Proportionate methods for evaluating a simple digital mental health tool

E Bethan Davies,1,2,3 Michael P Craven,1,3,4 Jennifer L Martin,1,2,3 Lucy Simons1,2,3

1NIHR MindTech Healthcare Technology Co-operative, Institute of Mental Health, University of Nottingham, Nottingham, UK; 2Division of Psychiatry and Applied Psychology, Institute of Mental Health, University of Nottingham, Nottingham, UK; 3NIHR Nottingham Biomedical Research Centre, University of Nottingham, Nottingham, UK; 4Bioengineering Research Group, Faculty of Engineering, University of Nottingham, Nottingham, UK

Correspondence to Dr E Bethan Davies, NIHR MindTech Healthcare Technology Co-operative, Institute of Mental Health, University of Nottingham, Triumph Road, Nottingham NG7 2TU, UK; bethan.davies@nottingham.ac.uk

ABSTRACT

Background Traditional evaluation methods are not keeping pace with rapid developments in mobile health. More flexible methodologies are needed to evaluate mHealth technologies, particularly simple, self-help tools. One approach is to combine a variety of methods and data to build a comprehensive picture of how a technology is used and its impact on users.

Objective This paper aims to demonstrate how analytical data and user feedback can be triangulated to provide a proportionate and practical approach to the evaluation of a mental well-being smartphone app (In Hand).

Methods A three-part process was used to collect data: (1) app analytics; (2) an online user survey and (3) interviews with users.

Findings Analytics showed that >50% of user sessions counted as ‘meaningful engagement’. User survey findings (n=108) revealed that In Hand was perceived to be helpful on several dimensions of mental well-being. Interviews (n=8) provided insight into how these self-reported positive effects were understood by users.

Conclusions This evaluation demonstrates how different methods can be combined to complete a real world, naturalistic evaluation of a self-help digital tool and provide insights into how and why an app is used and its impact on users’ well-being.

Clinical implications This triangulation approach to evaluation provides insight into how well-being apps are used and their perceived impact on users’ mental well-being. This approach is useful for mental healthcare professionals and commissioners who wish to recommend simple digital tools to their patients and evaluate their uptake, use and benefits.

INTRODUCTION

Mobile health (mHealth) involves using handheld and typically internet-connected digital devices, such as smartphone and tablets, for the purpose of healthcare. These devices run a wide range of software applications (apps). Evaluation of digital technologies for young people’s mental health has focused largely on internet and computer-delivered cognitive behaviour therapy (eCBT) for depression and anxiety and computerised treatments for diagnosed conditions, for example, attention deficit hyperactivity disorder.1 Researchers have typically adopted traditional health technology assessment approaches for eCBT, such as randomised controlled trials (RCTs), because of the need to demonstrate the promised benefits and to justify the healthcare resources that are required to deliver them. While the literature does regularly cite that eCBT is more cost-effective compared with traditionally delivered CBT, findings from the recent Randomised Evaluation of the Effectiveness and Acceptability of Computerised Therapy trial found little cost-effectiveness differences between two eCBT programmes and GP treatment as usual.2

Because of the resources and time needed to plan, undertake and implement findings from traditional evaluation assessments, RCTs and other ‘big’ trial designs are also considered to be out of proportion to the rapid development and obsolescence of digital technologies,3,4 especially self-help tools (often ‘apps’) which can be accessed directly by users and may not be used in conjunction with clinical services. Several methods have been used to evaluate apps,4 and it is unlikely that one methodology will fit all. It is also argued that formal mHealth trials may not represent their intended real-world use, and evaluations should appraise technologies within the settings where they are intended to be used.3 In particular, the emergence of relatively unregulated (non-medical device) apps aimed at public mental health and well-being indicates the need for alternative approaches to evaluation.

The focus of this article is evaluation methods for simpler digital tools, such as mobile apps for assisting well-being, which are widely and publicly available and intended for use without direct clinical supervision.5 These products may have potential in helping young people overcome some of the traditional barriers to accessing support and reducing stigma.6,7 Although there has been discussion of the need to stratify evaluation methods of healthcare apps based on complexity and risk,8 there has been little research investigating how best to evaluate examples of simple, self-help digital tools. Conceptualisation of engagement with digital health interventions has recently been recognised as an important area for research.9 We propose that the quality and value of these digital tools can be assessed through analysis of real-world usage data and assessment of user experience, methods which match the relative simplicity of these tools and anticipated size of effect on users’ mental health. Furthermore, the gap being addressed is the need for better quality early evaluation of new digital health products, whether this is an initial summative evaluation of an established product to help decide whether it is worth adopting, or as in this case the formative evaluation of an app during its development.

