the European Health Interview Survey, which was used and validated in 30 countries [33,34]. Additional instruments included a scale to measure religious affiliation [35,36], an instrument to measure experiences of discrimination [37], and an instrument to measure subjective social status before and after migration [38]. The additional instruments were partly validated in other surveys.

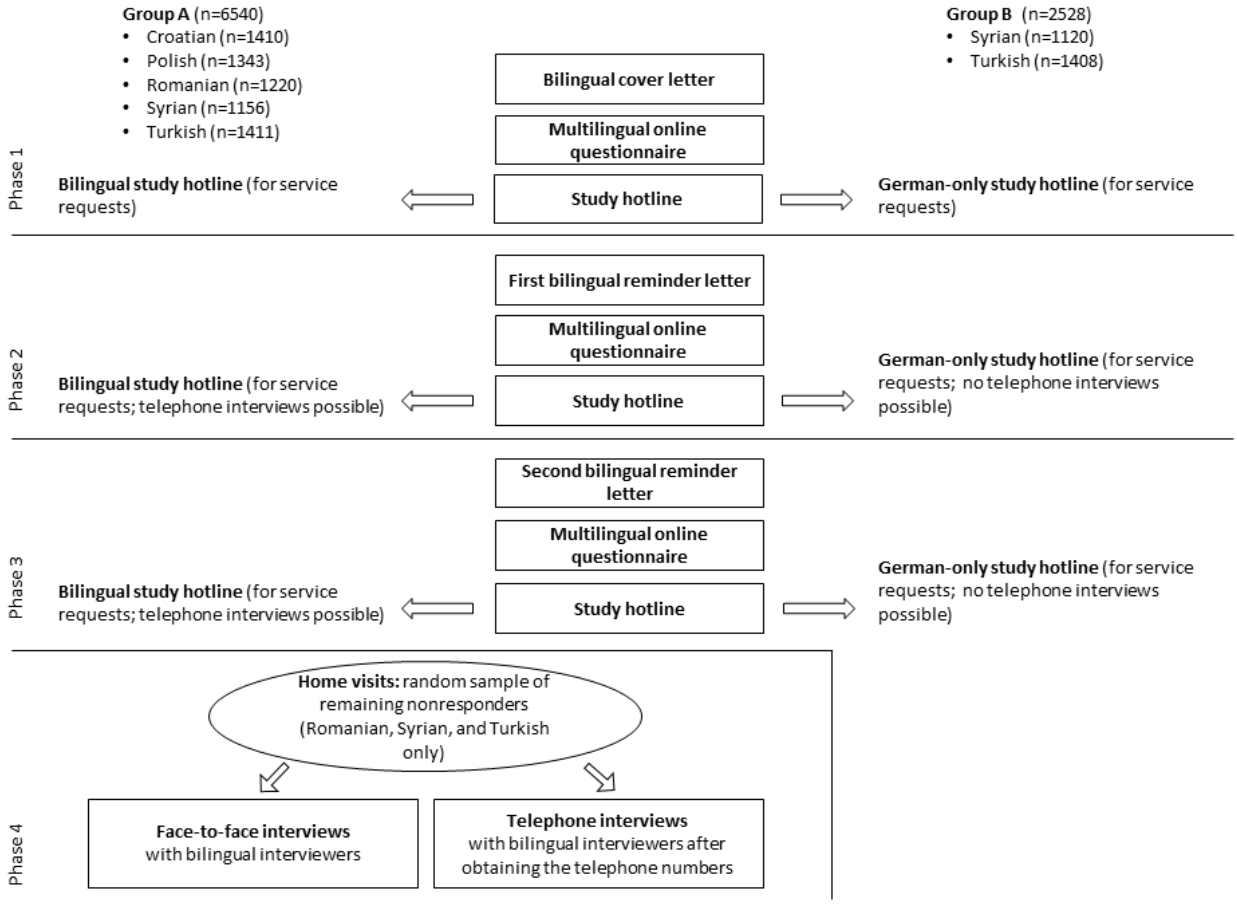
Data Collection

Survey Design

To evaluate the effectiveness of the interventions, we established 2 groups for Turkish and Syrian participants in an experimental design. Whereas group A (intervention group) received further options to participate in the survey, such as face- to-face interviews with bilingual interviewers, group B (reference group) did not (see Figure 2). All participants received a €10 (about US $11) shopping voucher as a conditional incentive. The interventions in group A were presented in a sequential mixed-mode design and are described below.

Qualitative methods such as focus groups with representatives of the migrant populations can be used to identify and evaluate specific measures to attract participants for surveys [24]. The information gained can, for example, be incorporated into the survey materials [22] or help to develop a diversity-sensitive survey design: in addition to the use of multilingual study materials, questionnaires, and interviewers, all materials should use a diversity-sensitive tailored language [21]. To ensure the cultural sensitivity of IMIRA's cover letters and study materials, we conducted a focus group with representatives from the target populations. The content and design of the cover letter and facilitating factors for survey participation, such as the amount and choice of the incentive, were discussed. Based on the results of the focus group, the cover letters were designed by a professional graphic design company, using pictograms instead of photographs.

Figure 2. Survey design.



Phase 1

After we selected the sample in December 2017, we sent the survey information materials by mail to participants in January 2018. These consisted of a cover letter, an information sheet with detailed information on the survey, and a document explaining the data protection measures to the participants. The survey information materials for both groups were in German and in the participant's language based on their citizenship. All information materials and the questionnaires were translated by a professional translation agency. For the translation, certified translators with native-language level in the required languages were selected by the company. The translators also had experience in translation in the field of health and scientific surveys. The translations were subjected to quality control by the translation agency and subsequent proofreading by a second translator. We collected data for the IMIRA survey from January to May 2018.

The cover letter described the option to participate online using a multilingual computer-assisted Web interview and promoted a toll-free study hotline for queries regarding the survey. The online questionnaire was available in Turkish, Arabic, Romanian, Croatian, Polish, and German in both groups (A and B) of the survey design. To access the computer-assisted Web interview, a personal pincode was necessary. Changing languages during the completion process was possible due to a drop-down language selection menu. This allowed participants to easily switch between the provided languages. The online questionnaire was programmed with Voxco Online software version 5.5.1.205 (Groupe Voxco Inc).

The participants could contact the toll-free study hotline for further information about the survey, to resolve issues with the questionnaire, or to refuse participation. In group A multilingual staff could be contacted through the study hotline. In group B the study hotline was available in German only. Reasons for calling the study hotline were recorded in a process questionnaire, which was also a measure for process standardization and quality assurance. It was programmed with the call-assisted telephone interview software Voxco Command Center version 1.10.5. The process questionnaire recorded queries regarding contact frequency, contact person, result of contact, reasons for nonresponse, and information about the language used during contact. The interviewers also had the option to add further notes about the contact. If the study hotline was contacted to refuse participation, the interviewers were trained to try to convince callers to participate or at least document the reasons for refusal. The interview training took place before the actual start of the survey (field start) in a 2-day intensive training session. On the first day interviewers were informed about the survey objectives and the survey design in detail. In the training participants could address questions to the interviewers, and possible issues and their resolutions were discussed. Interviewers were familiarized with the health interview questionnaire and the process questionnaire that was

used to record the hotline calls. On the second day interviewers had the opportunity to practice the use of the software and to click through the questionnaires by themselves. Afterward they practiced possible situations on the study hotline as a role play. During the whole training interviewers could address questions to the project researchers. To remember the training interviewers received a handbook with information about the survey, as well as the most frequently asked questions and problems that might occur. For quick help there was also an overview sheet on the survey with the most important information (Who is conducting the survey and why? What are the survey objectives? Who was invited and how? How can I participate?) at each workstation. Throughout the survey fieldwork the interviewers on the study hotline were supported by experienced supervisors to ensure the quality of data and to provide assistance in difficult situations. The supervisors underwent the same training as the interviewers. In addition, all supervisors had extensive prior experience with scientific surveys and the software used, and were thus able to provide assistance with technical questions. Supervision was also supplemented by project staff who were able to answer questions and problems regarding content.

Phase 2

After 2 weeks a reminder was sent to people who had not participated after the first invitation. The reminder promoted the multilingual computer-assisted Web interview again. Participants in group A were given the option to take part in the survey through a call-assisted telephone interview with bilingual interviewers.

Phase 3

After 3 weeks a second reminder was sent to participants to promote the survey again, and home visits were offered to group A. Group B participants received a second reminder to complete the online questionnaire. We conducted a focus group with the telephone hotline interviewers after the end of phase 3 to learn about their experiences. The interviewers discussed their personal experience with the IMIRA survey during their work on the study hotline. The focus group was moderated by 1 of the IMIRA survey's researchers (LB), using an interview guideline. The interview guideline included the following topics: (1) evaluation of the contact design with a focus on reasons for calling the study hotline, mentioned issues or problems, reasons for refusal, and whether these objections could be resolved, and the response to the offer of bilingual interviews; and (2) the interviewers' opportunity to express their own opinion with a focus on the bilingual materials, study hotline, and translated questionnaire. A log was kept during the focus group by 2 researchers (MLZ) to enable later analysis. We evaluated the log using qualitative content analysis [39]. We formulated categories to analyze the material deductively taking into account the interview guidelines.

Phase 4

For the majority of the people sampled (ie, people with Polish or Croatian citizenship or people living in the 2 sample points in Brandenburg), the IMIRA survey ended with the last reminder letter in phase 3. Home visits took place only in a random subsample. The inclusion criteria for the random subsample were as follows: Romanian, Syrian, or Turkish citizenship; allocation to group A; and no response to the further invitation steps. In group B no further interventions were initiated. In this last survey phase we carried out 2 interventions, aiming at a comparison of their effectiveness: (1) home visits for face-to-face interviews using bilingual interviewers, and (2) home visits to obtain the participant's telephone number to conduct a subsequent call-assisted telephone interview with bilingual interviewers.

The face-to-face interviews were carried out using a tablet with internet access and the online questionnaire. To evaluate the 2 approaches, participants were randomly allocated to the face-to-face interview group or the phone number group. Both interventions were aimed at increasing the response rate further and gaining more information about the target population, for example, reasons for nonresponse, German language skills, or validation of the addresses. An address was not replaced if it was wrong or unavailable. These data were documented in a process questionnaire similar to that used for the study hotline.

