User involvement in developing the MYPLAN mobile phone safety plan app: a case study

Abstract:

Background: The efficacy of safety planning for people in suicidal crisis is not yet determined but using safety plans to mitigate acute psychological crisis is regarded ‘best practice’. In 2016-17, Australian and Danish key stakeholders were systematically involved in revising and translating the Danish MYPLAN mobile phone safety plan app into a culturally appropriate version for Australia.

Objectives: The objective of this study was to examine stakeholders’ suggestions and significant developments of MYPLAN during the user involving processes.

Method: We utilized a case study design where audio recordings of four focus groups and five workshops in Denmark and Australia were subjected to thematic analysis.

Results: The analyses identified three consecutive phases in the extensive development of the app: From phase one, ‘Suggesting core functions’, through phase two, ‘Refining functions’, to phase three, ‘Negotiating the finish’. The user involving processes continued to prevent closure and challenged researchers and app developers to repeatedly rethink the app’s basic design and functions. It was a limitation that the analysis did not include potentially determinative ‘backstage’ dimensions of the decision-making process.

Conclusion: The extended user involvement prolonged the development process, but it also allowed for an extensive exploration of different user perspectives and needs.

Introduction

Applications (apps), which can be downloaded to a mobile device, have the potential to provide health interventions to people who would not otherwise receive these services due to barriers, such as limited recourses, access, as well as poor help seeking capabilities [1]. mHealth, the use of mobile and wireless technologies to promote health, can assist people in assessing, monitoring, and tracking their mental and physical health; acquiring health information and psychoeducative resources; accessing real-time, recorded, or virtual psychotherapy; and connecting with social networks, cf. [2]. However, these potential benefits are counterbalanced by current limitations to mHealth.

Unlike most other health care interventions, the provision of health apps is not highly regulated [3]. There is a plethora of mental health apps, but very little research has explored their efficacy [1, 4, 5], and there is a growing concern about the effects, usefulness or potential harmfulness of health apps [6]. Seemingly, there is no published research on the effectiveness of suicide prevention apps [5, 7], although two published reviews of suicide prevention apps provide some insights: Aguirre et al. [8] identified 27 apps that could be linked to suicide prevention and assessed them
according to: 1) research/evaluation of the app, 2) privacy, 3) usability and accessibility, and 4) appropriateness of functionality. In particular, they noted that 12 of the apps did not include a direct link to a crisis hotline. Larsen et al. [6] reviewed 49 apps that included at least one suicide prevention feature and were available in the Australian Google Play store or the Australian iTunes store. Each app’s tools were mapped against 18 suicide prevention strategies and ranked according to these strategies’ level of evidence (from one to four) as identified in the research literature. Only 10 of the 24 suicide specific apps contained a crisis support/helpline, which was rated as the highest level of evidence, i.e. ‘strong evidence’. Further, the two most employed strategies were: peer-support (16 apps) and safety planning (13 apps), which were considered as having ‘some evidence’ or being ‘best practice’, respectively. However, it is difficult to make balanced interpretations of this latter study, as Larsen et al. counted the 10 apps available for both Android and iPhone as 20 apps. Both reviews concluded that suicide prevention apps need to be supported by stronger research evidence.

In contrast, Nicholas et al. [9] made a case for abandoning the traditional evidence base for mobile phone applications, including randomized controlled trials (RCTs). The authors reasoned that the fast-paced mobile market requires alternative, more rapid, evaluation methods, such as iterative participatory research and single case designs. Further, they argued for developing new, alternative ways of accrediting high-quality apps. In accordance with Nicholas et al.’s suggestions, we present a case study of the user involving processes that led to the revised version of the mobile safety plan app, MYPLAN [10](Authors, in press).

MYPLAN was originally a Danish app modelled after Stanley and Brown’s [11] paper based safety planning tool. This tool combines at least three preventive strategies: 1) encouragement to contact peers and professionals; 2) cognitive, problem solving, and personalized safety planning; and 3) encouragement to limit access to lethal means [11]. The initial Danish version of MYPLAN from 2013 augmented the encouragement to contact professional support by including map with directions to the nearest emergency room. This existing version of MYPLAN is currently available in English, Danish, and Norwegian for Android and iPhone [10].

