Original Paper

Usability and Usefulness of a Mobile Health App for Pregnancy-Related Work Advice: Mixed-Methods Approach

Monique van Beukering¹, MD; Adeline Velu¹, MD; Liesbeth van den Berg², MSc; Marjolein Kok¹, MD, PhD; Ben Willem Mol³, MD, PhD; Monique Frings-Dresen⁴, PhD; Robert de Leeuw⁵, MD; Joris van der Post¹, MD, PhD; Linda Peute^{2,6}, MSc, PhD

Corresponding Author:

Monique van Beukering, MD Department of Obstetrics and Gynecology Amsterdam UMC University of Amsterdam Room H4-222 PO Box 22660 Amsterdam, 1100 DD Netherlands

Phone: 31 657514795 Fax: 31 (0) 206963489

Email: m.d.vanbeukering@amc.uva.nl

Eman. <u>m.c. van</u>

Abstract

Background: Pregnant women are often unaware of the potential risks that working conditions can cause to them and their unborn child. A mobile health (mHealth) app, the *Pregnancy and Work* (P and W) app, developed by a multidisciplinary team and based on an evidence-based guideline for occupational physicians, aims to provide advice on work adjustment during pregnancy.

Objective: This study evaluates the usability of the mHealth P and W app and the perceived usefulness of the *work advice*, the main goal of the app, by potential end users.

Methods: A total of 12 working pregnant women participated in think aloud usability sessions and performed 9 tasks. All think aloud sessions were recorded, transcribed, and coanalyzed. The usability problems were rated for their severity in accordance with Nielsen severity scale. The completion rates and time taken for completion of tasks were registered. In addition, participants were questioned on demographics and user characteristics and were asked to evaluate the value of the app by filling in the Intrinsic Motivation Inventory (IMI) score and the System Usability Scale (SUS) questionnaire.

Results: In total, 82 usability problems with a severity ≥ 1 were identified, of which 40 had severity ≥ 3 . The main usability problems concerned the interpretation of terminology used in the app's questionnaires and difficulties in finding and understanding the work advice. Furthermore, 10 out of 12 participants were able to open the work advice page in the app. Only 7 out of these 10 participants understood and intended to follow the work advice. The overall mean IMI score was relatively high (5 out of 7), indicating that the participants did indeed value the use of the app. This IMI score corresponded to the overall mean SUS score (68 out of 100) and the mean grade given to the P and W app (7 out of 10).

Conclusions: This think aloud usability study showed that the information provided in the P and W app was considered valuable by the end users, working pregnant women, and it meets their needs; however, usability issues severely impacted the perceived usefulness of the work advice given in the app.



¹ Department of Obstetrics and Gynecology, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

²Department of Medical Informatics, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

³Department of Obstetrics and Gynaecology, Monash Medical Centre, Monash University, Melbourne, Australia

⁴Coronel Institute of Occupational Health, Amsterdam Public Health Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

⁵Department of Obstetrics and Gynecology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

⁶Department of Medical Informatics, Center for Human Factors Engineering of Health Information Technology, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

(JMIR Mhealth Uhealth 2019;7(5):e11442) doi: 10.2196/11442

KEYWORDS

mHealth; eHealth; mobile phone; pregnancy; work; occupation; occupational exposure; qualitative research

Introduction

Background

Many women continue to work during their pregnancy. In the United States, more than 65% of pregnant women work, whereas in the Netherlands, around 80% [1,2]. During pregnancy, exposure to certain working conditions, such as physically demanding work, long working hours, working in night shifts, and stress, are associated with preterm birth, low birth weight, and fetal abnormalities [3-12]. As pregnant women are often not aware of these risks, they do not adjust their working conditions [13].

Mobile health (mHealth) apps can offer a suitable solution to this problem as women of reproductive age who are expecting a child are frequent consumers of Web-based health information [14-17]. mHealth, defined as the use of mobile devices for medical and public health practice [18], could therefore inform pregnant working women about work-related pregnancy risks, to increase their awareness of these risks and their associated need for change in working conditions.

However, evidence on the effectiveness of mHealth apps in general is limited [17,19]. Prior studies provide little information as to how best to design them [20-24]. Adequate consideration of the needs of their intended users is necessary so that they are easy to use and perceived as useful [25,26]. The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use is the definition for applied usability, based on the International Standardization Organization [27]. To assess and improve upon the usability of mHealth apps, a wide range of usability evaluation methods (UEMs) is available to detect problems in the user-system interaction. UEMs thus assess human interaction with a system for the purpose of identifying those facets of this interaction which can be improved [28]. Ideally, the design process of any health-related product is conducted in an iterative fashion to better fit with the end user population. Utilizing UEMs in such an iterative design process in the health care domain is especially important as the poor design and usability of medical products can lead to harmful consequences [29,30]. Therefore, the utilization of UEMs during the development and testing process of health apps is widely recommended throughout research [31,32].

In this study, we developed an mHealth solution (the Pregnancy and Work [P and W]) app that aimed to provide information and advice about work-related pregnancy risks [33]. With this advice, pregnant women can adjust their work. The P and W app content is based on the evidence-based guideline for occupational physicians, *pregnancy*, *postpartum period*, *and work* [34]. In a prior study, the results of 2 multidisciplinary

focus group meetings provided content and design instructions for the development of the P and W app [35].