Whenever possible the interviewers were accompanied by an experienced supervisor to ensure data quality and reduce uncertainties in handling the tablet. Based on the experience of the first interviewers' focus group, a second one was carried out at the end of the home visit phase. The interviewers' experiences enriched the quantitative data gathered with the process questionnaire. The focus group was moderated by an IMIRA researcher (LB) using an interview guideline. The interview guideline was structured in 3 main parts: (1) experiences with the home visits, including situations at the door, experiences during face-to-face interviews or the retrieval of telephone numbers, and reactions to bilingual contact; (2) technical and organizational framework regarding address quality, tablet handling, or operating times; and (3) the interviewer's personal opinion about the effectiveness of the home visits. A log was kept by 2 researchers (MLZ) for subsequent qualitative content analysis. This was done in a similar way to that in the first interviewers' focus group.

Phase 4 can be considered a benchmark test for a specific subsample of nonresponders. No evaluation in comparison with group B was possible due to the study design.

Results

Response to the Survey

Data collection was completed in May 2018. The response rate was 15.88% (1190/7494) over all target populations, such that 1190 questionnaires were completed by our participants. We published the first results on response rates, the effectiveness of the interventions, and sample composition in 2019 [40].

Gross Sample

We selected 9800 people, stratified by age and sex, for the gross sample and requested their data from the residents registries in Berlin and Brandenburg. Originally, we aimed at an equal distribution of age groups and sexes within the citizenships for sampling. Our final gross sample consisted of only 9068 people due to the lack of people aged over 65 years in some citizenships. In particular, there were few Syrian and Romanian people aged over 65 years in the lists provided by the residents

registries. In Brandenburg this was also problematic for Polish people (see Table 2). The sample thus systematically lacked older age groups.

Table 2. Frequency distribution of the gross sample by sample point, citizenship, group, sex, and age.

Characteristic	Sample point[a]										Total (N=9068)
	Berlin (n=8255)							Brandenburg (n=813)			
Citizenship	Turkish		Syrian		Romanian	Croatian	Polish	Syrian		Polish	
Group	A (n=1411)	B[b] (n=1408)	A (n=905)	B (n=880)	A (n=1220)	A (n=1410)	A (n=1021)	A (n=251)	B (n=240)	A (n=322)	
Age range (years): male	706	703	473	463	606	705	511	131	126	155	4579
18-44	255	255	217	212	300	265	185	68	66	68	1891
45-64	242	240	151	148	213	217	183	52	50	68	1564
≥65	209	208	105	103	93	223	143	11	10	19	1124
Age range (years): female	705	705	432	417	614	705	510	120	114	167	4489
18-44	257	259	212	207	301	249	183	68	66	68	1870
45-64	237	236	159	152	207	227	181	46	44	66	1555
≥65	211	210	61	58	106	229	146	6	4	33	1064

[a]In Berlin it was possible to draw people with all 5 target citizenships from the residents registry. In Brandenburg this was not practicable, and only people with Syrian and Polish citizenship were selected due to low numbers of people with Turkish, Romanian, and Croatian citizenship.

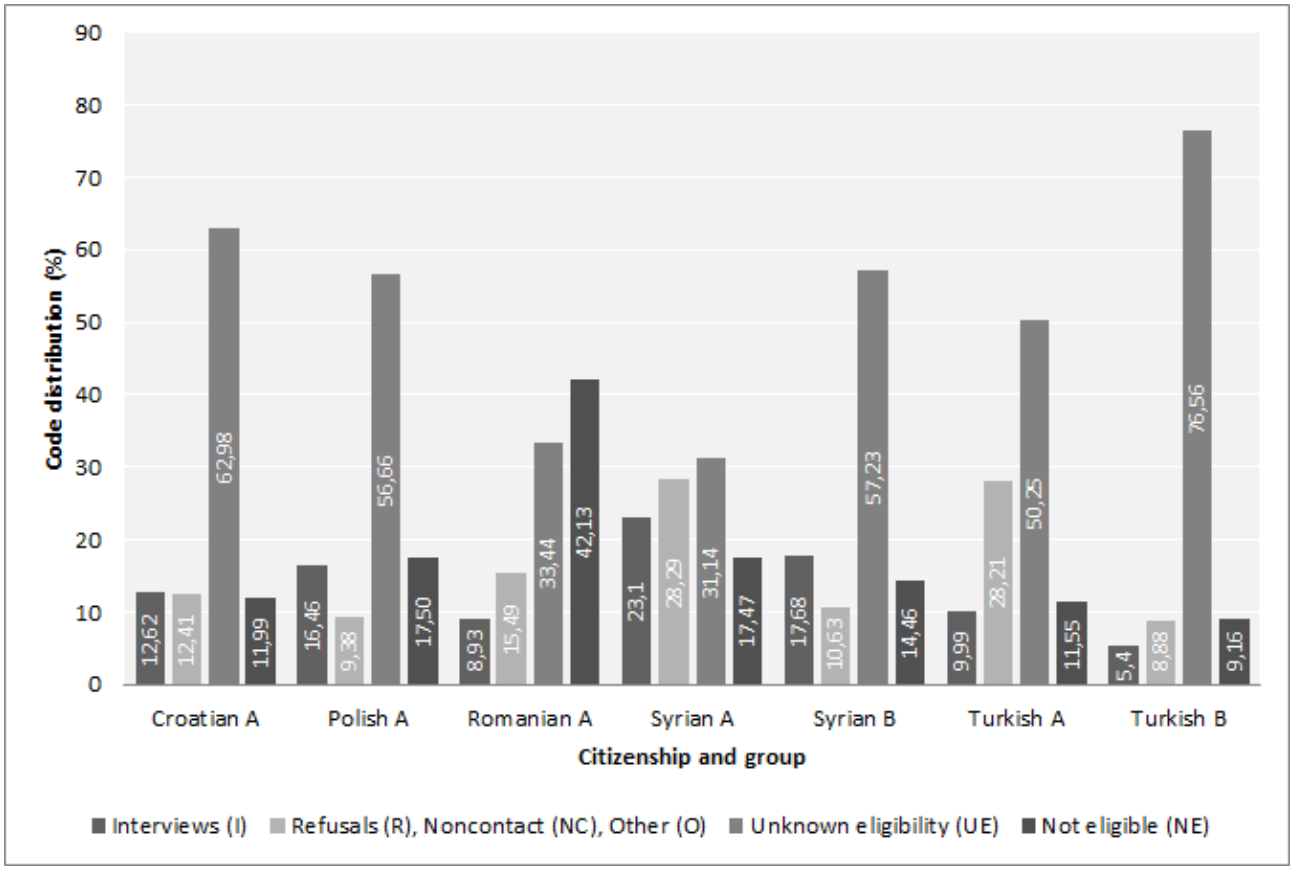
[b]Reference group B was drawn from Turkish and Syrian citizens only.

Final Disposition Codes

At the end of the survey, we assigned every sampled case a final disposition code to enable the later analysis of response rates. We adapted the disposition codes from the guidelines of the American Association for Public Opinion Research (AAPOR), differentiating the sample into 4 categories: (1) interviews (ie, cases with a completed questionnaire); (2) refusals (R), noncontacts (NC), or other (O) (ie, cases we had contact with, not resulting in participation, eg, people who called the study hotline to refuse participation; (3) cases of unknown eligibility (UE) (ie, cases we did not know anything about, eg, people who did not react to our contact attempts at all; and (4) cases that were not eligible (NE) for the survey (ie, the selected person was unknown in the household or the address was not correct, eg, when mailed cover letters were returned as undeliverable or we received information that someone had died) [41].

As Figure 3 shows, the distribution of the disposition codes varied greatly with regard to citizenship. Most completed interviews were conducted with Syrian (group A: 267/1156, 23.10%; group B: 198/1120, 17.68%), Polish (221/1343, 16.46%), and Croatian (178/1410, 12.62%) participants. Significantly lower proportions of Turkish (group A: 141/1411, 9.99%; group B: 76/1408, 5.40%) and Romanian (109/122, 8.93%) participants than participants with other citizenships completed the interviews. Contact resulting in no participation (R NC O) were most frequent with the Syrian (327/1156, 28.29%) and Turkish (398/1411, 28.21%) participants of group A, followed by Romanian (189/1220, 15.49%) and Croatian (175/1410, 12.41%) participants. All other citizenship groups had R NC O shares of around 10% (Polish group: 126/1343, 9.38%; Syrian group B: 119/1120, 10.63%; Turkish group B: 125/1408, 8.88%). Differences between the citizenships and the survey design for groups A and B can be identified regarding cases with no further information about eligibility (UE). UE had the greatest share in the Turkish group B (1078/1408, 76.56%), Croatian (888/1410, 62.98%), Syrian group B (641/1120, 57.23%), and Polish (761/1343, 56.66%) participants. Among Romanian participants, only 33.44% (408/1220) were UE, in contrast to 42.13% (514/1220) of the cases that were NE. The percentage of Romanian NE cases was 3 times higher than for the other citizenships, with only 13.50% (1060/7848) of NE cases on average, ranging from 9.16% (129/1408) in the Turkish group B to 17.47% (202/1156) in the Syrian and 17.50% (235/1343) in the Polish group A. When comparing the intervention group A with the control group B, it was striking how the measures in group A led to an increase in participation and, furthermore, to more information about the sample.

Figure 3. Distribution of American Association for Public Opinion Research disposition codes (%) by citizenship and assignment to intervention group A or control group B.