Democratizing knowledge and ensuring the relevance of research and design to end-users are key elements of user involvement [12]. Between 2015 and 2017, the initial version of MYPLAN was revised by means of involving Danish and Australian users with the purpose of developing a cross-cultural adaptation [13] and translation into Australian English. In addition, the app’s storage changed from being app based to cloud based, which, for instance, allowed users to share strategies online through a strategy-bank. In this paper, we examine significant developments of MYPLAN during the extended user involving processes.

**Methods**

We conducted an instrumental case study [14], which is a useful method for gaining insight into a particular event, such as the gradual development of MYPLAN’s design and functions.
Study context and participants:
Users were involved through focus groups in Denmark and participatory workshops in Denmark and Australia. The specific focus group and workshop design are detailed in Figure 1 and Table 1.
Figure 1. Organisation of the study.
Table 1: Focus groups and workshops

<table>
<thead>
<tr>
<th>Focus group #1</th>
<th>Participants</th>
<th>Date</th>
<th>Length (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denmark: ‘DK FG #1’</td>
<td>Young users (n=5) Moderators (n=2)</td>
<td>December 2015</td>
<td>126</td>
</tr>
<tr>
<td>Focus group #2</td>
<td>Adult users (n=8) Moderators (n=2)</td>
<td>November 2015</td>
<td>132</td>
</tr>
<tr>
<td>Denmark: ‘DK FG #2’</td>
<td>Relatives (n=3) Moderators (n=2)</td>
<td>December 2015</td>
<td>129</td>
</tr>
<tr>
<td>Focus group #3</td>
<td>Clinicians (n=10) Moderators (n=2)</td>
<td>November 2015</td>
<td>120</td>
</tr>
<tr>
<td>Denmark: ‘DK FG #3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focus group #4</td>
<td>Users (n=3) Technician (n=1) Moderators (n=2)</td>
<td>January 2016</td>
<td>155</td>
</tr>
<tr>
<td>Denmark: ‘DK FG #4’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workshop #1</td>
<td>Users (n=2) Moderators (n=2)</td>
<td>August 2016</td>
<td>156</td>
</tr>
<tr>
<td>Denmark: ‘DK WS #1’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workshop #2</td>
<td>Users (n=80) Moderators (n=3)</td>
<td>August/September 2016</td>
<td>Up to 180</td>
</tr>
<tr>
<td>Australia: ‘AUS WS #3’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workshop #3</td>
<td>Users (n=3) Moderators (n=2)</td>
<td>December 2016</td>
<td>141</td>
</tr>
<tr>
<td>Australia: ‘AUS WS #4’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workshop #4</td>
<td>Users (n=3) Moderators (n=3)</td>
<td>January 2017</td>
<td>188</td>
</tr>
<tr>
<td>Australia: ‘AUS WS #5’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. In late 2015, four Danish (DK) focus groups (FG) (DK FG #1-4) were held with key stakeholder participants (adult and young users, relatives, and clinicians) (Authors, in press). The focus groups were thematically organized to focus on discussing personal experiences of using MYPLAN and participants’ suggestions for improving design and functionality. In particular, participants were asked to consider whether a safety plan on a smartphone should include auto-generated communication (e.g. notifications and pre-written messages), digital memory (e.g. a hope box), GPS (e.g. monitoring), and self-assessment (e.g. monitoring and testing). The moderator summarised the suggestions that came up during the focus groups. Notable suggestions included incorporating an alarm button, a safety plan for relatives, GPS monitoring, notifications, pre-written messages, a hope box (that could possibly also be shared by a group of users), capacity to share coping strategies, and tests.

2. In 2016, two Danish participatory workshops (WS) (DK WS #1-2) were held where end-users were invited to evaluate the updated design and functionality and to suggest further changes. The workshop participants were recruited from the ‘adult users’ focus group, which had been the most active in suggesting changes to the app. The app technician/programmer took part in the first of these workshops.

3. In 2016, 10 Australian (AUS) participatory workshops (AUS WS #3) were held with 80 participants, including young people, supportive others and headspace clinician/service staff. Over a three-hour period, the participants examined the wireframes of several apps, including MYPLAN, through a process of discovery, evaluation, and prototyping. Data were collected through: 1) written comments made by participants, 2) hand-drawn mock-ups, and 3) qualitative field notes written during the workshops (Authors, in press).