Using the context of an independent evaluation of In Hand, a mental well-being smartphone app, this paper aims to show how a proportionate evaluation can be realised while using elements of web-based and app-based software design that can be readily introduced by product developers into existing applications. In Hand (www.inhand.org.uk, launched 2014) is an app developed by young people with experience of mental health problems to support well-being through focusing the user on the current moment and bringing balance to everyday life. In Hand is a simple, free-to-download digital self-help tool publicly available on iOS and Android, intended to be used independently of healthcare services. Using a traffic light inspired system, the app takes the user through different activities depending on how they are feeling (see figure 1). Its development was led by a UK arts organisation, working with a digital agency and a public mental health service provider. The project’s clinical lead drew on principles of cognitive behaviour therapy and Five Ways to Well-being10 during the development process, but the primary influence on the content arose from needs derived in the co-design process that explored coping strategies used by young people at times of stress or low mood.

NIHR MindTech HTC was commissioned to undertake an independent evaluation of a suite of digital resources produced by Innovation Labs11; In
Hand was one of these products. As these digital resources were about to be publically launched as the evaluation stage commenced, the team proposed to observe how users engaged with the tools by capturing background usage data and seeking feedback directly from users through embedding user elicitation into the tools. This approach would enable the capture of insights from naturally occurring users and gain understanding of how people interact with it in the ‘real world’.

METHODS

Research design
We wanted to evaluate how In Hand was engaged with in real life, and what kind of benefits users gained through use. Three methods of data collection were used to gain insight into naturalistic use of In Hand: (1) sampling and analysis of quantitative mobile analytical data (eg, number of individual user sessions, number of interactions with each section of the app); (2) a user survey with questions adapted from a validated well-being measure (the Short Warwick-Edinburgh Mental Well-being Scale12 (SWEMWBS)) and (3) semistructured individual interviews with a subsample of survey respondents.

Procedure

Mobile analytical data
Flurry Analytics (now part of Yahoo!, http://developer.yahoo.com/ analytics), a tool that captures usage data for smartphone/tablet apps, was used to securely access anonymous, aggregated data about users’ interactions: this captured time spent using the app, frequency of use, retention over time and key app-related events such as visits to specific content. No identifying information for any user was directly available to the research team, and interactions across sessions by individual users could not be tracked. Data were captured and analysed for the iOS and Android versions of In Hand between May and October 2014.

User survey
A software update to In Hand was implemented in August 2014 to add an invitation to complete the survey into the app. On first opening the app after the update, users were presented with a ‘splash’ page which invited them to feed back on the app, which was also added to the app’s menu (see figure 2). Once users clicked this, they were taken to the In Hand website (via internet connection outside the app), where further information about the survey was presented. Interested participants then clicked a URL to access the survey (hosted on SurveyMonkey) and consented to complete it. As an incentive to participate, respondents could opt into a prize draw to win a shopping voucher on survey completion. The online survey was open for 6 weeks.

Semistructured interviews
At the end of the online survey, users could enter their email address to register their interest in participating in a semistructured interview about their experience of In Hand. Twenty-four users registered their interest and were emailed an information sheet that included a request for an informal chat with a researcher about the interview procedure and to answer any questions. From this, eight people chose to participate: an arrangement for an interview was made and an online consent form emailed to them and completed prior to interview Six interviews were carried out by telephone and were audio recorded, with two asynchronously conducted by email. Participants received a £15 voucher as an inconvenience allowance.

Sampling and recruitment
The target group for In Hand is young people (aged up to 25 years), but given it was publicly available, people of all ages could access and use the app and therefore take part in the evaluation. All In Hand users (aged ≥16 years) were eligible to participate in the survey and interview. Those who entered their age as ≤15 (meaning parental consent would be required) at the start of the survey were automatically directed to an exit webpage. One hundred responses to the survey were sought: using data from another digital tool evaluation,11 a conservative estimate of 10% of all app users completing the survey was made, suggesting that the survey needed to run for 6 weeks to achieve 100 responses.

For the interviews, a purposeful sample of 12 respondents with varied demographic characteristics was sought to gain a range of perspectives. Survey respondents who opted in to the interviews were sampled according to their age, gender, ethnicity, sexuality, disability, geographical location and their use and experience with In Hand.