Objectives

Think aloud (TA), an UEM method, was chosen in this study to assess the usability of the P and W app with potential end users to reveal cognitive processes in the app's user interaction that result in user-interaction problems. The TA method requires participants to talk aloud (ie, verbalize their thoughts) while performing or solving a task to reveal their cognitive processes while interacting with the app, which may result in user-interaction problems [36-38]. In this way, the TA helps to understand how pregnant woman think—or believe they think—the P and W app works (ie, their mental model) [38]. Mismatches in the end users' mental model of an app and the app's design can severely influence its usability and subsequently its use in practice. This study therefore evaluated the usability of the P and W app and also how potential end users experienced the usefulness of the work advice; this was the main goal of the app.

Methods

Participant Recruitment

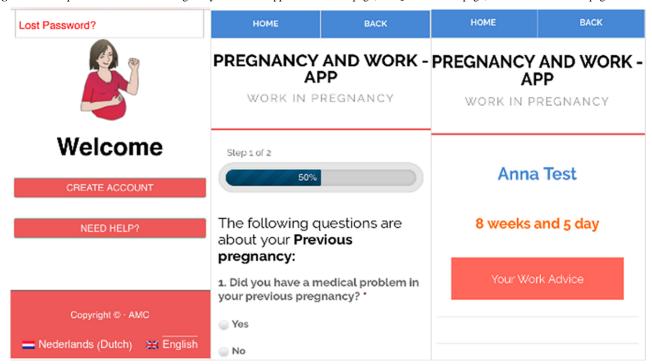
A total of 2 obstetric care facilities, representing a broad variety of patient groups, participated in this study. Posters and flyers were distributed in both locations. The inclusion criteria were drawn up by an obstetrician and occupational physician. If patients met the inclusion criteria, they were invited to participate in the study. The inclusion criteria were Dutch working women, who were less than 20 weeks pregnant. The criterion of being less than 20 weeks pregnant was deliberately stated as the work advice for pregnant women under 20 weeks of pregnancy can be different than that for those after 20 weeks of pregnancy. Eligible participants were recruited in the waiting area of the physician's office. Recruitment of participants continued until a total of 12 female patients agreed to participate in the TA sessions and evaluate the app; this was the first time they used the P and W app. All participants included in the study were offered a gift card worth €15. An app for this research was submitted to the ethical board of the Amsterdam University Medical Center. The board confirmed that the Medical Research Involving Human Subjects Act did not apply to this study. All data from the 12 participants were anonymously processed. Informed consent was obtained from all participants, allowing us to use the data for analysis.

P and W Pregnancy and Work App and Study Flow

The P and W app (Dutch and English) was created as a Web-based app, accessible from every type of mobile browser, with an adaptive design for desktop and mobile phone use.



Figure 1. Examples of screenshots of Pregnancy and Work App: the Welcome page, the Questionnaire page, and the Work Advice page.



The P and W app requires the user to create an account to gain access to its content. After creating an account, a user needs to complete a questionnaire about her pregnancy-related medical and work conditions (Figure 1). When completing this questionnaire, the user will be directed to the home page of the app, from where she can navigate to all other pages. On the home page, users can view monthly pregnancy- and work-related advice messages, which are also sent by email. In addition, the app provides messages about the growth of the unborn baby as the weeks pass. Next to the baby messages, a video with tips and information about pregnancy-related work advice can be viewed on the home page. Participants were given access to a Dutch beta test version of the P and W app.

Phase I: Preparation

Participants were informed about how the TA session would be performed; see Figure 2 for the full study setup. After a 2-week reflection period, a condition for participation in the research, an appointment was made with those women who wanted to participate in the study. The TA session then took

Figure 2. Overview of study setup.

place at their next visit (follow-up consultation) to the obstetrics department. After signing an informed consent form, the participant completed a short survey, the validated health literacy (HL) assessment tool— *the Newest Vital Sign*, translated to Dutch—to analyze its potential influence on the TA outcomes (Stage I, Multimedia Appendix 1 [39,40].

Phase II: Think Aloud Usability Testing

Participants started with practice tasks on how to *think aloud* (Multimedia Appendix 2). Each participant was informed that the researcher (LvdB) was solely interested in the app's performance and would only interrupt the participant to provide new tasks and to encourage her to keep talking to break silences longer than 5 seconds [41]. A participant had to complete 9 tasks in total that were centered around the core purpose of the app (Multimedia Appendix 3). Tasks were developed in collaboration with the developer and project supervisors of the P and W project. All TA sessions were recorded via video camera. Voice and screen (of their mobile phone) were also recorded (Figure 3).

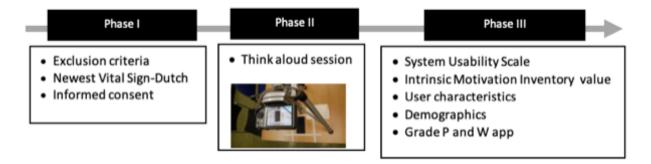




Figure 3. Think aloud session set-up.





Phase III: Usability and Motivation Questionnaires

After the TA test was finished, the SUS survey was given to the participant to assess the perceived usability of the P and W app [42] (Multimedia Appendix 4). The SUS comprises 10 statements which the participant had to rate on a scale from 1 (strongly disagree) to 5 (strongly agree) to indicate the extent to which she agreed. Then, a short survey selection of the Intrinsic Motivation Inventory (IMI) was given to assess a self-reported evaluation on how much the participant valued the P and W app [43] (Multimedia Appendix 4). The IMI value subscale comprised 7 statements which the participant had to rate on a scale from 1 (do not agree) to 7 (strongly agree) to indicate the extent to which she agreed. An additional short survey was developed to gain more insight into participants' demographics, medical history related to pregnancy, prior experience with (pregnancy-related) mobile apps, and working hours (Multimedia Appendix 4). We asked all participants whether they had received and would follow the work advice (Multimedia Appendix 4). Finally, participants were asked to give the P and W app a grade on a scale from 1 to 10, where 1 was the lowest and 10 the highest grade.