4. Finally, in late 2016 and early 2017, there were two Australian participatory workshops (AUS WS #4-5) where end-users were invited to evaluate the app for the Australian context. The workshops focused primarily on functionality and the culturally appropriate wording of the Australian version, in particular of the progressive on-boarding, which gradually provides users with information when they use functions of an app for the first time. Young users were recruited through mental health services, such as headspace and ReachOut. All participants were given access to an English prototype of MYPLAN prior to the workshops. Moderators took notes at all five Danish and Australian workshops, which were used to develop prototypes and a list of additional suggestions that were evaluated by the app developers and technicians with respect to importance and cost.

Data:
The four focus groups (DK FG #1-4) and four of the workshops (DK WS #1-2, and AUS WS #4-5) were digitally audio recorded. Data included verbatim transcriptions of the focus groups and the first and the second author’s extensive summaries of the suggestions and discussions in the workshops. Field notes and design artefacts generated during workshops (AUS WS #3) were also independently knowledge translated by a representative end-user team (young people, supportive others, clinician and service staff) (authors, in press). In addition, all written notes produced during the workshops and the developers’ list of suggestions were collected and
analysed.

**Analysis:**
Data were subjected to a thematic analysis [15, 16], which included: 1) reading/listening to the full data set to familiarize researchers with content; 2) mapping trajectories of suggestions from inception to rejection, or from inception to implementation; 3) parallel coding by two researchers of full data set and written summaries; 4) gradual development of descriptions of thematic content/discrete phases; and 5) corroboration of the description of phases by re-examining data.

**Ethics:**
We notified the relevant Danish regional research ethics committee and the Danish Data Protection Agency about the Danish focus groups and workshops (DK FG #1-4 and DK WS #1-2); neither institution reported any reservations towards the study. The University of Sydney Human Research Ethics Committee (HREC) approved the research of Australian participatory workshops (AUS WS #3: ref. no. 2016/529 and AUS WS #4-5: ref. no. 2016/749). All participants gave their informed consent to participate, based on written and oral information about the study. Interview responses were handled in full confidentiality and all details that could potentially be used to identify individual participants have been altered in the data extracts presented in the Results section below.

**Results**

The analysis identified three temporal phases during the user involving processes, which were characterized by distinct types of negotiations. The first phase, ‘Suggesting core functions’, was characterized by a focus on discussing the potential inclusion of basic app features. The second phase, ‘Refining functions’ was characterized by testing and negotiating the design of newly implemented app features. The third phase, ‘Negotiating the finish’ was characterized by tests and discussions about the final layout and wording. The three data extracts presented in the findings section below were selected because they were characteristic for the different phases.

Table 2 illustrates the different phases with two examples, speed dialling buttons and mood rating.
Table 2. Examples of the phases in the user involvement processes

<table>
<thead>
<tr>
<th>Speed dial buttons</th>
<th>Phase 1: Suggesting core functions. DK Focus group #1-4</th>
<th>Phase 2: Refining functions. DK Workshop #1 and #2 and AUS Workshop #3</th>
<th>Phase 3: Negotiating the finish. AUS Workshop #4 and #5</th>
<th>After Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users suggested speed-dialling buttons. The colours of the buttons should reflect ‘the degree of emergency’ (yellow/red). They should be simple to use: an emergency call should be made with no more than two clicks</td>
<td>Development of wireframes with different types of menus and different opportunities for placing an alarm button</td>
<td>The location of the two buttons was discussed and they were placed at the bottom of the front page. The yellow button should be named ‘Help’ and was assigned a telephone icon. The red button should be named ‘Alarm’ and was assigned an exclamation mark icon. The way to assign contacts to buttons was discussed and the number of possible contacts to assign to the yellow button should be ten</td>
<td>Implementation of two buttons on front page. A yellow ‘Help’ button with a telephone icon and a red ‘Emergency’ button with a white exclamation mark. Alarm call in two clicks from front page (but a four-digit access code was also added). Maximum of 10 ‘yellow’ contacts</td>
<td>The need for better on-boarding information, explaining the way to add contacts to buttons, was noted. The icon on the yellow button should be a telephone and a head in a circle. The icon on the red button should be a white cross in a red circle</td>
</tr>
</tbody>
</table>
The development and implementation of the speed dial buttons was gradual and characterised by minor edits from the programmers that were slightly different than what users had originally suggested.