Survey and interview guide design
The user survey and semistructured interview topic guide were developed specifically for this study. Young people involved with Innovation...
Labs and In Hand development collaborated with the NIHR MindTech Team in generating the study’s design and areas to be explored in the survey and interview, reviewing and testing out the online survey and interview schedule and finalising the study materials. To explore the kinds of benefits In Hand had to users, the survey questions were based on the SWEMWBS, an evidence-based measurement tool with seven dimensions of mental well-being which has been shown to be valid and reliable with young people. In this real world, observational evaluation, it was not possible to assess participants’ well-being before and after use of the app, but rather, users were asked to rate to what extent In Hand had helped them on specific dimensions of mental well-being. During the co-design process, one of these dimensions was reworded to be more accessible to young people (‘feel optimistic’ changed to ‘have a positive outlook’), and three other dimensions were selected for inclusion in the survey to reflect the young people’s experiences (‘feel ready to talk to someone else’, ‘feel less stressed’, ‘feel more able to take control’). These dimensions were also used to guide the interview topics.

Data analysis
Aggregated analytical data from Flurry were tabulated and summary statistics calculated. To assess users’ engagement, the In Hand team were asked to advise on the time a user would need to spend on the app to have a ‘meaningful engagement’: that is, open the app, make a selection of how they were feeling and perform at least one activity based on their response to the front screen (eg, take a photo). Flurry gives data on session length in fixed ranges (see table 1). It was likely that many sessions recorded as less than 30 s would have been too short for the user to have completed an activity. Therefore, user sessions in the range 30–60 s and above were classed as ‘meaningful engagement’.

Survey data were downloaded from SurveyMonkey and imported into a database. Descriptive statistics were calculated using SPSS V22 (Chicago, Illinois, USA). For the demographic description of the sample, the whole dataset is reported (n=131), but for data relating to whether In Hand helped with mental well-being, only data from respondents who ran the app once or more are reported (n=108).

Audio-recorded interviews were transcribed verbatim. A top-down, deductive thematic analysis was taken which focused on: how the user discovered In Hand, their use of In Hand, whether it was helpful, any risks in using In Hand and potential areas for improvement. Each transcript was reviewed and codes assigned to content which reflected the defined areas of interest. These codes were reviewed to identify any overlap between codes and to group them into overarching themes.

FINDINGS
Analytics
From launch of In Hand on 14 May to 31 October 2014, there were 22 357 user sessions on In Hand across both mobile platforms (14 981 on iPhone; 7376 on Android). Seventy-five per cent of these were returning users. Sixteen per cent remained active 1 week after first use (likely to be at installation), 7% after 4 weeks and 2% after 20 weeks. Around half of the users (52%) opened In Hand once a week, with 34% using it 2–3 times per week, 10% 4–6 times and 4% more than six times per week.

Table 1 shows engagement with In Hand measured by the length of time of each user session (data were provided by the analytics in the ranges shown). More than half of users’ sessions (58%) were in the ‘meaningful engagement’ ranges of 30–60 s or longer and a further fifth...
Table 2  Participants ratings of whether In Hand helped with their mental well-being (n=108), ranked in order by numbers endorsing it helped ‘a lot’

<table>
<thead>
<tr>
<th>Would you say In Hand has helped you…</th>
<th>No way, n (%)</th>
<th>Not really, n (%)</th>
<th>Yes, a little bit, n (%)</th>
<th>Yes, a lot, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>More able to take control</td>
<td>4 (3.7)</td>
<td>28 (25.9)</td>
<td>43 (39.8)</td>
<td>33 (30.6)</td>
</tr>
<tr>
<td>Think clearly*</td>
<td>3 (2.8)</td>
<td>25 (23.1)</td>
<td>51 (47.2)</td>
<td>29 (26.9)</td>
</tr>
<tr>
<td>Feel relaxed*</td>
<td>2 (1.9)</td>
<td>23 (21.3)</td>
<td>58 (53.7)</td>
<td>25 (23.1)</td>
</tr>
<tr>
<td>Deal with problems well*</td>
<td>5 (4.6)</td>
<td>37 (34.3)</td>
<td>41 (38.0)</td>
<td>25 (23.1)</td>
</tr>
<tr>
<td>Less stressed</td>
<td>4 (3.7)</td>
<td>20 (18.5)</td>
<td>60 (55.6)</td>
<td>24 (22.2)</td>
</tr>
<tr>
<td>Make my own mind up about things*</td>
<td>6 (5.6)</td>
<td>34 (31.5)</td>
<td>47 (43.5)</td>
<td>21 (19.4)</td>
</tr>
<tr>
<td>Have a positive outlook*</td>
<td>3 (2.8)</td>
<td>17 (15.7)</td>
<td>68 (63.0)</td>
<td>20 (18.5)</td>
</tr>
<tr>
<td>Feel useful*</td>
<td>5 (4.6)</td>
<td>39 (36.1)</td>
<td>49 (45.4)</td>
<td>15 (13.9)</td>
</tr>
<tr>
<td>Ready to talk to someone else</td>
<td>15 (13.9)</td>
<td>38 (35.2)</td>
<td>41 (38.0)</td>
<td>14 (13.0)</td>
</tr>
<tr>
<td>Feel close to other people*</td>
<td>18 (16.7)</td>
<td>43 (39.0)</td>
<td>29 (26.9)</td>
<td>14 (13.0)</td>
</tr>
</tbody>
</table>

*These items are taken from the Short Warwick-Edinburgh Mental Well-being Scale.
The interviewees further expanded on why In Hand was useful to them. They talked about In Hand being discreet, private, and not requiring them to be in specific locations to access it. Interviewees described how the anonymity and perceived non-judgmental nature of a digital tool was important to them. It gave them the ability to think about how they were feeling at any time without having to involve other people:

“There’s no-one to trust on your app – it’s just asking you how you are feeling. There’s no kind of come back or no-one’s going to say anything back.” (Interviewee 1)

**DISCUSSION**

**Strengths of the evaluation approach**

First, we believe the approach to evaluation set out in this paper is useful as it was intended to be proportionate to this type of digital tool and its anticipated impact on health outcomes—a simple, non-clinical tool intended for personal, unsupervised use as one part of an individual’s self-management strategies for their emotional well-being. The evidence generated by this approach, informed by the principles of Health Technology Assessment, has provided quantifiable insights into the app usage and patterns of engagement, an assessment of how the app has supported users with their mental well-being and identified some descriptive insights to how the tool works to support users. The approach has value because it goes beyond user ratings in app stores, while being timely and cost-efficient to implement.

Second, the approach was able to gather data from actual users of the tool in real world settings. By accessing the analytical data of all users within a specified time period, we were able to analyse, in aggregate, how people used the tool and how this changed over time. This may be different (or similar) from how people interact with a tool within a controlled setting: digital interventions in formal trials tend to have greater adherence than naturalistic evaluations.15 Moreover, the survey and interview respondents were people who had selected the tool independently from a range available (via the app stores) and used it in naturally occurring ways, rather than a sample of volunteers using the tool under experimental, controlled conditions.

Third, the evaluation plan (along with the app) was guided by a team of young people and proved achievable within a timescale to fit with the development cycles of digital tools. Evaluation commenced at the end of beta testing when the app first became available from app stores; data and analysis were communicated to the development team for implementation at their 6-month review point. The co-design process assisted in ascertaining that the evaluation methods were understandable and used language and a design familiar to its target audience. The methods adopted did not require extensive digital development—which would have been outside the resources of the development and evaluation teams—and good response rates were achieved. Nor did it require retention of participants over time to generate useful insights.

Fourth, while responses to the survey were good and numbers exceeded our initial target (see figure 3), this probably represents a small proportion of overall users: because usage of the app is recorded in user sessions, rather than individuals, it was not possible to accurately estimate the response rate to the survey.

Third, as figure 3 shows, interest in taking part in the interviews was very limited. In the event, we interviewed all those that agreed and achieved only eight in total: this was lower than our target number and it was not possible to sample from the specified criteria of interest (eg, gender, different experience with the tool). The interview findings are therefore to be interpreted with caution as they are based on a small number of experiences with a homogenous sample. Users of other backgrounds and demographics may have different perspectives on In Hand.

Fourth, the time period between users first interacting with In Hand and the completion of the survey was not recorded. Therefore, an assessment of the influence of recall bias on the reliability of the results is not possible. However, users would have accessed the survey through the hyperlinks within the app, so it is probable that they completed it during an app session.

Finally, the Flurry analytics software is designed for developers, rather than a research tool, which brings limitations. In particular, data were aggregated into predefined numerical ranges which acted to limit the analysis. For example, Flurry categorises each individual user session into time ranges, rather than providing the actual time spent on the app. The precise format of data returned by Flurry depended on the smartphone operating system and did not always align well between these. Moreover, it was not possible to track an individual user’s interactions with the app over time, for example to assess how consistent the usage was. The importance of using a more sophisticated individualised metric combining different aspects of user engagement (an ‘App Engagement Index’) is a current topic in the mHealth literature19 which could be adopted in future research.