Process and Qualitative Data From the Study Hotline

General Results

The study hotline was available during the first 3 survey phases from January to April 2018. Outside the interviewers' working hours an answering machine was set up. We received 409 contacts during this period; 21 of these were repeat contacts. Most calls came from Syrian (130/405, 32.1%), Turkish (101/405, 24.9%), and Croatian (83/405, 20.5%) participants. There were no significant differences between the contact frequency of group A (multilingual hotline; 296/6540, 4.52%) and group B (German hotline; 109/2528, 4.31%). The study hotline was evaluated positively in the interviewers' focus group. Participants who had been living in Germany for a long time associated a "warm and nostalgic feeling" with the multilingual hotline. The interviewers described the use of the native language as an opportunity "to get a foot in the door." Personal stories, hobbies, or the same origin facilitated the communication process and built "a feeling of connectedness," Polish, Romanian, and Croatian interviewers reported that the language changed frequently during a single call. The Syrian interviewers reported the use of the Arabic language in most cases, but they faced great difficulties due to the variety of Arabic dialects. Similar experiences were reported by Turkish interviewers, who claimed that communication was impossible with Kurdish-speaking people. Communication in group B was not entirely possible due to the lack of bilingual interviewers.

Call times varied between 5 and 41 minutes in both groups, excluding calls during which telephone interviews took place. The hotline was contacted slightly more often by women (198/384, 51.6%) than by men (186/384, 48.4%), which holds true for all citizenships except Syrian, where more men called the study hotline. In general the hotline was most often used by elderly people. In the Croatian, Polish, and Romanian groups, the share of elderly people (≥65 years) was more than 50% (96/168, 57.1%). In the Syrian group there was no difference in contact frequency by age. The interviewers reported that loneliness and the ability to speak to someone were relevant factors in calling the study hotline for older respondents. A total of 70.7% (289/409) of contacts were established by the target person (the person invited to participate in the IMIRA survey). The only exception was in the Turkish group, where only 58.4% (59/101) of contact was with the target person. When the call was made through a contact person, it was mainly a close family member, such as a child or partner. The reasons for calling the study hotline were documented in a process questionnaire in 4 categories: (1) queries regarding the survey, (2) issues regarding participation, (3) telephone interview (possible from phase 2), and (4) refusal.

Queries

Of all contacts through the study hotline, 23.6% (107/453) involved general queries about the survey. The most queries reported by the interviewers in the focus group involved participation ("Do I need to go somewhere?"), content of the survey ("Why is the survey being conducted?"), and the contracting authority of the survey ("Who is conducting the survey?"). There were slightly fewer questions about the sampling of respondents ("Why was I chosen to participate in

the survey?") and data protection ("Will my data be shared with others?"). Data from the process questionnaire provided further information about the citizenship of the callers. Syrian people mostly had queries regarding participation, the content of the survey, and data protection. Questions about the contracting authority or sampling were more often asked by Croatian and Polish people.

Issues

Issues with participation were the reason for calling the study hotline in 9.7% (44/453) of contacts. The problems documented in the process questionnaire were largely technical: the online questionnaire could not be opened, the code to participate in the survey was not working, or there was no internet access or device available for online participation. Some callers requested a paper questionnaire. These technical issues were more often reported by Polish, Croatian, and Syrian participants. The Syrian interviewers reported that literacy problems and illiteracy made it impossible for some participants to take part in the survey without a telephone interview.

Telephone Interview

A telephone interview was promoted in the first reminder in group A, resulting in 93 telephone interviews lasting an average of 50.5 minutes. Of all the callers, 9.3% (42/453) asked directly for the telephone interview as promoted in the cover letters. Participation via telephone interview varied greatly between the citizenships. Most telephone interviews took place in the Croatian (n=29), Polish (n=25), and Syrian (n=22) groups, while Romanian (n=9) and Turkish (n=8) callers did not use this option.

Issues with the telephone interview were reported in the interviewers' focus group, which could not be found in the descriptive data of the process questionnaire. According to the interviewer, the quality of the translated questionnaires was not very good. Although the questionnaire was translated by a professional translation agency, incorrect translations were belatedly identified in the process of data collection. The questionnaire was also perceived as very long and complex, and therefore unsuitable for call-assisted administration. According to the interviewers, participants rated the questionnaire as very intimate and "psychologically stressful," especially if they had serious diseases or traumatic experiences that affected their health status permanently. Interviewers were thus forced into the role of a psychological therapist when participants told them about their traumatic experiences due to war, torture, or violence. In these cases contact information was provided for advisory offices specializing in these issues, to lessen the burden on the interviewers. Specific problems were also identified, such as questions regarding religion, which were considered offensive in the Syrian and Turkish groups, or sociodemographic questions about health, which were considered to be irrelevant in the Romanian group.

Refusals

Half the callers intended to refuse participation (237/453, 52.3%). They were predominantly Turkish, Syrian, or Croatian citizens. Even though interviewers were trained to promote the survey, the majority could not be persuaded.

The most frequent reason for refusal was a lack of interest in the survey (148/384, 38.5%). Interviewers in the focus group reported that they tried to convince people who commented on the topic of the survey that it was "none of my business." Polish interviewers referred to different living environments and explained that Polish people were predominantly in Germany for work, not to live permanently. Thus, questions about health were not of interest because health-related action, such as visits to the doctor, took place in Poland rather than in Germany. The 2 reminder letters, and especially the announcement of a home visit as part of the last reminder, were perceived as "nuisance," "invasion of privacy," and "involuntary." According to the interviewers, the objectives of the survey were not well understood. Participants reported having "no trust that something will change due to their participation" and criticized the focus on PMB ("Are there diseases that only affect special population groups?"). The interviewers reported that some callers complained about being assigned to the group of migrants and felt discriminated against in the survey. Croatian, Syrian, and Turkish participants most commonly refused due to having "no interest."

The second most frequent reason for refusal was a lack of time (57/384, 14.8%), which was particularly common for Croatian, Syrian, and Turkish participants. Other reasons that prevented participation were health restrictions or literacy issues, predominantly given by Syrian and Turkish participants. The interviewers reported that there were various calls reporting that participants had moved to another address or were deceased. Although these interviewers reports could not be confirmed quantitatively, confrontation with death was considered challenging by the interviewers and was highlighted in the focus group.

Process and Qualitative Data of the Home visits

General Results

Of the 1822 participants selected for the home visits, only 945 (51.87%) could be visited by the Romanian, Syrian, and Turkish interviewers in the given time. Overall the interviewers reported that contact with target people was "mostly polite and friendly." The Turkish interviewers in particular evaluated their strategies of persuasion in direct personal contact with the sample targets as particularly convincing. Romanian interviewers emphasized the positive effect of the home visits with the following statement: "If you are standing at the front door, you will be let in for sure," meaning that the main challenge was not convincing people to participate, but making first contact. When personal contact was accomplished, it was with the target person directly in 78.5% (551/702) of cases. Of these contacts with the target persons, 58.9% (334/567) led to survey participation.

Differences could be identified between the groups: Syrians consented in 74.4% (203/273) of all cases, whereas Turkish participants did so in only 43.7% (101/231). This is in line with the overall response rate to the survey. More Turkish participants than people with other citizenships also needed to be contacted more than once. The contact language was mainly not German. Interviewers reported that it was important for some participants to be able to use their mother tongue during the interview. This was especially demonstrated by the fact that no interviews in

the last survey phase were conducted in German, although interviewers estimated that German was possible in 9.7% (45/465) of communication. Interviewers reported different characteristics of language as related to citizenship. Older Turkish participants were more frequently not German speaking; thus, bilingual approaches should be taken into account for this group. Romanian participants could not be contacted in German during the home visits at all. Similar experiences were reported by the Syrian interviewers, where communication took place in Arabic only. They stated that some medical terms were not understood in Arabic but were understood in English. Language barriers that could not be attributed to the questionnaire itself were reported for people of all citizenships, for example, challenges with Kurdish-speaking people in the Turkish or Syrian group or Russian-speaking people in the Romanian group.

Interviews

Of all contacts, 11.0% (104/945) led to a face-to-face interview performed in the participant's household directly, and 5.5% (52/945) led to a subsequent telephone interview after the telephone number was recorded (100/945, 10.6%). The interviewers' focus group noted that obtaining the phone number for a subsequent telephone interview should be considered an emergency option if the respondent desired no alternative. Both participation options during the fourth survey phase were taken up by Syrian or Turkish participants. Of the participants targeted, 51.2% (334/652) were female. Slightly more often women participated in face-to-face interviews (57/100, 57.0%) than in the telephone interviews (25/46, 56%), which had a more balanced sex distribution. Participants were generally 18 to 64 years old (114/146, 78.1%). The youngest age group (18-44 years) was mainly represented by Romanian (14/58, 24%) and Syrian participants (31/58, 53%), and more people from the older age group (≥65) participated from the Turkish group (18/32, 56%) (see Table 3).

Table 3. Age and citizenship frequencies in the interviews in the home visit phase.

Citizenship	Age range (years)			Total (N=146)
	18-44 (n=58)	45-64 (n=56)	≥65 (n=32)	
Romanian	14 (24.1)	8 (14.3)	2 (6.3)	24 (16.4)
Syrian	31 (53.4)	25 (44.6)	12 (37.5)	68 (46.6)
Turkish	13 (22.4)	23 (41.1)	18 (56.3)	54 (37.0)

The interviewers described various situations in which they felt stressed, for example, when they were confronted with diseases, misfortunes, or participants "tending to show depressive symptoms." In these situations they stated that it was difficult to keep their distance or interrupt the flow of speech without being offensive. The traumatic experiences of some participants (eg, during war) led to a feeling of helplessness for the interviewers. In this respect the questionnaire was described as very formal, as opposed to the personal character of the home visits. Similar to the experiences in the study hotline, the interviewers reported that questions regarding migration history or religious affiliation offended some participants so much that interviewers had difficulties continuing the interview. Questions regarding income or the consumption of alcohol were also viewed critically.

Refusals

Of all contacts during phase 4, 33.2% (314/945) resulted in refusal, especially in the groups of Syrian (173/314, 54.8%), Turkish (76/314, 24.2%), and Romanian (42/314, 13.4%) participants. The main reason for refusing participation was "no interest" in the survey (138/235, 58.7%). Interviewers stated that there were some issues involving the RKI that led to a misunderstanding about the survey's aims or the importance of research in general. Interviewers were, for example, forced to explain the RKI and its function because it was not known by the respondents. In the Turkish and Romanian groups, skepticism about and mistrust in research caused refusal. There was, for example, no understanding of data security and data sharing. Some participants felt as if they were under surveillance by the German government and believed that hidden information was being gathered with the survey. As in the process of standardization for the study hotline, interviewers were asked to evaluate the effectiveness of their attempts to convince participants after refusal. The interviewers were not successful in handling objections in 96.1% (172/179) of all refusals. Syrian participants were most likely to be convinced, but only half participated in the end, which was the same finding as in the study hotline.