Mood ratings

<table>
<thead>
<tr>
<th>Phase 1: Suggesting core functions. DK Focus group #1-4</th>
<th>Between Phase 1 and Phase 2</th>
<th>Phase 2: Refining functions. DK Workshop #1 and #2 and AUS Workshop #3</th>
<th>Between Phase 2 and Phase 3</th>
<th>Phase 3: Negotiating the finish. AUS Workshop #4 and #5</th>
<th>After Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users discussed the usefulness of tests and here the idea to use smileys for the mood rating arose. In general, users argued against using tests in the app. Suggestion of making mood rating customizable</td>
<td>Introduction to different types of emoji-based mood trackers</td>
<td>The selected emojis were reviewed and edited (they looked angry rather than sad). The wording of the mood tracking function was discussed and changed. Suggestion of a clear introductory text. The function of mood rating, including its reminders, should be customizable</td>
<td>Implementation of emoji-based mood tracker and SIDAS suicide ideation measure. The presence of mood ratings in the menu became customizable by individual user</td>
<td>Objection to the wording of SIDAS’s response categories. An introductory text to SIDAS is reformulated. The feedback from SIDAS should be gentler and less demoralising</td>
<td>Implementation of revised on-boarding information about SIDAS and its response categories</td>
</tr>
</tbody>
</table>

The development and implementation of the mood rating was gradual, but characterised by reluctance from users regarding its usefulness, in particular the SIDAS questionnaire. The text linked to SIDAS was completely rewritten by users before final implementation.
Discussions in all three phases were characterized by very low levels of displayed disagreement; when different opinions were voiced, participants would most often resolve disagreement by suggesting that a given app feature should ultimately be optional and adaptable by individual users.

**Phase 1: Suggesting core functions**
The first phase primarily took place during the Danish focus groups, where the functions of the original MYPLAN app were discussed along with ideas for new functions, which were introduced by the researchers. Participants raised some principal issues about the app’s design and functions: *First*, some participants voiced different understandings of what end-users might enter into the app’s core problemsolving function: ‘Warning signs’. They noted that the concept was unclear and could be interpreted as ‘signs of a potential crisis’ or as ‘signs of a current acute crisis’, which seemed very far from the intention with this function. In Phase 1, these discussions did not lead to explicit suggestions for changes to the app. However, the discussions reappeared in Phase 3 as discussions of whether to frame the app as a safety plan or a suicide prevention tool, see below. *Second*, users were aware not to over-complicate the app by suggesting too many functions, which occasionally happened when participants started outbidding each other with numerous interactive Facebook inspired functions. For participants, the focus on ‘need to have’ functions rather than ‘nice to have’ functions was related to concerns about developing an app that could be used in an emergency where simplicity would be paramount.

In the following data extract from focus group DK FG #4, which included relatives of MYPLAN users, the participants discussed the potential use of GPS. They had discussed a vignette about a young girl who switched on her GPS when she felt suicidal, which would automatically alert parents and clinicians who would then be able to follow her phone’s location. After discussing that MYPLAN users could potentially be bluffing and that there would be a need for sincerity from the offspring’s side, they started reflecting on the personal costs such alerts would have on themselves as well as issues of intrusions that GPS surveillance might inflict on their sons or daughters.

**Data extract 1:**

**Moderator:** I hear you describing something like the existing ‘Find My iPhone’ feature that parents use to check if the kids are at school or other places. What are your thoughts?

**Emma:** No. That would not be OK for me. Ian [her son] would definitely be too old for that, so I would find that it was intruding on his private space.

**Jenny:** That would send Rachel [her daughter] round the bends; you would be stalking her, right? I think it is good, you could use it in crisis situations or that they can use it in crisis situations.