Data Collection and Analysis

The TA sessions were videotaped, reviewed multiple times, and transcribed to verbal protocols by 2 researchers (LvdB and LP). To gain insight into the effectiveness and efficiency of the participants in performing tasks, each TA session transcription comprised text spoken by the participant and included task

completion time stamps and time taken for task completion. To analyze the usability problems participants encountered in more detail, we performed a thematic analysis for which a coding scheme was developed bottom-up in 3 iterative cycles as described by Jaspers [44]. We analyzed 2 TA sessions in-depth to develop a raw coding scheme (first cycle). Usability issues encountered by participants were then given a specific description. We subsequently discussed the resulting codes and grouped them to determine the main themes in the data (second cycle). The developed coding scheme was then applied to code and analyze all verbalizations, this was performed by LvdB and checked by LP. All new issues were discussed to determine whether they were within the branches of the coding tree or if a new main theme had emerged. Usability problems were rated on severity in accordance with Nielsen severity scale [45]. Nielsen severity scale is a rating scale from 0 to 4 (Textbox 1), that allows for the prioritization of usability problems that need to be revised in the development process. The questionnaires were completed on paper and put in a database for data analysis.

All data filled in by the participants in the P and W app during the TA sessions were specifically transcribed into a different file to test for task efficacy in relation to the IMI-given work advice by the system. Verbalizations of task 6 in the TA sessions (find the work advice) were assessed to analyze whether participants would follow the work advice. These results were compared with the results of the IMI on participant level and the questions about the work advice from questionnaire 3 (Multimedia Appendix 4). Finally, the SUS was used to assess the perceived usability of the P and W app.

Textbox 1. Nielsen's severity scale

- 0-I do not agree that this is a usability problem at all.
- 1-Cosmetic problem only: need not to be fixed unless extra time is available on project.
- 2-Minor usability problem: fixing this should be given low priority.
- 3-Major usability problem: important to fix, so should be given high priority.
- 4-Usability catastrophe: imperative to fix this before product can be released.

Results

Participant Characteristics

The TA sessions with the participants (N=12) took place between April and June 2017. Most participants scored high (=adequate) HL. All participants had paid jobs and used a mobile

phone. The average gestational age of the participants was 15 weeks and 50% (6/12) of the participants were pregnant for the first time (Table 1).

Task Completion

The effectiveness and efficiency of the participants in performing tasks were measured by completion rates and times



and the usability problems. The completion rates and times can be found in Table 2. The average duration of a TA session was 19 min 55 seconds (SD 5 min 25 seconds). Task 1, *create an account*, had a much higher completion time than the other tasks. Tasks 2, 3, 5, and 9 were completed by all participants. Tasks 1, 4, 6, 7, and 8 were not completed by all participants. The first 3 tasks took, on average, the longest time to complete, ranging from 1.5 min to 4 min. Task 9 had the fastest mean completion rate of 4 seconds.

Usability Problems

The TA study identified a total of 101 usability issues, 82 of which were considered *real* usability problems (ie, severity \geq 1),

whereas 40 usability problems were rated with a severity of 3 (major) or 4 (catastrophic). In addition, the participants encountered 11 unique bugs when using the P and W app. An overview of the most severe usability problems can be found in Table 3. The high completion time with *create an account* (Table 1) seemed to have a connection with the many usability problems in this area (Table 3). None of the participants experienced (severe) usability problems when completing tasks 5, 7, and 9. In the following section, we give an in-depth analysis of the severe usability problems detected regarding *terminology interpretation* and *finding and understanding the work advice* that directly impacted the participants' perceived usefulness of the advice given in the app.

Table 1. Participants' basic demographics and characteristics (N=12).

Characteristics	Statistics	
Age (years), mean (SD)	33 (3.8)	
Education (secondary school), n		
Higher education	8	
Intermediate vocational education	4	
Health literacy, n		
High	11	
Low	1	
Paid job, n	12	
Working time (hours per week), mean (SD)	37 (6.15)	
Gestational age (weeks), mean (SD)	15 (3)	
Previous pregnancy, n	6	
Children, n	5	
Mobile phone (operating system), n		
Android	7	
iPhone	5	

Table 2. Completion rates and time taken per task (N=9) by participants.

Task	Completion rate	Time taken for completion (seconds), mean (SD)
1. Create an account	10/12	240 (83)
2. Fill in a questionnaire	12/12	179 (101)
3. Adjust answers to the questionnaire	12/12	96 (74) ^a
4. Find your rights and tips for consultation page	11/12	31 (38)
5. Find baby message(s)	12/12	16 (10)
6. Find the your work advice page	10/12	10 (8)
7. Find the <i>print/save</i> button	10/12	9 (9)
8. Find the goal of the Pregnancy and Work P and W app	11/12	32 (18)
9. Log out of the app	12/12	4 (4)

^aA total of 2 participants initially did not understand this task.



Table 3. Overview of severe usability problems per main problem type.

Usability problem ^a	Frequency	Severity	Source of main problem
Unclear buttons	12	2 to 4	Create account
Functionality with layout	11	4	Create account/home page
Terminology interpretation problems	8	4	Create account/home page
Finding and understanding work advice	8	4	Home page/work advice

^aMultimedia Appendix 5 shows an overview of all the usability problems.

Qualitative Assessment

Terminology Interpretation Problems

Participants had to complete a questionnaire about their pregnancy-related medical conditions, previous pregnancy (if relevant), and work conditions using the app. Several terminology interpretation problems arose during the TA study, which consequently prevented the participants from receiving accurate personal work advice. For example, when asked whether problems had been experienced during the previous pregnancy, participants were unsure whether *previous pregnancy* implied the immediate previous pregnancy or also the pregnancies before that. One participant who had not experienced problems during her previous pregnancy, but did experience issues during the pregnancy before that, assumed it implied her direct previous pregnancy. Her confusion in answering the question correctly affected the outcome of the work advice, as relevant information was missing:

Okay. Um. "Did you have a medical problem in your previous pregnancy?" This is about my last pregnancy, I think, and not the pregnancies before. So, I'm assuming that. And then it's a no. [Participant 5]

Problems were also prevalent when, in closed-ended questions, the participant did not find the answer that applied to her within the limited selection of possibilities of medical disorders. When given a list of potential problems during a previous pregnancy, participants experienced troubles in selecting the best suited option to describe their problem:

...But I do not know if that should be put under "deceased child" or "child born before a gestational age of 37 weeks"? You know what I mean? [Participant 3]

Another example of a terminology interpretation problem that affected the outcome of the work advice was related to the question of *being exposed to any chemical agents in the work environment*, followed by a list of examples. Several participants did not notice the list of examples and answered *no*. Furthermore, 2 other participants did not know whether an agent that they worked with should be considered chemical, as it was not on the list of examples:

... Yes, with hair dye. Is that chemical? [Participant 9]

...I'm having doubts. I work with laughing gas. That's not very chemical, but...I don't know whether I should answer yes or no. [Participant 11]

Finding and Understanding the Work Advice

Participants also experienced problems in understanding the work advice because of central design problems in the interface. One of the first issues encountered was that the participants expected the app to show them something different than what it actually did. Participants expected the app to show their work advice directly on the homepage, as they perceived this to be the essential goal of the app. They did not expect to have to search for it in the interface or take any other action to find it. For example, participant 6 did not understand that the *your work advice* button was clickable and therefore sought work advice elsewhere or stated that she could not find it (Figure 4):

...Oh, let's see if that is somewhere. No idea. [Scrolls down and up] Have a look. Here is my work advice. Uh... [Scrolls up and down, multiple times] No, I have no idea. [Participant 6]

A different example related to the participants stating that they saw their work advice depicted on the home page. However, the home page only provided a small section with tips and information about pregnancy-related work advice, which some clearly interpreted as the entire personal work advice. A total of 2 participants thought this was the case; therefore, both of them missed the actual content of the *your work advice* page:

I've just seen my work advice. [Scrolls up and down. Scrolls to top of the page. Taps the back button. Loads page] Yes, your work advice. I have already read it. So, it is here. [Participant 8]

Another usability issue was related to the fixed structure in which the work advice was presented in the mobile interface. Depending on the answers given in the questionnaire, specific information followed on the work advice page. The resulting advice therefore included some sections without advice and some sections with the advice, spread over the mobile interface. One participant did not get work advice below the *work header*; however, she did receive work advice with regard to issues during her previous pregnancy, but this would only have become visible if she had scrolled the page down. She therefore missed the advice given:

None? That's easy. I don't need to make any work adjustments. I don't think so either, because I have an office job. [Participant 1]



Figure 4. The "Your Work Advice" button on the home page with examples of work advice (infectious diseases and stress) when the button is clicked.

Anna Test

18 weeks and 5 day

Your Work Advice

If you are pregnant you are entitled to extra breaks, a total of 1 / 8th of your working time. You can just relax in between or later start or earlier home. Use this if it is really like, you're very tired.

Want to know more about your rights

as a pregnant employee? Get the

app from the FNV.

Infectious Diseases
Infectious Diseases You
have indicated that in yo

Infectious Diseases You have indicated that in your work you can get an infection. Infections caused by viruses, bacteria, parasites or fungi. They are transmitted in a variety of ways. For example, transmission by animals or sick people. Some



Stress

Stress

You're more likely to stress by work. If you are pregnant and have stress because of work you yourself more likely to have high blood pressure, your child may be born with less weight. This stress therefore try to reduce the pregnancy as early as possible.

User Evaluation: Intrinsic Motivation Inventory and System Usability Scale

The task efficacy of task 6, find the "your work advice" page, was analyzed in relation to the detected usability problems in finding and understanding the work advice and combined with the results of the IMI, SUS, and questions about the work advice from questionnaire 3 (Multimedia Appendix 4). Some participants never reached the work advice page on the app (17%) but thought they did, whereas 3 out of 12 participants (25%) were convinced that they had not received this advice

(Table 4). However, all participants did actually receive some form of pregnancy-related work advice. Among the 9 participants who stated that they had received work advice, 2 indicated that they would not follow it.

Using the IMI, we assessed the self-reported evaluation of how much the participants valued the P and W app; the overall mean IMI value score was 5 (SD 0.9) out of 7. The perceived usability of the P and W app was stated by the SUS. The overall mean SUS was 68 (SD 11). Finally, the participants were tasked to give the P and W app a grade on a scale from 1 to 10; the mean grade given to the P and W app was a 7 (SD 0.89; Table 4).