Sample Dropouts

A total of 26.8% (253/945) of the participants could not be reached during the home visit phase. More than half of these sample dropouts (138/253, 54.2%) were Romanian participants, confirming the previous results shown for the AAPOR disposition codes. The reasons for sample dropout were very different between populations. Whereas Romanian participants were mainly not available due to wrong addresses or the lack of a name on the front door or mailbox, it was predominantly relocations that caused nonparticipation in the Syrian group. Interviewers also reported that they applied different strategies to overcome these barriers; for example, 1 of the Romanian interviewers walked up the staircase in the residential houses to identify the right apartment if no name was on the doorbell nameplate. Syrian interviewers reported that sometimes there was no possibility of reaching participants living in shared accommodation, since no name was on the doorbell or because access was blocked by security guards. The interviewers thus evaluated the reachability of participants and quality of addresses very differently.

Discussion

Principal Findings

The main objective of the IMIRA survey was to identify recruitment measures to improve the response of PMB to health surveys by the RKI. Although the IMIRA survey focused on people with non-German citizenship only, the findings are of value for designing subsequent health surveys at the RKI. We tested multiple interventions in the IMIRA survey, which focused primarily on overcoming language barriers, enhancing personal contact with bilingual interviewers, and optimizing bilingual survey information materials or a multilingual questionnaire. The emphasis of the IMIRA survey was on diversity-sensitive information with a special focus on language and the design of the documents. The materials consisted of a cover letter and a survey brochure, and were developed with representatives of the target population in a focus group. Although the lack of varying materials for different citizenships was criticized in the focus group, this idea of varying materials was waived due to limitations to the capacity of survey management. Our experience with this survey shows that alterations for people with different citizenships should be implemented, although it might lead to additional costs. Otherwise, there is a risk that participants would feel discriminated against or misunderstand the survey's objectives, as happened in this IMIRA survey. The main reason for this was that the survey materials were not tailored for each citizenship group. Germans with additional foreign citizenship and people with only foreign citizenship were not differentiated in the IMIRA survey. Addressing the whole sample as migrants resulted in an external attribution that might have been in contrast to people's own personal attribution of origin or affiliation, and this might lead to a feeling of discrimination. This aspect has been discussed in recent research and should be taken seriously when designing survey materials or a communication strategy in subsequent surveys [42].

Personal contact is very often described in the literature as the "silver bullet" for achieving better response rates in hard-to-reach populations [20,21,28,43]. For PMB, who can also be considered hard to reach, personal contact was an essential component of the IMIRA survey design. The study hotline was set up with bilingual interviewers offering the opportunity to address queries regarding the survey or participation directly. Home visits with bilingual interviewers were carried out in the last survey phase, turning the previously rather indirect contact into a direct contact. The interviewers were members of the target population themselves and thus capable of addressing the survey topics not only in the relevant language, but also on a diversity-sensitive basis. This was intended to minimize inhibitory factors, such as lack of trust in research or literacy issues, which were often discussed in the literature, or a lack of internet access [8,21]. The attempt to eliminate language barriers worked very well, as the findings show. Interviewers reported the frequent use of languages other than German on the study hotline. Interviewers on the German-only hotline (group B) in particular reported that sometimes no communication was possible due to language issues. Communication during the home visits took place almost exclusively in languages other than German and would not have been possible in German according to interviewer assessments. Home visits increased participation frequency. Personal contact through the telephone was not as effective as we had assumed it would be and had no effect on response. The study hotline was primarily used for refusals and less for queries regarding the survey, and thus primarily had an informative value.

Another important emphasis of the IMIRA survey was on offering different participation options and interventions to lessen the burden of participation. Our respondents had the option to participate online, by telephone, or (only in a small subsample) face-to-face, using multilingual questionnaires and bilingual interviewers to overcome language barriers as recommended in research [21,22]. The questionnaire was administered in a mixed-mode design, known to be connected to higher response rates, since it takes a variety of participation preferences into account [44]. Participants were first invited to participate online. With each new contact, other modes, such as telephone or face-to-face interviews, were offered sequentially in group A. This did increase the response rate and in previous research was shown to be more effective than a simultaneous mixed-mode design [45]. In consideration of the heterogeneity of the target population, a combination of self- and interviewer-administered questionnaires seemed to be of value to facilitate survey participation [46], although it might have had implications for data quality in terms of, for example, interviewer effects or the discrepancy between auditory and visual comprehension of the questionnaire [47,48]. The online questionnaire was accessible via mobile phone: participants had the opportunity to answer the questionnaire "on the go." This was intended to reduce refusals for time management reasons. The questionnaire could also be interrupted, further promoting compatibility with the participant's leisure behavior. Research has shown that younger people might be more attracted by the opportunity to participate online, whereas older people tend to prefer interviewer-administered modes (telephone or face-to-face interviews) [49]. This was confirmed in the IMIRA survey with regard to the use of the study hotline and home visits, where the composition of the sociodemographic sample differed significantly between self- and interviewer-administered modes.

Alongside the survey's objective to increase responses, another important aim was to identify reasons for nonresponse of PMB. Every contact with the participants, through the study hotline or through the home visits, was therefore documented in a process questionnaire. The results showed that more than half of the callers wanted to refuse participation. Interviewers were required to document the reasons as accurately as possible. We identified differences between citizenship groups: Croatian, Syrian, and Turkish participants primarily called the study hotline to refuse participation. During home visits Turkish and Syrian participants more frequently refused participation. The most common reasons in both cases were "no interest in the survey" or "no time." Interviewers in the focus groups described the refusal reasons in more detail than was possible in the process questionnaire. They stated that participants felt disturbed by the survey, did not understand the survey's aims, or had no interest in the research and the possible implications of the survey results in their everyday life. This is similar to the

increasing trend of survey fatigue found in other research in recent decades [50,51]. More frequent headlines in the press concerning the falsification of survey results could also explain this negative trend. Other reasons for refusal were given by the interviewers as well, involving the subject of the IMIRA survey. Although this could not be confirmed quantitatively, interviewers described the survey's focus on improving health monitoring in migrant populations as causing conflict, especially for those who did not identify with the label "migrant." According to the interviewers, there was suspicion that people who had lived for a long time in Germany were being put on a level with recent refugees or newly migrated people. Complaints about this were predominantly made by Croatian and Romanian participants, whose migration histories are from a more distant past. The Turkish interviewers described other issues regarding the focus on migrants. Many participants did not react to the survey invitations. Only during home visits could participants be questioned about their unwillingness to participate. The reported indifference to survey participation might be the result of being ignored by German institutions and politics for a long time and a resulting development of social segregation from the host population [24,52]. Interviewers stated that the only way to convince Turkish people to participate was to argue at the institutional level, for example, emphasizing that the government was now interested in their health status and that it was therefore important to provide the information.

In addition to providing detailed reasons for refusal, the process questionnaire also improved the quality of the sample with regard to dropouts due to relocation, death, or wrong addresses. This was striking especially in the Romanian group, who had the greatest share of noncontactable (NE) people in the whole sample. A large proportion of these cases were identified by undeliverable mail; their address was wrong or nonexistent. Other cases could only be identified by the interviewers in the home visits phase, especially when no name was on the doorbell or mailbox, or if the apartment was sublet. The high mobility of people with Romanian citizenship was in contrast to the relatively rigid system of population registration, which makes the sampling of mobile groups more difficult and does not seem appropriately representative [16,53].

Limitations and Challenges

We were bound by the general procedure at RKI when designing the IMIRA survey, using population registers for sampling. Two challenges became clear as a result. First, we had to sample people according to their citizenship, since there was no other practicable way to sample PMB for our survey, as described earlier in the theoretical considerations. As a result we excluded naturalized PMB and subsequent generations of PMB from our target population, who had no citizenship other than German. Second, the sample size decreased more than expected, beginning with insufficient numbers of people who could have been selected for the sample. We requested a disproportionate sample from the residents registration offices, stratified by sex and age for each citizenship. Each cell should have been filled equally for a better comparison of the groups in later analysis. The sample could not be delivered as requested, especially for the Syrian and Romanian groups 65 years of age or older, and for the 2 sample points in Brandenburg for all citizenships in the oldest age category. Of 9800 requested individuals, a sample of only 9068 people was delivered. The sample size was also decreased by invalid addresses due to relocation or for other reasons. In previous RKI studies, the quality-neutral dropout rate due to invalid postal addresses was about 10% to 15% of the gross sample [12]. According to research it can be assumed that the percentage may be even higher in mobile population groups, such as newly arrived migrants or asylum seekers [16]. The quality-neutral dropout rate in the IMIRA survey was 17.36% (1574/9068), with great differences between the citizenships, ranging from 10.74% (169/1574) in the Croatian group to 32.66% (514/1574) in the Romanian group. There were a variety of reasons for the reduced gross sample, and these should be considered when drawing a sample from population registers. In future surveys it would be advisable to use a proportional stratified sample of the target populations. In addition, the possible dropout rate should be taken into account and the gross sample should be increased accordingly. For some target populations this might still be an inadequate sampling method, and other ways of sampling should be taken into account.

Although the survey design was sequential, the effectiveness of the various interventions cannot clearly be separated, since the prior interventions were still accessible when the new interventions were initiated. We also used the survey's experimental approach to evaluating the interventions only for Syrian and Turkish participants and for only the first 3 survey phases. The effectiveness of the interventions could not be evaluated for the other citizenships, and the analysis was descriptive only. The impact of home visits on response rate could be evaluated only for Romanian, Syrian, and Turkish participants.

The questionnaire was available to all participants in Arabic, Croatian, German, Polish, Romanian, Syrian, and Turkish. According to the interviewers, who were native speakers, the translation of the questionnaire was not entirely adequate. In addition, the questionnaire was not culturally adapted for the target populations. As described in the results, some interviewers reported that respondents had strong reservations about certain questions, for example, whether a respondent was a member of a religious community. The negative experiences with the quality of the translation will have consequences for future surveys. We will conduct a cognitive test of instruments that are not already validated in the requested languages and that might be identified as particularly sensitive with representatives of the target populations. We are endeavoring to develop standard operating procedures for externally commissioned translations at the RKI to ensure the quality of future studies. We will use the team approach for translations, in which 2 independent translators simultaneously produce a translation. Differences between the 2 translations are discussed by the team, meaning between the 2 translators and a moderator who acts as a mediator [54].