**Evaluation of In Hand**

As a result of this evaluation, we can describe how people used In Hand and the nature of its benefit on users’ mental well-being:

---

1. Evid Based Mental Health November 2017 Vol 20 No 4

2. www.org) – which used the same standardised measure in a similar manner.10 This evaluation found that DocReady was reported as useful in different domains of mental well-being to In Hand: the top three rated domains were ‘Able to think clearly’, ‘Ready to talk to someone else’ and ‘More able to take control’, which reflects how DocReady aimed to benefit its users through changing preparatory intentions and behaviour in seeking out help from a GP. As with In Hand, most users rated the help as ‘a little bit’ rather than ‘a lot’, confirming that both these tools have a limited, specific effect and, as would be expected, one mHealth app would not provide all the functions required for overall mental health.16

3. Research participation bias is not clear.
Each interaction with the app was brief, but the majority of interactions were long enough to allow an active interaction with the app.

The majority of users interacted with the app more than once and although use did taper off over time, there was a level of sustained use.

In Hand supported users’ mental well-being through changing attitude and point of view, helping with decreasing feelings of stress and increasing feelings of relaxation and supporting clear thinking.

Users described these supports to their mental well-being were encouraging because: the app prompted users to actively reflect on their current mental state; provided an easily available strategy to ease anxiety and helped build confidence and empowerment in how users coped with their mental health day to day.

These usage patterns are similar to those seen in other studies of digital interventions, especially the low retention rate for many, but sustained use for some, which is similar to the 1-9-90 rule observed in online forums. This rule suggests that in participating in online communities, the majority (90%) are passive users (‘lurkers’ who observe but do not actively participate in the online community), a minority (9%) are occasional contributors, and an even smaller proportion (1%) are the biggest participants who are responsible for the majority of contributions to the online community. While other study designs can identify whether an effect is achieved, for example, increased emotional self-awareness, the methods adopted in this evaluation provide an understanding of how users perceive the benefits of a digital tool. The next stage in our proposed proportionate approach would be a closer examination which would explore any adverse effects, as well as further evaluation of value and benefit, if uptake of the tool proved sufficient to warrant it.

CLINICAL IMPLICATIONS

This paper has demonstrated how simple, self-help digital tools can be evaluated in a proportionate and practical way. Triangulating data sources provided an understanding of how In Hand was used and the ways it can support mental well-being. In particular, the survey of naturally occurring users provided insights into the value of the app from the user perspective and, in this case provided evidence of the app having the intended effect for users. This is important for healthcare decision-makers who need to be assured of the quality of an app before recommending to patients. In addition, analytical data provides evidence of how and when a tool is being used (or not), which enables health providers and commissioners to make a judgement on the value of a tool in order to ensure any cost is justified (not applicable for in this case as the tool is free to use). Furthermore, we have demonstrated cost-efficient, timely methods, which can be easily incorporated into digital tools, thus providing scope for audits and service evaluations.

Acknowledgements

We would like to thank the In Hand team and their young persons’ panel for their support and all the participants for their input and time. The research reported in this paper was conducted by the NIHR MindTech HTC and supported by the NIHR Nottingham BRC. The views represented are the views of the authors alone and do not necessarily represent the views of the Department of Health in England, NHS, or the National Institute for Health.

Contributors

MPC, JLM and LS conceived the study, collected and analysed the data and interpreted the results. All authors drafted and revised the manuscript and gave final approval of the version submitted for publication.

Funding

This study was funded by Comic Relief as part of Innovation Labs (a consortium of Comic Relief, Mental Health Foundation, Paul Hamlyn Trust and Nominet Trust).

Competing interests

None declared.

Provenance and peer review

Not commissioned; externally peer reviewed.

REFERENCES

Proportionate methods for evaluating a simple digital mental health tool

E Bethan Davies, Michael P Craven, Jennifer L Martin and Lucy Simons

Evid Based Mental Health 2017 20: 112-117 originally published online October 9, 2017
doi: 10.1136/eb-2017-102755

Updated information and services can be found at:
http://ebmh.bmj.com/content/20/4/112

These include:

References
This article cites 18 articles, 1 of which you can access for free at:
http://ebmh.bmj.com/content/20/4/112#BIBL

Open Access
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Open access (17)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/