**Emma:** It must not turn into surveillance, because they will just switch it off.

**Rachel:** It has to be optional.

**Moderator:** So it would be ok if you had an agreement with your daughter for a while?

**Rachel:** Yes, if I could cope with it, my daughter would have to control it herself.
Emma responds to the Moderator’s question by refusing the idea of monitoring her adult son, as it would be too intrusive. Jenny follows up by agreeing that it would not be good for her daughter, but she adds that it might be good in a crisis. She revises her own initial statement when she describes GPS monitoring as something ‘you’ could use in a crisis situation, to something ‘they’ [their sons and daughters] could use, which could indicate that she preferred to describe it as the offspring’s tool, and not the parents’. Emma continues by voicing a similar concern and emphasizes that it must not have a character of surveillance as that would be unacceptable and the children would simply not use it. Rachael emphasizes that it has to be optional for the children to be followed by GPS and concludes by stressing that it would require an agreement and that her daughter would have to control it. However, she also states that she, herself, would need to be able to cope with being able to monitor her daughter, which could indicate that it would be emotionally taxing having access. The option of GPS seems to trigger these participants’ ambivalence about, on one hand, having control and certainty, and, on the other hand, the anxiety of a hands-off approach to their vulnerable children.

In Phase 1, the researchers ultimately controlled the discussions as they introduced the focus group agenda highlighting the core app features. However, participants were invited and most often able to voice their concerns, which had a direct impact on the design of the core functionality.

Phase 2: Refining functions
The second phase primarily took place during workshops DK WS #1-2 and AUS WS #3 after the designers had done a first revision of the app. The majority of changes were only available as printed wireframes and discussions were focused on developing additional ideas. As in Phase 1, there was a continual flow of suggestions that were never discussed in any significant depth because of their apparent complexity (e.g. when you speed dial the emergency services, the phone should automatically send your GPS coordinates, or the phone should be able to prompt you or your network if you come near certain risky locations according to codes you have entered yourself, for instance a bar or a tall building), and suggestions that were explicitly negotiated in the groups. Only one new core function, the Rant Box, a place where self-selected text and images can be destroyed by applying pressure to the screen, was introduced following the AUS WS #3 and developed for the app.

In workshops DK WS #1 and 2, relatively structured discussions organized by the moderators took place and they ended the workshop by restating the key points raised and decisions made. In DK WS #1 and 2, moderators both implicitly and explicitly drew on insights from the focus groups. Points were mostly discussed by stating ideas one after the other without much explicit disagreement. There were several instances in DK WS #1, where the participating technician funnelled the discussions in a particular direction based on arguments linked to the concrete programming of features and economy. Finally, there was an incident where the app designers insisted on keeping a feature, the structured SIDAS suicidality rating scale, despite repeated pushback from user participants who perceived it as redundant and unnecessary, see table 2.

The following data extract 2 illustrates the typical collaborative nature of negotiating
and refining a function during a workshop. The participants were discussing different ways of designing a speed dialling function and had previously settled on a solution with two buttons at the bottom of the home screen, a yellow for sub-acute situations and a red for acute crisis. Now, the key issue was to discuss how much flexibility a user had in terms of assigning particular functions to each of the buttons. Interactions were fast and in the data extract “… indicates that the next speaker started talking by interrupting the previous speaker slightly.

Data extract 2:

James: My idea would be that [the red button] would do only one thing. When you press the red one then you call a particular person or 112 [emergency call].

Linda: I think that too. Alternatively, you could, if you wanted both 112 and Lifeline, then you could add two…

Moderator 1: Or you made 112 the red [button] and Lifeline the yellow [button].

Linda: Yes, yes, then you’d…

James: Yes, of course…

Moderator 1: I believe that the red one should be - only need one click. You shouldn’t have to make any decisions…

James: That’s also what I think. The yellow is meant for when you feel terrible and you need to talk to someone. When I have been so far out that I simply needed to talk to someone. That would be how I saw the yellow phone [button] whereas the red phone is when I’m almost out of reach. I need to get hold of someone before things go really wrong. Therefore, as you say, the red one is for one person whereas with the yellow you might choose between several. There needn’t be a maximum, but between 5 and 6 persons. Having 20 persons to choose from should not confuse you.

James starts out by explaining his ideas about the differences between the buttons’ functions and Linda follows up by explicitly agreeing, but at the same time adding that the red button could be linked to several numbers. The moderator disagrees by emphasizing that he believes that the red button should be simple to use. James expresses his agreement and expands on his idea for the two buttons by explicitly drawing on his personal experiences of being in a crisis. He concludes by re-wording the issue regarding simplicity that had previously been forwarded by Moderator 1.