Further studies in this field of research should acknowledge these limitations. In particular, the evaluation of various aspects of the survey recruitment measures should be studied in more detail. For example, to our knowledge, there is still no research

into the effect of multilingual questionnaires on response rate. Experimental studies are thus needed.

Conclusion

The IMIRA survey met the challenges of low response rates of PMB with tailored interventions, which were shown to be more or less effective. Bilingual information material and questionnaires were received very positively by the participants, although there were some issues regarding translation that we hope to avoid in future surveys through the team approach. The study hotline did not have the expected effect, but it proved to be of value for participants with queries, with technical or other survey-related issues, or wanting to refuse participation. It is therefore advisable to establish a contact channel for respondents in future surveys. The study hotline also enabled us to draw lessons for the future in terms of structuring survey information materials and providing the required support for participants. Home visits with bilingual interviewers had the greatest effect on responses and will be a firm component of our future surveys when considering PMB. We gathered qualitative and procedural data through focus groups with the interviewers and contact protocols during the fieldwork phases, providing information about reasons for refusal or issues leading to nonresponse, and helping to better characterize the sample in general. Nevertheless, some of the mentioned reasons for nonresponse were still very unspecific and should be investigated in more detail in subsequent surveys.

Acknowledgments

The Improving Health Monitoring in Migrant Populations (IMIRA) survey was a cross-sectional feasibility study funded by the German Ministry of Health (grant number: ZMVI1-2516FSB408) and conducted by the Robert Koch Institute (RKI). The RKI is the German national public health institute with a mandate for epidemiology and health monitoring. The authors especially thank the participants for taking the time to participate in the IMIRA survey, the interviewers, and the project management team at USUMA GmbH, who implemented the field work, and all colleagues at RKI who support the IMIRA survey. The contents of this publication are solely the responsibility of the authors.

Conflicts of Interest

None declared.

References

1. Statistisches Bundesamt. Bevölkerung und Erwerbstätigkeit: Bevölkerung mit Migrationshintergrund. Ergebnisse des Mikrozensus. Fachserie 1 Reihe 2.2 2017. Wiesbaden, Germany: Destatis; 2018. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Publikationen/Downloads-Migration/migrationshintergrund-2010220177004.html [accessed 2018-09-25]
2. Butterwegge C. Von der “Gastarbeiter”-Anwerbung zum Zuwanderungsgesetz. Migrationsgeschehen und Zuwanderungspolitik in der Bundesrepublik. Grundlagendossier Migration. Bonn, Germany: Bundeszentrale für politische Bildung; 2005. URL: http://www.bpb.de/gesellschaft/migration/dossier-migration-ALT/56377/migrationspolitik-in-der-brd?p=all
3. Bundesamt für Migration und Flüchtlinge. Migrationsbericht des Bundesamtes für Migration und Flüchtlinge im Auftrag der Bundesregierung: Migrationsbericht 2015. Berlin, Germany: Bundesministerium der Innern; 2015. URL: https://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/Migrationsberichte/migrationsbericht-2015.pdf?__blob=publicationFile [accessed 2018-09-30]
4. Salentin K, Schmeets H. Sampling immigrants in the Netherlands and Germany. Comp Migr Stud 2017;5(1):21 [FREE Full text] [doi: 10.1186/s40878-017-0062-2] [Medline: 29264233]
5. Saß AC, Grüne B, Brettschneider AK, Rommel A, Razum O, Ellert U. [Participation of people with migration background in health surveys of the Robert Koch Institute]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015 Jun;58(6):533-542. [doi: 10.1007/s00103-015-2146-1] [Medline: 25896496]
6. Feskens RC. Difficult Groups in Survey Research and the Development of Tailor-Made Approach Strategies. Utrecht, Netherlands: Centraal Bureau voor de Statistiek Utrecht University; 2009.
7. Méndez M, Ferreras M, Cuesta M. Immigration and general population surveys in Spain: the CIS Surveys. In: Méndez M, Font J, editors. Surveying Ethnic Minorities and Immigrant Populations: Methodological Challenges and Research Strategies. Amsterdam, Netherlands: Amsterdam University Press; 2013:271-291.
8. 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;14:42 [FREE Full text] [doi: 10.1186/1471-2288-14-42] [Medline: 24669751]
9. Gabler S, Häder S. Sampling in Theory. GESIS Survey Guidelines. Mannheim, Germany: GESIS Leibniz-Institut für Sozialwissenschaften; 2016. URL: https://www.gesis.org/fileadmin/upload/SDMwiki/GablerH%C3%A4der_Sampling_in_Theory.pdf [accessed 2018-10-22]
10. Cornesse C, Bosnjak M. Is there an association between survey characteristics and representativeness? A meta-analysis. Surv Res Methods-Ger 2018;12(1):1-13 [FREE Full text] [doi: 10.18148/srm/2018.v12i1.7205]

11. Lange C, Finger JD, Allen J, Born S, Hoebel J, Kuhnert R, et al. Implementation of the European health interview survey (EHIS) into the German health update (GEDA). Arch Public Health 2017;75:40 [FREE Full text] [doi: 10.1186/s13690-017-0208-6] [Medline: 28936356]
12. Hoffmann R, Lange M, Butschalowsky H, Houben R, Schmich P, Allen J, et al. Querschnitterhebung von KiGGS Welle 2 -- Teilnehmendengewinnung, Response und Repräsentativität. J Health Monitor 2018;3(1) [FREE Full text] [doi: 10.17886/RKI-GBE-2018-015]
13. Kamtsiuris P, Lange M, Schaffrath Rosario A. [The German Health Interview and Examination Survey for Children and Adolescents (KiGGS): sample design, response and nonresponse analysis]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2007;50(5-6):547-556. [doi: 10.1007/s00103-007-0215-9] [Medline: 17514438]
14. Lange M, Butschalowsky HG, Jentsch F, Kuhnert R, Schaffrath Rosario A, Schlaud M, et al. Die erste KiGGS-Folgebefragung (KiGGS Welle 1). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2014 Jun 21;57(7):747-761. [doi: 10.1007/s00103-014-1973-9]
15. Gößwald A, Lange M, Dölle R, Hölling H. Die erste Welle der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2013 May 25;56(5-6):611-619. [doi: 10.1007/s00103-013-1671-z]
16. Salentin K. Sampling the ethnic minority population in Germany. The background to 'Migration Background'. Methods Data Anal 2014;8(1):25-52 [FREE Full text] [doi: 10.12758/mda.2014.002]
17. Kamtsiuris P, Lange M, Hoffmann R, Schaffrath Rosario A, Dahm S, Kuhnert R, et al. [The first wave of the German Health Interview and Examination Survey for Adults (DEGS1): sample design, response, weighting and representativeness]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2013 May;56(5-6):620-630. [doi: 10.1007/s00103-012-1650-9] [Medline: 23703478]
18. Liebau E, Humpert A, Schneiderheinze K. Wie gut funktioniert das Onomastik-Verfahren? Ein Test am Beispiel des SOEP-Datensatzes. 976-2018. SOEPpapers on Multidisciplinary Panel Data Research. Berlin, Germany: German Socio-Economic Panel, DIW Berlin; 2018. URL: https://www.diw.de/documents/publikationen/73/diw_01.c.595744.de/diw_sp0976.pdf [accessed 2018-11-16]
19. Schnell R, Trappmann M, Gramlich T. A study of assimilation bias in name-based sampling of migrants. J Off Stat 2014;30(2):231-249. [doi: 10.2478/jos-2014-0015]
20. Frank L, Yesil-Jürgens R, Born S, Hoffmann R, Santos-Hövener C, Lampert T. Maßnahmen zur verbesserten Einbindung und Beteiligung von Kindern und Jugendlichen mit Migrationshintergrund in KiGGS Welle 2. J Health Monit 2018;3(1). [doi: 10.1055/s-0037-1605753]
21. Méndez M, Font J. Surveying immigrant populations: methodological strategies, good practices, open questions. In: Méndez M, Font J, editors. Surveying Ethnic Minorities and Immigrant Populations: Methodological Challenges and Research Strategies. Amsterdam, Netherlands: Amsterdam University Press; 2013:271-291.
22. Schenk L, Neuhauser H. [Methodological standards for migrant-sensitive epidemiological research]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2005 Mar;48(3):279-286. [doi: 10.1007/s00103-004-0995-0]
23. Kroh M, Brücker H, Kühne S, Liebau E, Schupp J, Siegert M, et al. Das Studiendesign der IAB-BAMF-SOEP-Befragung von Geflüchteten. SOEP Survey Papers 365: Series C: Data Documentation. Berlin, Germany: German Socio-Economic Panel, DIW Berlin; 2016. URL: https://www.econstor.eu/bitstream/10419/155353/1/880984651.pdf [accessed 2018-11-26]
24. Reiss K, Dragano N, Ellert U, Fricke J, Greiser KH, Keil T, et al. Comparing sampling strategies to recruit migrants for an epidemiological study. Results from a German feasibility study. Eur J Public Health 2014 Oct;24(5):721-726. [doi: 10.1093/eurpub/cku046] [Medline: 24872519]
25. Ipsos MORI Social Research Institute. GP Patient Survey--Technical Annex. 2018 Annual Report. London, UK: Ipsos MORI Social Research Institute; 2018 Aug. URL: https://gp-patient.co.uk/Downloads/archive/2018/GPPS%202018%20Technical%20Annex%20PUBLIC.pdf [accessed 2018-08-10]
26. Stadler B. Die Befragung von MigrantInnen in Stichprobenerhebungen. In: Weichbold M, Bacher J, Wolf C, editors. Umfrageforschung, Herausforderungen und Grenzen. Wiesbaden, Germany: Springer; 2009:275-291.
27. Häder M. Empirische Sozialforschung: Eine Einführung. Wiesbaden, Germany: Springer; 2015.
28. Blohm M, Diehl C. Wenn Migranten Migranten Befragen. Zum Teilnahmeverhalten von Einwanderern Bei Bevölkerungsbefragungen. Z Soziol 2001;30:223-242 [FREE Full text] [doi: 10.1515/zfsoz-2001-0304]
29. Mohammadi N, Jones T, Evans D. Participant recruitment from minority religious groups: the case of the Islamic population in South Australia. Int Nurs Rev 2008 Dec;55(4):393-398. [doi: 10.1111/j.1466-7657.2008.00647.x] [Medline: 19146549]
30. Statistisches Bundesamt. Bevölkerung und Erwerbstätigkeit: Ausländische Bevölkerung, Ergebnisse des Ausländerzentralrgisters 2017. Fachserie 1 Reihe 2. Wiesbaden, Germany: Destatis; 2018. URL: https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00035266/2010200177004_korr13062018.pdf [accessed 2018-05-10]
31. Statistik Berlin Brandenburg. Statistisches Informationssystem Berlin Brandenburg (StatIS-BBB). 2017. URL: https://www.statistik-berlin-brandenburg.de/datenbank/inhalt-datenbank.asp [accessed 2017-03-14]

32. Statistik Berlin Brandenburg. Bevölkerung, Einbürgerungen, Ausländer, Basisdaten. 2017. URL: https://www.statistik-berlin-brandenburg.de/BasisZeitreiheGrafik/Bas-Einbuergerungen_Auslaender.asp?Ptyp=300&Sageb=12025&creg=BBB&anzwer=7
33. European Union. European Health Interview Survey (EHIS Wave 2): Methodological Manual. eurostat Methodological Working Papers. Luxembourg: Publications Office of the European Union; 2013. URL: https://circabc.europa.eu/sd/a/20e4ad07-78a3-4873-b5b4-495ba2b0f8d6/European%20Health%20Interview%20Survey%20%28EHIS%20wave%202%29%20-%20Methodological%20manual%20-%202013%20edition.pdf [accessed 2018-07-05]
34. Eurostat. Europäische Gesundheitsumfrage (EHIS): Beschreibung des Datenbestands. URL: https://ec.europa.eu/eurostat/de/web/microdata/european-health-interview-survey
35. TNS Infratest Sozialforschung. Erhebungsinstrumente der IAB-BAMF-SOEP-Befragung von Geflüchteten 2016: Integrierter Personen- und Biografiefragebogen, Stichproben M3-M4. SOEP Survey Papers 362, Series A - Survey Instruments (Erhebungsinstrumente). Berlin, Germany: German Socio-Economic Panel, DIW Berlin; 2016. URL: https://www.diw.de/documents/publikationen/73/diw_01.c.570916.de/diw_ssp0362.pdf [accessed 2018-10-30]
36. Wittlif A, Beigang S. Methodenbericht: SVR-Integrationsbarometer. Berlin, Germany: Sachverständigenrat deutscher Stiftungen für Integration und Migration (SVR) GmbH; 2016. URL: https://www.svr-migration.de/wp-content/uploads/2016/04/Methodenbericht-zum-Integrationsbarometer-2016.pdf [accessed 2018-03-25]
37. Deutschses Jugendinstitut (DJI), München, Technische Universität Chemnitz-Zwickau. DJI-Ausländersurvey-Jugendliche (10/1996-02/1997). Cologne, Germany: GESIS Datenarchiv; 2001.
38. Hoebel J, Müters S, Kuntz B, Lange C, Lampert T. [Measuring subjective social status in health research with a German version of the MacArthur Scale]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015 Jul;58(7):749-757. [doi: 10.1007/s00103-015-2166-x] [Medline: 25986532]
39. Mayring P. Qualitative Inhaltsanalyse: Grundlagen und Techniken. Volume 12. Vollständig Überarbeitete und Aktualisierte Auflage. Weinheim, Germany: Julius Beltz; 2015.
40. Zeisler ML, Lemcke J, Bilgic L, Lampert T, Santos-Hövener C, Schmich P. Integration of migrant populations into health monitoring in Germany: results from a feasibility study. Survey Methods: Insights from the Field 2019 Mar 29 [FREE Full text] [doi: 10.2196/preprints.14747]
41. American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Revised 2016. Oakbrook Terrace, IL: AAPOR; 2016.
42. Bednaschewsky R, Supik L. Von Deutschen of Color und Deutschen mit Migrationshintergrund in der Statistik. In: Gomolla M, Kollender E, Menk M, editors. Rassismus und Rechtsextremismus in Deutschland. Figurationen und Interventionen in Gesellschaft und staatlichen Institutionen. Weinheim, Germany: Beltz Juventa; 2018:179-194.
43. Brücker H, Rother N, Schupp J. IAB-BAMF-SOEP-Befragung von Geflüchteten 2016: Studiendesign, Feldergebnisse sowie Analysen zu schulischer wie beruflicher Qualifikation, Sprachkenntnissen sowie kognitiven Potenzialen: Forschungsbericht 30. 2018 Mar 20. URL: http://doku.iab.de/forschungsbericht/2017/fb1317.pdf [accessed 2018-09-25]
44. Dillman DA, Smyth JD, Christian LM. Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method. Somerset, NJ: John Wiley & Sons; 2014.
45. Allen J, Schmich P. Gesundheit in Deutschland Aktuell (GEDA). In: Eifler S, Faulbaum F, editors. Methodische Probleme von Mixed-Mode-Ansätzen in Der Umfrageforschung. Wiesbaden, Germany: Springer; 2017:161-182.
46. Aschauer W. Besonderheiten und Problemlagen der quantitativen Befragung bei Migrantinnen. In: Weichbold M, Bacher J, Wolf C, editors. Umfrageforschung. Wiesbaden, Germany: Springer; 2009:293-315.
47. De Leeuw ED. To mix or not to mix data collection modes in surveys. J Off Stat 2005;21:233-255 [FREE Full text]
48. Couper MP. The future of modes of data Collection. Public Opin Q 2011 Dec 15;75(5):889-908. [doi: 10.1093/poq/nfr046]
49. Schilling R, Hoebel J, Müters S, Lange C. Pilotstudie zur Durchführung von Mixed-Mode-Gesundheitsbefragungen in der Erwachsenenbevölkerung (Projektstudie GEDA 2.0). Berlin, Germany: Robert Koch Institute; 2015.
50. Groves RM. Survey Errors and Survey Costs. New York, NY: John Wiley & Sons; 2004.
51. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
52. Brand T, Samkange-Zeeb F, Dragano N, Keil T, Krist L, Yesil-Jürgens R, et al. Participation of Turkish migrants in an epidemiological study: does the recruitment strategy affect the sample characteristics? J Immigr Minor Health 2019 Aug;21(4):811-819. [doi: 10.1007/s10903-018-0788-4] [Medline: 30027505]
53. Maehler DB, Martin S, Rammstedt B. Coverage of the migrant population in large-scale assessment surveys. Experiences from PIAAC in Germany. Large Scale Assess Educ 2017 Mar 1;5(1):9. [doi: 10.1186/s40536-017-0044-8]
54. Behr D, Braun M, Dorer B. GESIS Survey Guidelines: Measurement Instruments in Cross-National Surveys. Mannheim, Germany: GESIS -- Leibniz-Institut für Sozialwissenschaften; 2016. URL: https://www.gesis.org/fileadmin/upload/SDMwiki/BehrBraunDorer_Measurement_Instruments_in_Cross-National_Surveys.pdf

Abbreviations

AAPOR: American Association for Public Opinion Research
IMIRA: Improving Health Monitoring in Migrant Populations
KiGGS: Study of the German Health Interview and Examination Survey for Children and Adolescents
NC: noncontacts
NE: not eligible
O: other
PMB: persons with a migrant background
R: refusals
RKI: Robert Koch Institute
UE: unknown eligibility

Edited by G Eysenbach; submitted 17.05.19; peer-reviewed by V Vennu, A Karakosta; comments to author 03.10.19; revised version received 11.11.19; accepted 15.12.19; published 15.04.20.

Please cite as:
Zeisler ML, Bilgic L, Schumann M, Wengler A, Lemcke J, Gößwald A, Lampert T, Santos-Hövener C, Schmich P
Interventions to Increase the Reachability of Migrants in Germany With Health Interview Surveys: Mixed-Mode Feasibility Study
JMIR Form Res 2020;4(4):e14747
URL: https://formative.jmir.org/2020/4/e14747
doi:10.2196/14747
PMID:32305950

Viewpoint

The Postencounter Form System: Viewpoint on Efficient Data Collection Within Electronic Health Records

Philip Held[1], PhD; Randy A Boley[1], BA; Walter G Faig[2], PhD; John A O'Toole[3], MS; Imran Desai[1], MBA; Alyson K Zalta[4], PhD; Jawad Khan[3], BSc; Shannon Sims[5], MD, PhD; Michael B Brennan[1], PsyD; Rebecca Van Horn[1], MD; Angela C Glover[1], BA; Bala N Hota[6], MD, MPH; Brian D Patty[3], MD; S Shafiq Rab[3], MD, MPH; Mark H Pollack[1], MD; Niranjan S Karnik[1], MD, PhD

[1]Department of Psychiatry and Behavioral Sciences, Rush University Medical Center, Chicago, IL, United States

[2]Research Institute, Children's Hospital of Philadelphia, Philadelphia, PA, United States

[3]Department of Information Services, Rush University Medical Center, Chicago, IL, United States

[4]Department of Psychological Science, University of California, Irvine, Irvine, CA, United States

[5]Vizient, Inc, Chicago, IL, United States

[6]Department of Internal Medicine, Rush University Medical Center, Chicago, IL, United States

Corresponding Author:
Philip Held, PhD
Department of Psychiatry and Behavioral Sciences
Rush University Medical Center
325 S Paulina St
Suite 200
Chicago, IL, 60612
United States
Phone: 1 312 942 1423
Email: philip_held@rush.edu

Abstract

Electronic health records (EHRs) offer opportunities for research and improvements in patient care. However, challenges exist in using data from EHRs due to the volume of information existing within clinical notes, which can be labor intensive and costly to transform into usable data with existing strategies. This case report details the collaborative development and implementation of the postencounter form (PEF) system into the EHR at the Road Home Program at Rush University Medical Center in Chicago, IL to address these concerns with limited burden to clinical workflows. The PEF system proved to be an effective tool with over 98% of all clinical encounters including a completed PEF within 5 months of implementation. In addition, the system has generated over 325,188 unique, readily-accessible data points in under 4 years of use. The PEF system has since been deployed to other settings demonstrating that the system may have broader clinical utility.

(JMIR Form Res 2020;4(4):e17429) doi:10.2196/17429

KEYWORDS

electronic health record; data collection; veterans

Introduction

Background

The increasing use of electronic health records (EHR) across medical systems offers unique opportunities for research and quality improvement, and has the potential to significantly improve patient care [1]. Data captured in the EHR provides important insights about patients' characteristics, treatment needs, visit frequencies, and other types of information. The voluminous amount of information contained in EHRs also enables the development of advanced models to predict the course of diseases, examine treatment effectiveness, and determine treatment options [2,3]. Despite the wealth of information contained in EHRs, a major challenge exists when it comes to extracting usable data to advance the delivery of care and clinical operations. This is largely due to extensive amounts of text-based information existing within clinical progress notes, which can be labor intensive and costly to transform into usable data points [4,5].

Manual chart reviews and data coding or more advanced analytic techniques such as natural language processing (NLP) are strategies that have been employed to make text-based

information in EHRs more usable. However, reviewing charts and extracting data manually or training a machine learning model using NLP to identify the proper information requires significant time and are consequently associated with higher labor costs [6]. Thus, even if organizations desire to use their data to drive clinical or operational decisions, it is often not feasible without significant resources devoted to this process.

Even when the availability of resources is not a barrier and advanced NLP is utilized to extract text-based data from clinical progress notes, there are still challenges to obtaining usable data and developing accurate models [6]. For example, many clinical progress notes contain text and information that can be extremely similar or the same as a result of providers copying-and-pasting to save time [4,7,8]. As a result, data extracted from these notes can be highly skewed and may result in overfitted models with limited generalizability [6,7]. In other cases, content in clinical progress notes may significantly vary, especially with regard to behavioral health information [9]. Whereas some providers include detailed documentation of factors that may be likely to impact treatment outcomes, such as the type and duration of interventions provided, referrals to other providers, or changes in clients' symptoms over the course of therapy, other providers may only include the minimal required documentation into their progress notes. Overall, clinical progress notes are highly prone to documentation errors that can lead to missing or incorrect treatment information and the additional burden of extracting usable data. Consequently, it is critical to identify and use approaches that improve the data capture process without adding a significant burden to providers.

One strategy to improve the data capture process is by developing tailored flow sheets within EHRs that are specific to the needs of each clinic and type of encounter. These flow sheets can then be implemented at virtually all levels of patient interaction to allow for the collection of relevant, usable data whenever there is contact with the patient. For example, postencounter forms (PEFs) are flow sheets that can be attached to each clinical encounter and enable clinicians to quickly, accurately, and completely document relevant treatment information that can be easily extracted for analytic purposes. For behavioral health, this may include information about interventions used during sessions, minutes spent providing the interventions, clinical severity and progress, referrals provided, or termination status. The adaptability and convenience of PEFs allow organizations and providers to easily and accurately capture information relevant to their specific needs.

In this case report, we documented the development and implementation of the PEF system in the Road Home Program at Rush University Medical Center. Specifically, we detailed the process of developing the PEF content and system, demonstrated the types and volume of data that can be captured and easily extracted using this system, and described the implementation of the system.

Case Description

The PEF system was first developed and implemented at the Road Home Program within the Department of Psychiatry at Rush University Medical Center in Chicago, IL in response to a need to easily document and evaluate program use and effectiveness and be able to provide detailed program information to its funders. Later, this system was disseminated to other departments within the academic medical center.

Methods

Development of the Postencounter Form

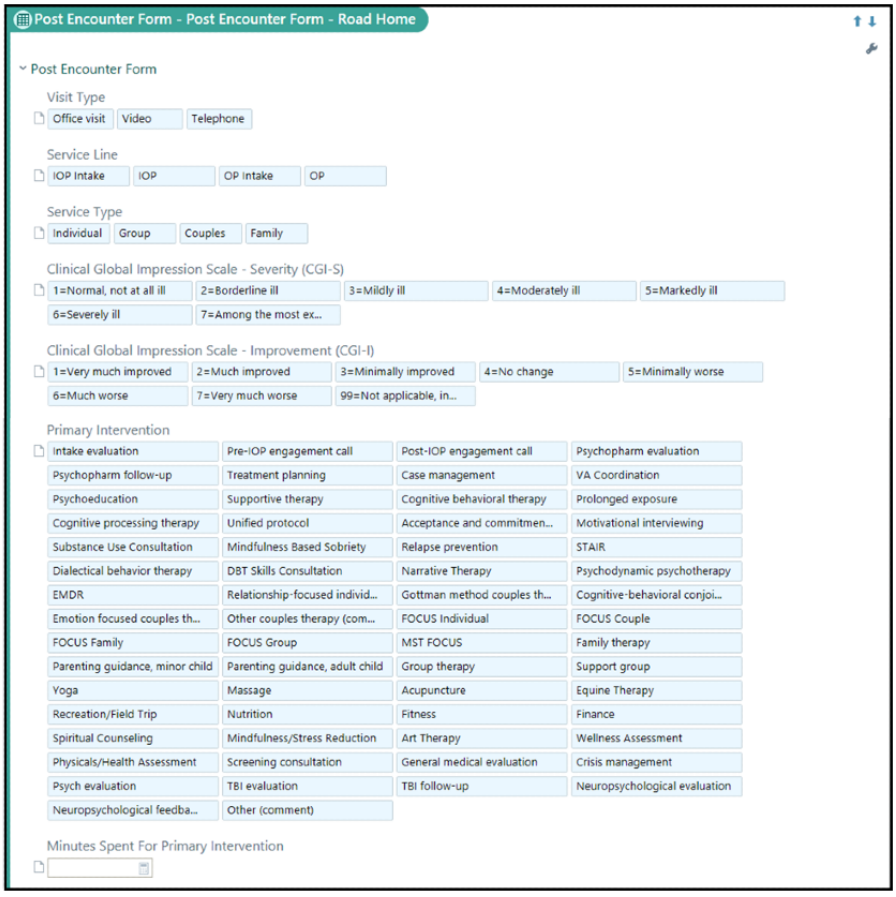
Members from various teams worked together during the PEF system development process to identify their unique data needs and note specific limitations. Teams included *clinical providers* who deliver services and were responsible for entering session data, members from the *knowledge management team* that combine data from multiple sources including the EHR and online survey tools, members from the *data team* who are responsible for cleaning and auditing the data to ensure its accuracy, and members from the *research team* who use data for descriptive and predictive analytics to generate insights that can help improve clinical care. The goal was to develop a comprehensive yet efficient system that enabled the collection of all relevant data in 90 seconds or less per patient encounter to prevent an excessive burden on clinicians who have limited time for documentation. Moreover, the PEF system needed to be *adaptable* since program needs would change as it expanded. It was critical for the system to be *easy-to-use* for all involved parties so that the tool could be quickly learned and errors could be minimized or easily detected and corrected. The PEF system also needed to be *accessible* and enable the immediate extraction of data for analytic purposes. In addition, as described earlier, it was important for the system to be *cost-effective* as the cost of implementing more advanced approaches involving NLP can be prohibitive. The clinical data that were determined to be important to capture in the PEF for the Road Home Program along with a brief description of their purposes are shown in Table 1.

Figure 1 shows the Road Home Program PEF. The specific fields of the PEF may differ based on a clinic's needs.

Table 1. Road Home Program postencounter form fields with descriptions

Postencounter form field	Description
Visit type	Specify whether encounters took place in person, over the phone, or through video
Service line	Specify whether service provided was an intake vs regular appointment and whether the service was part of the outpatient or 3-week intensive treatment program
Service type	Specify whether services were provided for individuals, groups, families, or couples
Clinical Global Impression Scale (severity)	Validated clinician-rated scale to indicate a patient's symptom severity level
Clinical Global Impression Scale (improvement)	Validated clinician-rated scale to indicate how much a patient has improved
Primary intervention	Specify the type of intervention delivered
Minutes spent for primary intervention	Specify the number of minutes for which the intervention was delivered
Secondary intervention	Specify the type of any secondary intervention delivered, when applicable
Minutes spent for secondary intervention	Specify the number of minutes for which the secondary intervention was delivered
Referral given	Specify the type of referral that was made during the session
Referral target	Specify for whom the referral was made
Referral reason	Specify the reason the referral was made
Termination	Specify whether termination took place during the session
Termination date	Specify the date of the termination
Leave reason	Specify the reason for leaving or terminating care

Figure 1. Sample Road Home Program post-encounter form. IOP: intensive out patient; OP: out patient; TBI: traumatic brain injury; VA: Department of Veterans Affairs; STAIR: Skills Training in Affective and Interpersonal Regulation; DBT: dialectical behavior therapy; EMDR: eye-movement desensitization and reprocessing; FOCUS: Families OverComing Under Stress; MST: Military Sexual Trauma.