Table 4. User evaluation based on the use of work advice, Intrinsic Motivation Inventory (IMI), System Usability Scale (SUS), and grade.

Participant number	Did you receive work advice from the app? ^a	If so, do you intend to do something with this work advice? ^b	IMI ^c	SUS ^d	Grade
1	No ^e	N/A ^f	5.57	85	8
2	No	N/A	4.29	77.5	7
3	Yes	Yes	3.71	55	5
4	Yes	Yes	5.00	77.5	7
5	Yes	Yes	5.14	65	8
6	Yes	No	4.43	77.5	6
7	Yes	Yes	5.57	57.5	7
8	Yes	No	3.00	70	7
9	No	N/A	4.29	75	7
10	Yes	Yes	6.29	55	6
11	Yes	Yes	5.29	50	6
12	Yes	Yes	4.86	72.5	6

^aMultimedia Appendix 4-III Questionnaire 3, Question 1.

Discussion

Principal Findings

The overall effectiveness and efficiency of the 12 participants in performing tasks in the TA sessions are gauged by the completion times and rates and the usability problems. The TA study identified 82 usability problems with a severity ≥ 1 , of which 40 had severity ≥ 3 . The high completion time of the task to create an account seemed to be connected to the many usability problems that participants experienced in this task. As creating an account in an mHealth app is not usually part of the core, there is a chance that the design of this first part of the app may be neglected. Design errors in creating an account, however, increase the risk of participants dropping out quickly.

We performed an in-depth analysis of the severe usability problems detected regarding terminology interpretation and finding and understanding the work advice as these issues directly impacted the usefulness of the app. As participants were unable to correctly interpret the terminology in the questionnaire about previous pregnancies, medical disorders, and chemical agents, they did not understand how to complete the questionnaires corresponding to their personal situation. They thus did not receive the correct personal work advice for their circumstances.

Participants also had a different expectation of what the app would show them. Their mental model, the way information is represented in the mind of the end user, affected how they acted in the system in filtering the relevant information. The mental model of the participants did not match how the designer developed the system, as the designer had based it on his own mental model of how future end users would act on the information presented. The mental model of end users, which encompasses values, beliefs, and knowledge, creates perspectives for filtering information and guiding problem solving [46] and has the ability to affect how a person acts [47], differed from that of the designers. The users therefore also experienced problems with understanding the work advice, as their expectations did not match how the designer developed the system (based on his mental model of how future end users should act on information).

Due to the usability problems in its design, 10 out of 12 participants were able to open the work advice page. Only 7 out of these 10 participants understood and intended to follow the work advice given in the app, which was the main goal of the app.

The overall mean IMI score was relatively high (5 out of 7), indicating that the participants did indeed value the use of the app. This corresponded to the overall mean SUS score (68 out of 100) and the mean grade given to the P and W app (7 out of 10).

Comparison With Prior Work

Our main results indicated the effect of the app's navigational structure and screen design on the ability of a specific group of participants—pregnant working women—to find work advice and their intention to follow it thereafter. Other studies in mHealth and electronic health that have applied the TA method have demonstrated that although participants *think* that they have achieved the main goal of using the apps, in reality its intended objective was not reached [48,49]. In one study the researchers observed that the majority of participants, older



^bMultimedia Appendix 4-III Questionnaire 3, Question 2.

^cIMI score; 1=not at all true to 7=very true.

^dSUS score; 1=strongly disagree to 5=strongly agree.

^eParticipants 1, 2, and 9 were convinced that they had not received work advice; however, all participants did receive work advice.

^fN/A: not applicable as the participant indicated that she did not receive work advice.

cancer patients, were not able to find the requested information although the participants themselves frequently commented during testing that it was easy for them to find it [48]. In a different study, patients with rheumatic diseases were enthusiastic about the possibilities of interactive apps such as peer support forums and online consultations; however, nearly all participants experienced difficulties and were not able to complete all the usability evaluation tasks while interacting with the system [49].

As in our study, other researchers and designers have underlined the importance of an iterative approach in designing mHealth apps to understand the needs of end users as well as improve app usability and feasibility [36,50]. The importance of performing usability studies on mHealth apps to be used in a clinical and patient setting therefore needs serious attention. User testing is an essential part of developing mHealth apps, especially when aiming to effectively change actual patient behavior and/or affect patient outcomes.

Strengths and Limitations

A limitation is that the TA sessions took place in a laboratory setting. In their own home, participants may have taken more time to take a look at the app again. One of the strengths of this study is that the sample size is adequate for obtaining usability problems and that we used a mixed-methods approach— we combined the results of a TA test with the results of questionnaires on demographics, user characteristics, SUS, perceived value (IMI), and evaluation of the app. Another strength of our study is that it was performed by a multidisciplinary team and that the TA study is part of a process in developing an mHealth app, which started with 2 multidisciplinary focus group meetings [29].

Due to a lack of variety in HL levels, we were unable to analyze its potential influence on the TA outcomes. However, the recruitment of only 1 out of 12 participants with limited HL is in line with the estimations of HL prevalence levels in the Netherlands [51]; this certainly applies to a working population.

It is possible that the intention to follow the work advice could change according to the end user's job. However, as a significant proportion of the participants was not able to open the work advice page in the app, and/or understand the work advice or intend to follow it, we think that the influence of profession is

limited in this study. For the next study, we would advise asking participants about their job.

To human factor specialists, it is well known that end users should be involved from the beginning when developing an mHealth app. However, those who are well informed about a particular health domain, but less so about medical informatics, should be aware that an iterative multidisciplinary approach with the involvement of the target group from the start by using UEM research in the project is essential and can be very valuable.

The mixed-methods approach provides an insight into the cognitive process of a specific user group—pregnant working women—and their intention to use the P and W app. The TA results, in combination with the questionnaires on the perceived usability and value and the evaluation of the app, showed that incorrect interpretation of terminologies in the system prevented the end users from receiving the correct work advice. They also experienced problems with understanding the work advice because of central design problems in the interface. Despite many usability problems, the participants were relatively positive about the P and W app; the information provided in the app is considered valuable to the end users and meets their needs. The usability findings of this research could then be used to drive recommendations for developers for the next iteration of the P and W app aimed at pregnant working women.

Conclusions

The overall conclusion of this study is that the information provided in the P and W app was considered valuable to the end users, working pregnant women, and meets their needs; however, the usability issues severely impacted the perceived usefulness of the work advice given in the app. The results of this study draw attention to the relation between effective health apps and how their design might hamper their effectiveness in changing patients' behavior. An iterative UEM multidisciplinary approach, with the involvement of the target group from the beginning, is therefore essential for the development of health apps.

The mHealth app will be redesigned and tested in an intervention study, a survey on the effect of the app on actual work adjustment by pregnant women. A future version of the P and W app will be a valuable tool for informing pregnant women about pregnancy-related work risks.

Acknowledgments

The authors would like to thank all the pregnant participants who participated in the TA sessions and the employees of the obstetric care facilities.

This pilot study received funding from ZonMw, the Netherlands Organization for Health Research and Development. This project is part of the Pregnancy and Birth Program.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire before think aloud sessions & NVS-D.



[PDF File (Adobe PDF File), 293KB-Multimedia Appendix 1]

Multimedia Appendix 2

Think aloud session protocol.

[PDF File (Adobe PDF File), 342KB-Multimedia Appendix 2]

Multimedia Appendix 3

Participant tasks during the think aloud sessions: description, achievement and inclusion motivation.

[PDF File (Adobe PDF File), 116KB-Multimedia Appendix 3]

Multimedia Appendix 4

Questionnaires after think aloud session: SUS, IMI, and questionnaire 3.

[PDF File (Adobe PDF File), 398KB-Multimedia Appendix 4]

Multimedia Appendix 5

Overview of all usability problems and bugs.

[PDF File (Adobe PDF File), 221KB-Multimedia Appendix 5]

References

- 1. Gao G, Livingston G. Pew Research Center. Working while pregnant is much more common than it used to be URL: http://www.pewresearch.org/fact-tank/2015/03/31/working-while-pregnant-is-much-more-common-than-it-used-to-be/ [accessed 2018-04-06] [WebCite Cache ID 6yTgtLQda]
- 2. Central Statistics Office, the Netherlands. [Labor participation by age and gender] URL: https://www.cbs.nl/nl-nl/achtergrond/2017/07/arbeidsparticipatie-naar-leeftijd-en-geslacht [accessed 2017-11-30] [WebCite Cache ID 6vMHGCxMk]
- 3. Bonde JP, Jørgensen KT, Bonzini M, Palmer KT. Miscarriage and occupational activity: a systematic review and meta-analysis regarding shift work, working hours, lifting, standing, and physical workload. Scand J Work Environ Health 2013 Jul;39(4):325-334 [FREE Full text] [doi: 10.5271/sjweh.3337] [Medline: 23235838]
- 4. Vrijkotte TG, van der Wal MF, van Eijsden M, Bonsel GJ. First-trimester working conditions and birthweight: a prospective cohort study. Am J Public Health 2009 Aug;99(8):1409-1416. [doi: 10.2105/AJPH.2008.138412] [Medline: 19542045]
- 5. van Beukering MD, van Melick MJ, Mol BW, Frings-Dresen MH, Hulshof CT. Physically demanding work and preterm delivery: a systematic review and meta-analysis. Int Arch Occup Environ Health 2014 Nov;87(8):809-834. [doi: 10.1007/s00420-013-0924-3] [Medline: 24390632]
- 6. Palmer KT, Bonzini M, Harris EC, Linaker C, Bonde JP. Work activities and risk of prematurity, low birth weight and pre-eclampsia: an updated review with meta-analysis. Occup Environ Med 2013 Apr;70(4):213-222 [FREE Full text] [doi: 10.1136/oemed-2012-101032] [Medline: 23343859]
- 7. Croteau A, Marcoux S, Brisson C. Work activity in pregnancy, preventive measures, and the risk of delivering a small-for-gestational-age infant. Am J Public Health 2006 May;96(5):846-855. [doi: 10.2105/AJPH.2004.058552] [Medline: 16571706]
- 8. Croteau AA, Marcoux SS, Brisson C. Work activity in pregnancy, preventive measures, and the risk of preterm delivery. Am J Epidemiol 2007 Oct 15;166(8):951-965. [doi: 10.1093/aje/kwm171] [Medline: 17652310]
- 9. Juhl M, Larsen PS, Andersen PK, Svendsen SW, Bonde JP, Nybo AA, et al. Occupational lifting during pregnancy and child's birth size in a large cohort study. Scand J Work Environ Health 2014 Jul;40(4):411-419 [FREE Full text] [doi: 10.5271/sjweh.3422] [Medline: 24647563]
- 10. Lee BE, Ha M, Park H, Hong Y, Kim Y, Kim YJ, et al. Psychosocial work stress during pregnancy and birthweight. Paediatr Perinat Epidemiol 2011 May;25(3):246-254. [doi: 10.1111/j.1365-3016.2010.01177.x] [Medline: 21470264]
- 11. Loomans EM, van Dijk DA, Vrijkotte TG, van Eijsden M, Stronks K, Gemke RJ, et al. Psychosocial stress during pregnancy is related to adverse birth outcomes: results from a large multi-ethnic community-based birth cohort. Eur J Public Health 2013 Jun;23(3):485-491. [doi: 10.1093/eurpub/cks097] [Medline: 22850187]
- 12. Mocevic E, Svendsen SW, Jørgensen KT, Frost P, Bonde JP. Occupational lifting, fetal death and preterm birth: findings from the Danish National Birth Cohort using a job exposure matrix. PLoS One 2014;9(3):e90550 [FREE Full text] [doi: 10.1371/journal.pone.0090550] [Medline: 24614129]
- 13. Hooftman WE, van den Bossche SJN. Original Ministry Social Affairs and Employment. 2007. Pregnancy and Work: counseling, measures and sick leave URL: http://www.monitorarbeid.tno.nl/dynamics/modules/SPUB0102/view. http://www.monitorarbeid.tno.nl/dynamics/modules/SPUB0102/view. http://www.monitorarbeid.tno.nl/dynamics/modules/SPUB0102/view. http://www.monitorarbeid.tno.nl/dynamics/modules/SPUB0102/view.



- 14. Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth Platform for Pregnancy on Nutrition and Lifestyle of the Reproductive Population: A Survey. JMIR Mhealth Uhealth 2016 May 27;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
- 15. Abroms LC, Johnson PR, Heminger CL, Van Alstyne J, Leavitt LE, Schindler-Ruwisch JM, et al. Quit4baby: results from a pilot test of a mobile smoking cessation program for pregnant women. JMIR Mhealth Uhealth 2015;3(1):e10 [FREE Full text] [doi: 10.2196/mhealth.3846] [Medline: 25650765]
- 16. Evans W, Nielsen PE, Szekely DR, Bihm JW, Murray EA, Snider J, et al. Dose-response effects of the text4baby mobile health program: randomized controlled trial. JMIR Mhealth Uhealth 2015 Jan 28;3(1):e12 [FREE Full text] [doi: 10.2196/mhealth.3909] [Medline: 25630361]
- 17. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
- 18. World Health Organization. 2011. mHealth New horizons for health through mobile technologies URL: http://www.who.int/goe/publications/goe/mhealth/web.pdf[WebCite Cache ID 6yTi2OKxp]
- 19. de la Vega R, Miró J. mHealth: a strategic field without a solid scientific soul. a systematic review of pain-related apps. PLoS One 2014;9(7):e101312 [FREE Full text] [doi: 10.1371/journal.pone.0101312] [Medline: 24999983]
- 20. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
- 21. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. Matern Child Health J 2012 Jul;16(5):1092-1101. [doi: 10.1007/s10995-011-0836-3] [Medline: 21688111]
- 22. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
- 23. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
- 24. Badawy SM, Kuhns LM. Texting and mobile phone app interventions for improving adherence to preventive behavior in adolescents: a systematic review. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e50 [FREE Full text] [doi: 10.2196/mhealth.6837] [Medline: 28428157]
- 25. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
- 26. 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]
- 27. Din E. International Organization for Standardization. 1998. 9241-11 Ergonomic requirements for office work with visual display terminals (VDTs). Part II: Guidance on usability URL: https://www.iso.org/obp/ui/ [accessed 2019-04-27] [WebCite Cache ID 77x2YWeRH]
- 28. Dumas J. The great leap forward: the birth of the usability profession (1988-1993). J Usability Stud 2007;2(2):54-60 [FREE Full text]
- 29. Kushniruk AW, Triola MM, Borycki EM, Stein B, Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. Int J Med Inform 2005 Aug;74(7-8):519-526. [doi: 10.1016/j.ijmedinf.2005.01.003] [Medline: 16043081]
- 30. Horsky J, Zhang J, Patel VL. To err is not entirely human: complex technology and user cognition. J Biomed Inform 2005 Aug;38(4):264-266 [FREE Full text] [doi: 10.1016/j.jbi.2005.05.002] [Medline: 15967732]
- 31. Peute LW, Spithoven R, Bakker PJ, Jaspers MW. Usability studies on interactive health information systems; where do we stand? Stud Health Technol Inform 2008;136:327-332. [Medline: 18487752]
- 32. Castensøe-Seidenfaden P, Reventlov Husted G, Teilmann G, Hommel E, Olsen BS, Kensing F. Designing a self-management app for young people with type 1 diabetes: methodological challenges, experiences, and recommendations. JMIR Mhealth Uhealth 2017 Oct 23;5(10):e124 [FREE Full text] [doi: 10.2196/mhealth.8137] [Medline: 29061552]
- 33. [Pregnancy and birth An impression of the knowledge network birth care and research projects The Hague]. [Project-The App to work healthy during pregnancy] URL: https://www.zonmw.nl/nl/onderzoek-resultaten/doelmatigheidsonderzoek/programmas/project-detail/zwangerschap-en-geboorte/de-app-gezond-werken-tijdens-de-zwangerschap/ [accessed 2017-11-30] [WebCite Cache ID 6vMIAS1YS]
- 34. Advice and guidance by the occupational physician. Practice Guideline-Pregnancy, Postpartum Period and Work URL: https://www.nvab-online.nl/sites/default/files/bestanden-webpaginas/Guideline Pregnancy Postpartum Period and Work 0. pdf [accessed 2019-04-24] [WebCite Cache ID 77sEBK6rX]



- 35. Velu AV, van Beukering MD, Schaafsma FG, Frings-Dresen MH, Mol BW, van der Post JA, et al. Barriers and facilitators for the use of a medical mobile app to prevent work-related risks in pregnancy: a qualitative analysis. JMIR Res Protoc 2017 Aug 22;6(8):e163 [FREE Full text] [doi: 10.2196/resprot.7224] [Medline: 28830851]
- 36. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
- 37. Van Engen-Verheul M, Peute L, Kilsdonk E, Peek N, Jaspers M. Usability evaluation of a guideline implementation system for cardiac rehabilitation: think aloud study. Stud Health Technol Inform 2012;180:403-407. [Medline: <u>22874221</u>]
- 38. Peute LW, de Keizer NF, Jaspers MW. The value of Retrospective and Concurrent Think Aloud in formative usability testing of a physician data query tool. J Biomed Inform 2015 Jun;55:1-10 [FREE Full text] [doi: 10.1016/j.jbi.2015.02.006] [Medline: 25733166]
- 39. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3(6):514-522 [FREE Full text] [doi: 10.1370/afm.405] [Medline: 16338915]
- 40. Fransen MP, Leenaars KE, Rowlands G, Weiss BD, Maat HP, Essink-Bot M. International application of health literacy measures: adaptation and validation of the newest vital sign in The Netherlands. Patient Educ Couns 2014 Dec;97(3):403-409. [doi: 10.1016/j.pec.2014.08.017] [Medline: 25224314]
- 41. Ericsson KA, Simon HA. Verbal reports as data. Psychological Review 1980;87(3):215-251. [doi: 10.1037/0033-295X.87.3.215]
- 42. Brooke J. SUS-A quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability evaluation in industry. London: Taylor and Francis; 1996:4-7.
- 43. Plant RW, Ryan RM. Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: an investigation of internally controlling styles. J Personality 1985 Sep;53(3):435-449. [doi: 10.1111/j.1467-6494.1985.tb00375.x]
- 44. Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002 [Medline: 19.046928]
- 45. Nielsen J. Usability Engineering. Fremont: Morgan Kaufmann; 1994.
- 46. Eckert E, Bell AA. Invisible force: farmers' mental models and how they influence learning and actions. J Ext 2005;43(3):2 [FREE Full text]
- 47. Rook L. Emerald Insight. Mental models: a robust definition URL: https://www.emeraldinsight.com/doi/abs/10.1108/09696471311288519 [accessed 2019-04-24] [WebCite Cache ID 77sFYSxes]
- 48. Bolle S, Romijn G, Smets EM, Loos EF, Kunneman M, van Weert JC. Older cancer patients' user experiences with web-based health information tools: a think-aloud study. J Med Internet Res 2016 Dec 25;18(7):e208 [FREE Full text] [doi: 10.2196/jmir.5618] [Medline: 27457709]
- 49. van der Vaart R, Drossaert CH, de Heus M, Taal E, van de Laar MA. Measuring actual eHealth literacy among patients with rheumatic diseases: a qualitative analysis of problems encountered using Health 1.0 and Health 2.0 applications. J Med Internet Res 2013 Feb 11;15(2):e27 [FREE Full text] [doi: 10.2196/jmir.2428] [Medline: 23399720]
- 50. White BK, Martin A, White JA, Burns SK, Maycock BR, Giglia RC, et al. Theory-based design and development of a socially connected, gamified mobile app for men about breastfeeding (Milk Man). JMIR Mhealth Uhealth 2016 Jun 27;4(2):e81 [FREE Full text] [doi: 10.2196/mhealth.5652] [Medline: 27349756]
- 51. Sørensen K, Pelikan J, 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]

Abbreviations

HL: health literacy

IMI: Intrinsic Motivation Inventory

mHealth: mobile health

P and W app: Pregnancy and Work app

SUS: System Usability Scale

TA: think aloud

UEM: usability evaluation method



Edited by C Dias; submitted 29.06.18; peer-reviewed by C Doarn, S Chokshi, R Marcilly, A Rosman; comments to author 26.07.18; revised version received 14.10.18; accepted 27.12.18; published 09.05.19

Please cite as:

van Beukering M, Velu A, van den Berg L, Kok M, Mol BW, Frings-Dresen M, de Leeuw R, van der Post J, Peute L Usability and Usefulness of a Mobile Health App for Pregnancy-Related Work Advice: Mixed-Methods Approach JMIR Mhealth Uhealth 2019;7(5):e11442

URL: https://mhealth.jmir.org/2019/5/e11442/

doi: <u>10.2196/11442</u> PMID: <u>31094353</u>

©Monique van Beukering, Adeline Velu, Liesbeth van den Berg, Marjolein Kok, Ben Willem Mol, Monique Frings-Dresen, Robert de Leeuw, Joris van der Post, Linda Peute. Originally published in JMIR Mhealth and Uhealth (http://mhealth.jmir.org), 09.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on http://mhealth.jmir.org/, as well as this copyright and license information must be included.