The workshops were highly interactive, and moderators engaged actively. The researchers had a high degree of control through their on-going engagement and a systematic and structured summarizing, which ultimately funnelled a consensus about each discussion point. However, the very explicit and consensus-seeking approach gave user participants a very direct influence on decisions. The user participants who had participated in both a focus group and the workshops seemed to display more ownership of the process and a better understanding of what was being developed, which strengthened and qualified their contributions in discussions.

Phase 3: Negotiating the finish

The third phase primarily took place during workshops AUS WS #4 and 5 when MYPLAN’s core functions had been designed, tested and refined. Fixing glitches
continued to take place as it had done in Phase 2 and all issues were listed and fed back to the app developers to prioritize and resolve.

Unlike in the Danish data collection context, Australian user participants were not introduced to the app by health care professionals they had met in the clinic. Therefore, the Australian participants (workshops AUS WS #3-5) relied heavily on the e-introductions, which they did not find intuitive or helpful. The Australian participants identified a need for better e-introductions and suggested a progressive on-boarding to MYPLAN’s functions. The participants from workshops AUS WS #4 and 5 were engaged in writing and editing the introductory texts for the app, both during the workshops and via mail after the workshops. These texts were sent to the app developers, who implemented them in both the Australian and Danish versions.

The Australian participants felt that the relative non-interactive introductions made the app’s particular clinical language-use even more unacceptable. The following data extract 3 is taken from AUS WS #4 where a participant, Sarah, highlighted her discomfort concerning the image and text she read after responding to the five items of the SIDAS questionnaire.

Data extract 3:

Sarah: “It says the word suicide a lot and it says a lot of words that probably when you are in distress you don’t need to be confronted with. And more importantly, it is talking about and not to you again, which is really kind of demoralising. I feel like it could be a lot shorter and more pleasant very easily.”

Sarah continued to elaborate on her position that the language use came from a medical and paternalistic position, which, to her, objectified the user and was not helpful for a person in a crisis. This led to a series of discussions about the general coherency: was it a problem that MYPLAN was, on one hand, a personalized self-help tool and, on another hand, employed medical surveillance and medical psychoeducation? The Australian participants strongly encouraged making a pure non-medicalized and non-pathologized self-help safety plan that would be acceptable for people who felt distress, but who would not identify as being in a crisis (a crisis management plan), let alone being suicidal (a suicide prevention plan).

The push back from the Australian workshops was a genuine surprise for the researchers who were forced to reconsider their own core assumptions about the app’s purpose and its users. Nevertheless, they appreciated the feedback and welcomed the user participants as authors of the app’s on-boarding texts and features.

Discussion

Models of user involvement often classify levels of involvement according to the distribution of power and control between user participants and researchers [17]. For example, Hanley et al. [18] differentiate between: 1. Consultation, where service users’ perspectives are explored by researchers and potentially brought in to decision-making processes. 2. Collaboration, where researchers are actively committed to engaging in on-going partnerships with service users, and 3. User control, where ‘the
locus of power, initiative and subsequent decision making is with service users rather than with the professional researchers’ [18, p. 10]. While the user involving processes analyzed in the present study clearly included elements of consultation, it remained debatable to which extent the researchers were committed to collaborate and genuinely share control. While study participants were consulted over extended periods of time and occasionally designed direct suggestions to the app, which resembled elements of user-controlled involvement, the researchers ultimately controlled the data-collection sessions and the information that was recorded, prioritized, and fed back to the app developers.

Although users are regularly involved in the evaluation, design and development of mental health apps, see for instance [19, 20], the actual level of involvement remains challenging to ascertain. Furthermore, there is no globally agreed upon terminology describing practices of ‘user involvement’. Terms, such as co-design, co-production, co-development, etc., often imply that users are only being consulted, i.e. not involved as a resource in their own right and continue to have very little actual power and control in health research and development. We were not able to identify any research of actual user involvement in research and development of mental health apps. Hence, more research in this area is needed to determine actual levels of involvement (Authors 2018).