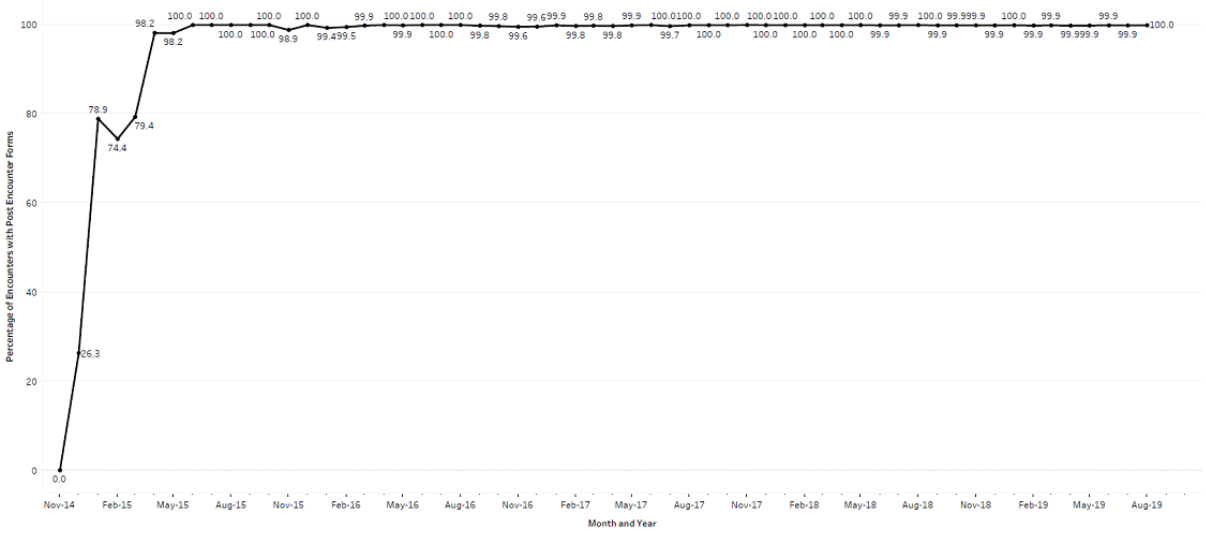


Implementation into the Road Home Program

Following the development of the PEF system, which involved evaluating various prototypes, the team implemented it into the Road Home Program clinical flow. All Road Home Program staff were required to attend a 1-hour training session where the purpose of the PEF was discussed. A demonstration was given on how to complete the form and questions about its use were answered. The difficulty of extracting data from clinical progress notes was specifically highlighted to help staff understand the importance of collecting data in easily extractable ways, even if this process added up to 90 seconds per encounter. Following the initial training, the team periodically met with select providers to ensure that everyone was aligned with regard to how the PEFs were completed, for example to ensure that providers would all code a cognitive behavioral intervention in the same way rather than using another, noncognitive behavioral intervention code. Several weeks following the training, the team met with the staff again to demonstrate the usefulness of PEFs by providing basic descriptive information about encounters that were already completed using PEFs. The descriptive information focused on the number of minutes spent delivering interventions such as cognitive behavioral therapy, and the types of referrals that had been made and tracked since the implementation of the PEF. The intention behind this process was to reinforce providers' use of the PEF. Clinicians quickly began to use the PEF following its implementation (Figure 2). In only 5 months, 98.2% (109/111) of all Road Home Program clinical encounters included a PEF. The rate of use has remained extremely high since the PEF system was implemented; on average, only 0.08% (260/323,026) of all encounters had a missing PEF since May 2015. Between December 2015, and August 2019, a total of 40,889 PEFs have been completed by Road Home Program providers, capturing more than 325,188 unique data points.

To facilitate the implementation, the team met with providers who tended to underutilize the PEF system to identify and resolve potential barriers. The biggest barrier to using the PEF was providers not being able to locate it in the EHR, which lead to the creation of a specific instructional guide on where to locate it. The team continued to meet with providers as needed when they encountered issues with locating or completing the form following the initial implementation and will continue to do so as needed. The team also continued to present descriptive and advanced analytical insights about the program to providers to further reinforce the utility and importance of the PEF system.

Figure 2. Percentage of Road Home Program encounters with post-encounter forms from November 2014 through September 2019.



Discussion

In this case study we described the development and implementation of the PEF system, which can be used to quickly and accurately capture clinically relevant data in a comprehensive yet efficient, adaptable, easy-to-use, accessible, and cost-effective way. Given the challenges associated with manual extraction of clinical progress note content or the use of more advanced and costly processes such as NLP, it was important to identify ways that clinical data could be easily captured and made immediately usable for analytic purposes. This study demonstrated the PEF system built in Epic; however, it can be built in to most modern EHR systems with minimal effort, as it is based on simple flow sheets that are attached to each encounter. Moreover, the PEF system is highly adaptable to the unique needs of any department or clinic. It is also flexible enough that its content can be changed as priorities or needs shift.

The implementation of the PEF system at the Road Home Program demonstrated the importance of collaboration between members and different teams. A common challenge to implementing new systems is that direction is given top-down with little input from individuals who are expected to use newly developed tools. In this case, members from a wide range of teams worked together to develop an end product that met the needs of everyone involved. Moreover, the development team worked closely with providers during the implementation phase and ensured that providers understood why the system was developed and the benefits of the system compared to existing practices (ie, clinical progress notes alone). The development

team then reinforced the use of the newly developed PEF system by repeatedly demonstrating to providers what data was being generated and how it can be used to inform clinical or operational decisions.

It is important to note that the use of flow sheets is not limited to PEFs, as flow sheets can be used to capture all types of data. Since the successful implementation of the PEF system, the Road Home Program has implemented this system for most other forms of data collection such as demographic information during an intake evaluation, medical information, or clinician-administered assessment data. Anecdotally, providers continue to remark about the ease of data collection. The data captured using the PEF system have been critical in providing data for the development of predictive models that have led to the improvement of clinical operations [10].

This case study has several limitations that should be noted. First, this case study focused on a single program within one academic medical center. Additional research needs to be conducted to determine whether the PEF is a viable option for other departments and medical systems. Second, the implementation of the PEF system was not randomized. Thus, it is impossible to determine in what ways the close collaboration between the development team and providers as well as the stepwise approach affected the ultimate uptake. Last, the PEF system was not directly compared to approaches using NLP. Thus, it cannot be determined whether the PEF system is more or less accurate, labor-intensive, or costly than NLP-based approaches.

Despite the aforementioned challenges, this case study highlighted how useful the PEF system can be for quickly and accurately capturing data without increasing provider burden or requiring significant time or funding to extract clinical data. Future research should closely examine the implementation process to determine the most effective way of rolling out new tools such as the PEF in a variety of health settings.

Acknowledgments

We thank the Wounded Warrior Project for their support of the Warrior Care Network and the resulting research. We would also like to thank the participating veterans and their families, as well as acknowledge the administrators, research assistants, and clinicians at the Road Home Program.

Author PH is supported by a career development award from the National Institutes of Health (5KL2TR002387-02) and receives grant support from the Wounded Warrior Project, and the Boeing Company. Author AZ is supported by a career development award from the National Institute of Mental Health (K23 MH103394). Author MP receives support from the Wounded Warrior Project and research funding from the National Institutes of Health and Janssen Pharmaceuticals; he provides consultation to Aptinyx, Clintara, and Palo Alto Health Sciences; he has equity in Argus, Doyen Medical, Mensante Corporation, Mindsite, and Targia Pharmaceuticals; he receives royalties from SIGH-A, SAFER interviews. Author NK is supported by grants from the Wounded Warrior Project and Welcome Back Veterans, an initiative of the Robert R. McCormick Foundation and Major League Baseball. NK is additionally supported by the National Center for Advancing Translational Sciences (UL1-TR002398, KL2-TR002387), the National Institute on Drug Abuse (R01-DA041071, UG1-DA049467), and the Cynthia Oudejans Harris, MD, Endowment Fund at Rush University Medical Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Wounded Warrior Project, or any other funding agency.

Conflicts of Interest

None declared.

References

1. Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. N Engl J Med 2010 Aug 05;363(6):501-504. [doi: 10.1056/NEJMp1006114] [Medline: 20647183]
2. Van Calster B, Wynants L, Timmerman D, Steyerberg EW, Collins GS. Predictive analytics in health care: how can we know it works? J Am Med Inform Assoc 2019 Dec 01;26(12):1651-1654 [FREE Full text] [doi: 10.1093/jamia/ocz130] [Medline: 31373357]
3. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. NPJ Digit Med 2018 May 8;1(1):18 [FREE Full text] [doi: 10.1038/s41746-018-0029-1] [Medline: 31304302]
4. Wang MD, Khanna R, Najafi N. Characterizing the source of text in electronic health record Progress notes. JAMA Intern Med 2017 Aug 01;177(8):1212-1213 [FREE Full text] [doi: 10.1001/jamainternmed.2017.1548] [Medline: 28558106]
5. Milinovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. Ann Transl Med 2018 Feb;6(3):42-42 [FREE Full text] [doi: 10.21037/atm.2018.01.13] [Medline: 29610734]
6. Liu F, Weng C, Yu H. Natural language processing, electronic health records, and clinical research. In: Clinical Research Informatics. London: Springer; Jan 23, 2012:293-310.
7. Gabriel RA, Kuo T, McAuley J, Hsu C. Identifying and characterizing highly similar notes in big clinical note datasets. J Biomed Inform 2018 Jun;82:63-69 [FREE Full text] [doi: 10.1016/j.jbi.2018.04.009] [Medline: 29679685]

8. Hammond KW, Helbig ST, Benson CC, Brathwaite-Sketoe BM. Are electronic medical records trustworthy? observations on copying, pasting and duplication. AMIA Annu Symp Proc 2003:269-273 [FREE Full text] [Medline: 14728176]
9. Hollister B, Bonham VL. Should electronic health record-derived social and behavioral data be used in precision medicine research? AMA J Ethics 2018 Sep 01;20(9):E873-E880 [FREE Full text] [doi: 10.1001/amajethics.2018.873] [Medline: 30242819]
10. HIMSS. 2019 Jan 30. Rush University Medical Center HIMSS Davies Enterprise Award URL: https://www.himss.org/library/rush-university-medical-center-himss-davies-enterprise-award [accessed 2019-10-14]

Abbreviations

EHR: electronic health record
NLP: natural language processing
PEF: postencounter form.

Edited by C Lovis; submitted 12.12.19; this is a non–peer-reviewed article;accepted 01.02.20; published 06.04.20.

Please cite as:
Held P, Boley RA, Faig WG, O'Toole JA, Desai I, Zalta AK, Khan J, Sims S, Brennan MB, Van Horn R, Glover AC, Hota BN, Patty BD, Rab SS, Pollack MH, Karnik NS
The Postencounter Form System: Viewpoint on Efficient Data Collection Within Electronic Health Records
JMIR Form Res 2020;4(4):e17429
URL: https://formative.jmir.org/2020/4/e17429
doi:10.2196/17429
PMID:32250276

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org