It was difficult to gauge if and how the user involving processes enhanced the MYPLAN app. Hawton et al. ’s [21] Cochrane review suggested possible mechanisms as to how the preventive strategies implemented in the app might mitigate suicidal crisis. These include enhancing problem solving and coping skills (achieved by linking warning signs and strategies) and an increased sense of social connectedness (achieved by listing social contacts, and quick messages). While Stanley and Brown’s [11] paper version of the safety-planning tool drew on at least three different preventive strategies, the initial as well as the revised MYPLAN apps added further features, in particular by making use of smartphone technology. More, control was handed over to the user in the discussions, which led to surprising (for researchers) suggestions, for example the use of non-pathologizing and clinical crisis language. In line with general aims of user involving strategies, these processes were perceived to be relevant by users, but added complexity to the designing process. It remains unclear whether the proposed changes could have unintended, negative consequences. Additionally, compared to a paper version of the safety plan, the numerous preventive strategies could hypothetically decrease transparency and obstruct ‘simple’ safety planning. This is currently being tested in a randomized trial [22]. Finally, reconsidering Nicholas et al.’s [9] introductory argument for abandoning RCTs in evaluations of health apps, it seems as if iterative participatory research and single case design (such as the present study) allow for intuitive new innovations, but remain very limited in terms of evaluating long term impacts, which would need different, more extensive methods of testing, such as RCTs.

As noted by Grundy et al. [3], adverse events and possible harm are rarely mentioned in disclaimers of mental health care apps. However, no reports of harm were voiced during any of the sessions and participants’ reports of potential ambivalence or adversity were managed in situ and in the ongoing design of the app. Grundy et al. [3] also noted that mental health care apps have a tendency to claim easy and rapid improvement of mental health in their presentations in app store. However, contrary
to these visible and positive claims, the formal app disclaimers tended to distance themselves from presenting the app as a medical service [3]. Arguably, this happens in order to clearly and conveniently differentiate a given app from a ‘medical device’, which would be subjected to extensive – and expensive – medico-legal regulation. This would happen if an app intends to be used in diagnosis, prevention, monitoring, treatment, or alleviation of disease. Phase three of the current study’s user involving process included a strong push towards de-medicalizing the management of everyday feelings of distress, which could make the app relevant to a much wider audience. This, in effect, moved the app away from potentially being classified as a medical device. To some extent, the revised design of the app begs the question of whether there might be a need for a parallel app with a strong explicit focus on suicide prevention?

Strengths and limitations
We collected observational data from a range of meetings over a long period of time, which allowed a basic mapping of the introduction and negotiation of ideas over time. However, we did not have observational data of negotiations of proposed suggestions that took place in consultations between app designers away from users. Such negotiations would, for example, include discussions among app designers and technicians about which of the listed suggestions should be prioritized in light of their complexity and costs, and decisions taken by the app programmers. From a methodological perspective, we believe that users’ involvement might be perceived as less impactful if these ‘back stage’ negotiations were part of the analysis. Arguably, users should be formally involved in as many as possible of these crucial design processes to achieve more genuine collaboration.

Most participants were offered to join several sessions, which allowed them to voice their opinion as the app was developed gradually over time, cf. the different phases, and to strengthen their sense of personal ownership. However, despite elaborate recruitment strategies, most workshops included only a very limited number of participants. Although we held several workshops and most participants had strong voices and opinions, it would have been advantageous to recruit slightly larger and more varied groups of participants.

Conclusion

The analyses identified three consecutive phases in the extensive development process of a safely plan provided as an app. While the phases reflected a gradual implementation process, the user involving processes continued to prevent closure and challenged researchers and app developers to continually rethink basic app design and functions. The implementation process remains without full closure as a current aim for developing the MYPLAN app is to secure further funding and to implement more items from the list of suggestions that has been compiled. This also implies to continually monitor the use of the specific functions and omit the ones that users do not use.

The variety of mHealth tools are likely to increase globally, which highlights a need for procedures for safe adapting, translation and tailoring of apps across countries and cultures. In line with Harper Shehadeh et al. [13], we believe that detailed reporting of adaption methods is crucial, and the systematic involvement of service users could be
an important way to increase the trustworthiness of such adoptions.

Funding:
This work was supported by TrygFonden under Grant 133722.
References:


