

e-Therapies in England for stress, anxiety or depression: how are apps developed? A survey of NHS e-therapy developers

Matthew Russell Bennion,¹ Gillian E Hardy,¹ Roger K Moore,² Stephen Kellett,¹ Abigail Millings¹

To cite: Bennion MR, Hardy GE, Moore RK, *et al.* e-Therapies in England for stress, anxiety or depression: how are apps developed? A survey of NHS e-therapy developers. *BMJ Health Care Inform* 2019;**26**:e100027. doi:10.1136/bmjhci-2019-100027

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjhci-2019-100027>).

Accepted 16 May 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Psychology, The University of Sheffield, Sheffield, UK

²Department of Computer Science, The University of Sheffield, Sheffield, UK

Correspondence to

Dr Matthew Russell Bennion; m.bennion@sheffield.ac.uk

ABSTRACT

Objective To document the quality of web and smartphone apps used and recommended for stress, anxiety or depression by examining the manner in which they were developed.

Design The study was conducted using a survey sent to developers of National Health Service (NHS) e-therapies.

Data sources Data were collected via a survey sent out to NHS e-therapy developers during October 2015 and review of development company websites during October 2015.

Data collection/extraction methods Data were compiled from responses to the survey and development company websites of the NHS e-therapies developers.

Results A total of 36 (76.6%) out of the 48 app developers responded. One app was excluded due to its contact details and developer website being unidentifiable. Data from the missing 10 was determined from the app developer's website. The results were that 12 out of 13 web apps and 20 out of 34 smartphone apps had clinical involvement in their development. Nine out of 13 web apps and nine out of 34 smartphone apps indicated academic involvement in their development. Twelve out of 13 web apps and nine out of 34 smartphone apps indicated published research evidence relating to their app. Ten out of 13 web apps and 10 out of 34 smartphone apps indicated having other evidence relating to their app. Nine out of 13 web apps and 19 out of 34 smartphone apps indicated having a psychological approach or theory behind their app.

Conclusions As an increasing number of developers are looking to produce e-therapies for the NHS it is essential they apply clinical and academic best practices to ensure the creation of safe and effective apps.

INTRODUCTION

While apps can have the potential to give great benefits, they also have the potential to cause physical, mental, reputational or financial harm to patients, healthcare professionals and their organisations if they are not evaluated for clinical safety. For example, an app may miscalculate a drug dose or give incorrect medical advice to a consumer or patient.

Summary

What is already known?

- There are 48 known NHS e-therapies that are used/recommended.
- The landscape for assessing apps is complex and ever-changing.

What does this paper add?

- A comprehensive analysis of the development processes behind the e-therapies used and recommended for common mental health problems across the National Health Service (NHS) in England.
- A list of areas that most need refinement to improve e-therapy suitability for NHS mental health services.

National Health Service (NHS) digital highlights that apps to be used by the NHS cannot be endorsed unless they have been evaluated for potential harm,¹ while Public Health England's (PHE's) health app assessment process requires developers to outline plans and policies to limit and mitigate potential risks associated with their apps. This paper is concerned with web and smartphone apps designed to offer treatment or support with the common mental health problems of depression, anxiety and stress—collectively herein termed 'e-therapies', which are used or recommended by the NHS in England. It is important to note that the landscape of e-therapies shifts rapidly, and indeed has done so since the data presented here were collected 4 years ago.

Regulatory approval provides patients and healthcare professionals with the assurance that an app is of high quality, safe and ethical.¹ There are two types of regulation presently available in the UK: the Medicines and Healthcare products Regulatory Agency (MHRA) Medical Device Registration and Care Quality Commission (CQC) Registration. Both are relevant to e-therapies. MHRA

provides a 'device determination' flowchart that enables developers to check whether their app is defined as a medical device. The two main questions in determining this are whether the app has a medical purpose and whether it works directly with data obtained *in vivo*. At the time of writing, developers of apps that meet these criteria and who want to market them to the public are required by UK regulation to register the app with the MHRA and to obtain for it a Conformité Européenne (CE) marking, indicating conformity with health, safety and environmental protection standards for products sold within the European Economic Area.²

CQC set out 14 regulated activities: personal care; accommodation for people who require nursing or personal care; accommodation for people who require treatment for substance misuse; treatment of disease, disorder or injury; assessment or medical treatment for persons detained under the Mental Health Act 1983; surgical procedures; diagnostic and screening procedures; management of supply of blood and blood-derived products; transport services; triage and medical advice provided remotely; maternity and midwifery services; termination of pregnancies; services in slimming clinics; nursing care and family planning services. If an app provides a health or social care service that fits one of these activities the developers are required by the PHE to register with the CQC before the app can be accessed via the PHE app assessment process.

It is essential for the public to know whether an e-therapy is effective. Some have argued that many apps have no evidence to support their effectiveness,³ but deciding what constitutes 'evidence' for apps is not straightforward. Within healthcare research, there is a hierarchical structure depicting the strength of evidence.⁴ The higher the level, the greater the internal validity and hence the more persuasive and trustworthy the evidence is. The randomised controlled trial (RCT) is currently the gold standard for providing evidence of clinical efficacy.⁴ However, RCTs take time to design, implement and publish and thus are poorly matched to the pace at which technologies and tools are evolving. This means that there is presently no clear consensus on how best to evaluate apps, although policymaker and researcher efforts are being directed to the issue, as outlined below.

An European Union Working Group was set up in February 2016 to create mHealth assessment guidelines but unfortunately failed to reach a conclusion.⁵ A report by the group highlighted that building the guidelines had been found to be a much more complex exercise than initially expected at the beginning of the process, and the work required went far beyond the original mandate of the group.⁶

Separately to this, a toolkit for appraising e-therapies was developed and released by MindTech in October 2017. The toolkit offers a standard set of criteria for evaluating existing digital mental health tools (apps and mobile websites) and a final report discussing the framework was published.⁷ Other examples of app assessment

methods that have been developed in recent years include the Mobile App Rating Scale (MARS), developed by an Australian research team in 2015.⁸ MARS is a scale that aims to provide researchers, clinicians and developers with a way to score digital tools based on a list of evaluation criteria.⁸ Item 19 of the scale regarding clinical evidence was ignored due to researchers having yet to test the impact of the mental health apps included in the study.⁸ Similarly, the British Standards Institution in conjunction with Innovate UK has developed the PAS 277:20 159 code of practice. The PAS recommends during the preliminary stages of app development that developers read academic research to ensure that their app is built on clinical evidence. It also recommends that app publishers/developers should collect data during testing with users to validate any clinical benefits that the app's intended use delivers⁹; such an exercise would likely require the involvement of academic researchers to ensure the evidence was of a sufficiently high standard. As well as helping to design new tools, the code can be used to evaluate existing ones.⁹

In October 2017, PHE released a health app assessment process developed to encourage the creation of effective health apps and to enable health professionals to consider health apps for use in General Practice.¹⁰ The process covered eight different areas: evidence of effectiveness; regulatory approval; clinical safety; privacy and confidentiality; security; usability and accessibility; interoperability; and finally, technical stability.

More recently, NHS digital have introduced a Beta Digital Assessment Questionnaire (DAQ) 1.2 for the assessment of mobile apps¹¹ and the National Institute for Health and Care Excellence (NICE), along with their partner organisations have published a set of evidence standards for digital health technologies which includes apps.¹² The evidence standards have been developed to ensure new technologies are clinically effective and represent value for money to the NHS, while also aiming to make it easier for innovators and commissioners to understand what good levels of effectiveness for digital technologies should look like. NHS Digital is working closely with NICE to incorporate these standards into future versions of the DAQ.¹¹

Current study

It is apparent that the landscape for assessing apps is complex and ever-changing. The aim of the present study was to examine the quality of apps in use by the NHS by examining the manner in which they have been developed. At the time of this study, the majority of existing app review methods either focused on the technical rules and regulations of app design and overlooked (often by necessity, MARS⁸) the question of effectiveness, that is, whether the actual app does what it says and meets the claims its developers make for it. While this has been rectified in the DAQ,¹¹ the current research precedes this. In an ideal world, before being released for general use, every app would have undergone rigorous user trials that

demonstrated its effectiveness. However, rigorous trials are a costly and challenging business, and current models of app development and publishing seem to encourage less rigorous approaches. Here, in an attempt to gauge the quality of existing apps as providers of therapy without performing user trials on each and every one, we have adopted the approach of probing more deeply into the processes employed in their development. More specifically, we are interested in the psychological model, theories, or therapies used, the extent of clinical and academic involvement, and any published (or otherwise) evidence in support of each app. The developers of each of the apps identified in our previous study¹³ were contacted and asked to provide the relevant details. It is worth noting that the quality assessment frameworks detailed above did not precede the development of many of the apps reviewed here, so developers were likely operating in a quality assessment vacuum at the time of creating their products.

Important indicators of quality

The purpose of this study was to evaluate the apps that were previously identified as being used or recommended by the NHS.¹³ We were specifically interested in the following four indicators of quality: clinician involvement, academic involvement, research or other evidence and use of a specific psychological approach or theory. These indicators were selected because they build on the premises that effective digital psychotherapy interventions come about as a result of rigorous theoretical and empirical work by experienced clinicians and academics, utilising a known psychological approach. We discuss the advantages of each below.

Clinician involvement

Healthcare staff routinely use apps to perform their roles.¹⁴ This makes it essential that the information given in these apps be grounded in the best and most up-to-date knowledge, derived from research, clinical experience and patient preference.¹ Unfortunately, many app stores do not carry out rigorous reviews regarding the accuracy of app content before publication, meaning some apps potentially have inaccurate information.¹⁵ Other publications have highlighted that when assessing digital mental health apps, it is important to assess whether clinicians have been involved in the development process.^{16 17} This is because clinician involvement can help to ensure that any established modes of treatment are appropriately deployed within the app. For instance, an app based on cognitive behaviour therapy (CBT) but made by someone who is not qualified to deliver CBT may fail to give an accurate implementation. The involvement of a clinician who specialises in CBT would improve the quality of app content by ensuring treatment fidelity.

Academic involvement

Academic involvement in the process of developing an app can help to ensure the implementation of empirically

supported interventions and principles, providing a foundation for an app's use in clinical practice. Responsible academics strive to bring neutrality and remove bias, to expose the app to peer review and publish evidence of an app's feasibility, acceptability and clinical effectiveness.

As mentioned previously it is essential that an app can show evidence of its effectiveness. In PHE's app assessment process, developers must provide evidence that their app improves outcomes for patients and users; provides value for money; meets user needs and is stable and simple to use, and that people use it. Independent research is weighted highly in the assessment criteria, and apps that have a high level of clinical evidence are considered by NICE for 'NICE evaluated' status. This status is considered to represent the gold standard for NHS health apps. In addition to this, all apps are required to show that they meet the criterion set out by NHS Digital covering: clarity of purpose and intended use; their evidence basis; the data that forms the basis their evidence and findings; any published academic studies.¹ The involvement of academics in the development of an app can be helpful in ensuring that data are collected in a manner that makes it possible to evaluate effectiveness, although it is also important that evaluation is conducted by researchers independent of the app, without a personal interest in the results.

Research evidence/other evidence

While RCTs are the gold standard, it is not expected that all apps will have published research evidence at the time of writing in part due to the rapid pace of change and the unwieldy nature of conducting RCTs. However, there might be other forms of evidence that can indicate whether an app may be beneficial to a patient. This evidence may take different forms such as practice-based evidence methodologies (eg, detailed case series) that assess the acceptability, feasibility and initial effectiveness of an app and may also include early pilot trials.

Specific psychological approach or theory/set of techniques/therapy

Apps claiming to help with mental health problems such as depression, anxiety or stress would be expected to use established approaches to treatment that have been found to be effective through high quality research studies. The psychological therapies that are designated by NICE regardless of disorder are all underpinned by a clinical theory. The risk of not having an organising theoretical framework for an app is that the change techniques that are used may be cherry-picked by developers on the basis of inappropriate criteria (eg, selecting techniques that can be gamified easily rather than those that are most effective) and so lack theoretical coherence and consistency.

Using indicators of quality to evaluate apps

If we accept the premise that effective psychotherapeutic interventions only come about as the result of rigorous

Table 1 All NHS e-therapy app survey responses by app

App	Web or phone based (w/p)	Clinical involvement	Academic involvement	Publications published or forthcoming	Other evidence	Psychological approach or theory
Aventurine Mood Improver	P	N	N	Y*	Y†	Y
Beating the Blues	W	Y	Y	Y	Y	Y
Big White Wall	W	Y	N	Y‡	Y‡	Y
Black Rainbow	P	N	N	N	N	N
Buddy App	W	Y	N	Y	Y§	N¶
Depression Calculator**	P	Y	N	N	N	N
DigitalMeds	P	N	N	Y*	N	Y
Don't Panic!**	P	Y	N	N	N	Y
Fear Fighter	W	Y	Y	Y	Y	Y
Five Ways to Wellbeing	P	Y	N	N	Y††	Y
Ginsberg App	P	Y	Y	N	N	Y
Happy Healthy App	P	Y	Y	Y	Y	N
Headspace	P	N	N	Y	Y	Y
Healthstored	P	N	N	N	N	N
Healthy Living**	P	Y	N	N	N	N
Hello Brain Health	P	Y	Y	Y‡‡	Y	Y
How Are You App	P	N	Y	N	N	Y
iCBT**	P	Y	N	N	N	Y
IESO Digital Health	W	Y	Y	Y	Y	Y
Kooth	W	Y	Y	Y*	Y	N¶¶¶
Living Life to the Full	W	Y	N	Y	N	Y
Mind Ed	W	Y	Y	Y***	N	N
Mindfulness Bell**	P	N	N	N	N	N
Mindfulness by Digipill**	P	Y	N	N	N	Y
Mindlogr	P	N	N	N	N	Y
Mindshift	P	Y	N	N	Y	Y
Moodbug	P	N	N	N	N	Y
MoodGym	W	Y	Y	Y	Y	Y
Moodkit—Mood Improvement Tools	P	Y	Y	Y	Y†††	Y
Moodometer	P	Y	N	N	N	Y
Moodscope	W	Y‡‡‡	Y	Y	Y	Y
MyMoodTracker	P	N	N	N	N	N
Panic Attack Aid**	P	N	N	N	N	N
Phobia Free	P	Y	Y	Y§§§	Y¶¶¶¶	Y
Puffelf	W	N	N	N	N	N
SAM: Self-help for Anxiety	P	Y	Y	Y	N	N
SilverCloud Health	W	Y	Y	Y	Y	Y
Stay Alive	P	Y	N	N	N	N

Continued

Table 1 Continued

App	Web or phone based (w/p)	Clinical involvement	Academic involvement	Publications published or forthcoming	Other evidence	Psychological approach or theory
Stress & Anxiety Companion	P	Y	N	N	N	Y
Stress Management App	P	Y	Y	N	Y****	Y
Take a break!	P	Y	N	N	N	N
Thought Diary**	P	N	N	N	N	N
Thought Diary Pro**	P	N	N	N	N	N
Virtual Hope Box	P	Y	Y	Y	Y	Y
WellMind	P	Y	N	N	N	N
WorkGuru	W	Y	Y	Y††††	Y‡‡‡‡	Y
Worry Watch**	P	N	N	N	N	N

*References studies related to the therapy on which the app was based, rather than the app itself.

†The developer assessed CBT apps in a non-scientific way.

‡Studies were not published at the time of the survey but were in preparation.

§Piloted within 12 NHS Trusts across a range of services and teams.

¶Designed to work with a behavioural activation framework.

**Failed to respond to survey. Answered by the researcher based on publicly available information.

††Had a testing/pilot phase mainly to see if the elements worked and what people thought about the app.

‡‡No publications but at very early stages of two projects that will produce independent research evidence.

§§PhD student rather than experienced research academic.

¶¶¶Declared no approach, but on inspection of the developer's website, it was found that the service offers one-to-one CBT over a messenger.

***Comprehensive systematic review of e-therapies.

†††Currently being evaluated in a large controlled efficacy trial.

‡‡‡Adaption of the PANAS test.

§§§Research carried out independently by Roehampton University and being prepared for submission.

¶¶¶¶Feasibility study to test acceptability, presented at Royal College Psychiatry 2013 International Congress.

****Piloted to ensure the UX was appealing/useful, and so on but no formal studies.

††††PhD researcher is planning research looking at how to increase engagement with online workplace therapies.

‡‡‡‡2 Non-academic studies without a control group. The developer is currently planning an randomised controlled trial.

CBT, cognitive behaviour therapy; NHS, National Health Service; PANAS, Positive and Negative Affect Schedule; UX, user experience.

theoretical and empirical work by experienced clinicians and academics, it follows that apps need clinician and academic involvement, psychological theory and research/other data to support their effectiveness. We have previously collated a list of NHS endorsed e-therapies (meaning therapeutic apps (both phone and web) that are used or recommended in NHS settings in England)¹³ designed to target stress, depression, or anxiety. In the current study, we evaluate these NHS e-therapies for compliance with the indicators of quality

described above. To do so, we surveyed the developers of all the apps identified in our previous study¹³ regarding the key indicators of quality described, and whether there were any differences between web and phone apps.

METHODS

Design

We documented development information surrounding each of the web and smartphone apps used and recommended by the NHS for stress, anxiety and depression identified in our previous study.¹³ Our data source was a survey sent to developers. In the survey, we asked for information about clinical involvement, academic involvement, publications published or forthcoming, other evidence of effectiveness and whether the app was based on a psychological approach or theory.

Participants

Participants were the developers of the 48 NHS e-therapies comprised 13 web and 35 smartphone apps. These apps were identified in our previous study in which data

Table 2 Summary of key findings by number (N) and percentage (%) where yes was the response given

	Web	Smartphone	Overall
Clinical involvement	12 (92%)	20 (59%)	32 (68%)
Academic involvement	9 (69%)	9 (27%)	18 (38%)
Research evidence	12 (92%)	9 (27%)	21 (45%)
Other evidence	10 (77%)	10 (29%)	20 (43%)
Psychological approach or theory	9 (69%)	19 (56%)	28 (60%)

regarding web and smartphone apps used or recommended were compiled from responses to: (1) freedom of information requests sent to all Improving Access to Psychological Therapies (IAPT) services and NHS Mental Health Trusts in England; and (2) NHS apps library search results.¹³

Procedure

Ethical approval was gained from the University of Sheffield Psychology Ethics Committee. During October 2015, each e-therapy developer was initially contacted by telephone before being sent an email invitation to be involved in a study. Those who expressed interest were sent an information sheet and consent form. Those who consented were sent a link to the online survey covering the five key areas of evidence for the quality of an app, namely: the extent of clinical involvement in its development; the existence of published evidence; the extent of academic involvement; any studies or trials which have been undertaken; any psychological approach or theory on which the app is based. The questions asked are reported in online supplementary table S1. Each initial question had a further open-ended subquestion asking for more details when a participant confirmed the existence of that particular evidence. Participants were then debriefed.

Statistical analysis

In order to establish whether the answers to these questions differed as a function of app type (web or phone), tests of association were used. A series of χ^2 tests were conducted for questions 2, 3, 4 and 5. Data from question 1 did not meet all assumptions for a χ^2 , so a Fisher's exact test was conducted instead.

RESULTS

Thirty-six (75%) of app developers responded to the survey. One app was excluded due to its contact details and developer website being unidentifiable. Data for the missing 10 were determined by using information on each developer's website relating to their web/smartphone app. Table 1 details, for each of the apps, the extent and nature of their indicators of quality, as derived from the responses of the developers to the survey or else from the interpretation of the apps' websites, and table 2 shows a summary of key findings relating to the questions asked.

Clinical involvement

Overall 32 (68%) developers had clinical involvement in the development of their app. According to a Fisher's exact test, there was a statistically significant association between app type and clinical involvement, $p=0.037$. The result suggests that web apps may have more clinical involvement than phone apps.

Academic involvement

Overall 18 (38%) developers had academic involvement in the development of their app. There was a statistically

significant association between app type and academic involvement, $\chi^2(1)=7.277$, $p=0.007$. There was a moderately strong negative association between app type and academic involvement, $\phi=-0.393$, $p=0.007$. The result suggests that web apps may have more academic involvement than phone apps.

Publications published or forthcoming

Overall 21 (45%) developers had publications published or forthcoming regarding their app. There was a statistically significant association between app type and whether there were any scientific articles published or forthcoming, $\chi^2(1)=16.492$, $p<0.000$. There was a strong negative association between app type and articles published or forthcoming, $\phi=-0.592$, $p<0.000$. The result suggests that web apps may have more papers published or forthcoming than phone apps.

Other evidence

Overall 20 (43%) developers had other evidence regarding their app. There was a statistically significant association between app type and other evidence, $\chi^2(1)=8.684$, $p=0.003$. There was a strong negative association between app type and other evidence, $\phi=-0.430$, $p=0.003$. The result suggests that web apps may have more other evidence than mobile apps.

Psychological approach or theory

Overall 28 (60%) developers used a psychological approach or theory in their app. There was a non-significant association between app type and psychological approach or theory, $\chi^2(1)=0.696$, $p=0.404$. The result suggests that no difference in the use of psychological approaches or theories between web and phone apps.

DISCUSSION

The present study documented self-reported data regarding development information surrounding each of the web and smartphone apps used and recommended by the NHS for stress, anxiety and depression as identified in our previous study.¹³ Our data source was a survey sent to developers requesting information about clinical involvement, academic involvement, publications published or forthcoming, other evidence of effectiveness and whether the app was based on a psychological approach or theory. In the absence of a response, we turned to websites. The purpose of doing this was to attempt to gauge the quality of the existing apps as providers of therapy without performing user trials on every one. The approach adopted has enabled a deeper examination of the processes that are employed in each app's development.

The data presented indicates a significant disparity between web and phone-based applications. Web applications are indicated as being significantly higher in all areas of quality except for psychological approach. This gap in quality measures may stem historically from how the two technologies have evolved. Web applications for

mental health such as Beating the Blues¹⁸ originated in academia in the 2000s before being brought to market. Mobile apps began to appear in 2008 when app stores began to open.¹⁹ However, the process of developing mobile apps for mental health took a drastically different route to market that had more in common with the early 1980s personal microcomputer boom²⁰ whereby anyone with access to a personal microcomputer and a programming book could publish their own software. Poor regulation of mental health app quality and privacy²¹ on the one hand, and dramatic growth of innovation in the digital space on the other,²² has meant mobile apps have taken a similar path with home coders being able to create and publish mental health apps with minimal questioning of their quality.

The varying quality between the two app types is supportive of the issues and solutions discussed in the introduction to this study. However, even with the slow introduction of regulation, there still seems to be a limited amount of action to regulate mental health apps coming from the app stores themselves. Responsibility is being placed on app developers to get regulated, for users to check apps are regulated and for governments and health organisations such as the NHS to flag up apps that are regulated. A number of studies have examined the potential dangers and safety issues associated with apps aimed at the public and whether these apps should be assessed or controlled by various regulatory agencies, such as the MHRA in England.²³

A systematic review by Donker *et al.*,²⁴ published 6 years ago supports the efficacy of mental health apps across all ages, but revealed a lack of published evidence for many e-therapy smartphone apps that were available. The results of the current survey have revealed a very similar trend to Donker's findings, with only 27% of smartphone apps surveyed having any evidence of effectiveness published or forthcoming.

The survey results indicate web apps fared much better with 92% of app designers saying that their app had evidence of effectiveness. However, overall 44.7% of apps had published or forthcoming evidence of effectiveness. The lack of a consistent evidence base makes the process of finding an effective NHS e-therapy haphazard. During 2018 the NHS rolled out a digital apps library designed to showcase a selected number of apps that have been through assessment and are safe to use.²⁵ The 'Mental Health' section of the library contains 18 apps, four of those apps appear in this study and account for 22.22% of the section: Big White Wall; Ieso; SilverCloud and Stress and Anxiety Companion.

It is vital that the NHS ensure this library contains apps that are supported by evidence, so service users are presented with effective apps. For a considerable amount of time NICE guidance for mental health practitioners was that computerised cognitive behaviour therapy (cCBT) could be offered for persistent subthreshold, or mild to moderate depression.²⁶ However, this guidance was withdrawn in July 2018 to allow OCFighter to be considered

for an IAPT assessment briefing.²⁶ Even with this withdrawal a strong emphasis is still placed on the use of CBT within mental health services. The results of the survey indicated that at that time 60% of apps used a known psychological approach or theory. On further analysis, 71% of the 60% used CBT, meaning 14% of those that declared using a psychological approach/theory were using an approach that was not NICE-approved, and 40% had no theory behind them at all.

Importantly, we have seen from this study that different criteria can be applied to evaluating apps. Regarding research evidence, developers have different interpretations of this. Some felt that findings from third-party studies of the conventional, face-to-face delivery of the therapeutic model used within the app would translate into an indication of the app's effectiveness, while others considered internal studies, evaluations and testing of elements such as the user interface to be a sufficient indication of app effectiveness.

This study is not without limitation, and there are areas which were not assessed, due to the need for brevity. However, future research would be well-advised to consider them. For example, questions querying service user and computer scientist involvement could have further expanded on the indicators of quality used within this study and opened further avenues of investigation. Assessment criteria that were released after this study was carried out such as NHS Digital's assessment questionnaire V.2.1¹¹ investigate service user involvement by asking questions around the use of user centred design. Approximately 120 apps have been evaluated using the questionnaire.²⁷ There are other potentially important areas that were unexplored in this study, such as the frequency of updates of the apps.²⁸ Finally, the nature of data collection (brief self-report and website review) may have inherent limitations on the depth and quality of the information available. Definitions were not provided to respondents regarding what did and did not qualify as 'academic involvement', for example, and there may have been variation in the way that this was interpreted. One developer felt that the involvement of a PhD student met this criteria, but a research student versus senior a member of faculty have very different levels of expertise. Future research could seek to rectify this, by defining explicit categories of involvement and levels of expertise. Future research could also examine not only whether clinicians, academics, computer scientists and service users were involved in the development process, but also at what stage. It would be reasonable to hypothesise that early involvement here may result in a better product, as there would be more iterative design cycles into which relevant expertise could feed.

CONCLUSIONS

As an increasing number of developers are looking to produce e-therapies for the NHS it is essential they apply clinical and academic best practices to ensure the creation

of safe and effective apps. The present study has provided a snapshot of the commercial e-therapy development landscape and the areas that need the most refinement to improve their suitability for NHS mental health services. These are as follows:

1. The building of better working relationships with clinicians to ensure e-therapy content is both beneficial and credible.
2. Creation of commercial app development projects with academics to enable increased innovation and to feed lessons learnt back to the research knowledge base.
3. Generation of published evidence through studies in areas such as piloting, feasibility, acceptability, effectiveness and efficacy to help enable easy distinction of the effectiveness of an e-therapy.
4. Collection, analysis and publication of other evidence to back up claims of effectiveness.
5. Ensuring that a psychological approach or theory that is known to be empirically sound is applied within an app to help build credibility.

It should be noted that while the alternative methods of evaluation discussed in this study are relevant and timely, RCTs remain the pinnacle of quality evidence of effectiveness. The best way to synthesise such evidence is through meta-analysis. As such, future work should extend the current investigation beyond this initial developer survey, to synthesise the published evidence available for all e-therapies used and recommended by the NHS in England, thus establishing NHS delivered e-therapy effectiveness.

Contributors MRB and AM conceived of and designed the research. MRB collected and analysed the data. MRB interpreted the results and drafted the manuscript, and AM, GEH, SK and RKM revised it. All authors approved the final version of the article. All authors had access to all study data and take responsibility for data integrity and accuracy of the analysis.

Funding This work was supported by a PhD studentship awarded by the University of Sheffield to the first author, and Economic and Social Research Council grant number ES/L001365/1.

Disclaimer The content is solely the responsibility of the authors and does not represent the views of the NHS England.

Competing interests The last author was formerly an employee (2010–2012) and minor shareholder of Ultras UK (makers of 'Beating the Blues'), which went into administration in October 2015.

Patient consent for publication Not required.

Ethics approval University of Sheffield Psychology Ethics Committee. Reference: 005047.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All data relevant to the study are included in the article or uploaded as supplementary information.

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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. NHS Digital. Digital Assessment Questions - Beta, 2017. Available: <https://developer.nhs.uk/digital-tools/daq/> [Accessed 19 Feb 2018].
2. gov.uk. Department for Business Energy & Industrial Strategy. CE marking, 2012. Available: <https://www.gov.uk/guidance/ce-marking> [Accessed 24 Feb 2018].
3. Leigh S, Flatt S. App-based psychological interventions: friend or foe?. *Evid Based Mental Health* 2015;18:97–9.
4. Joanna Briggs Institute. The JBI approach. Joanna Briggs InstT, 2014. Available: <http://joannabriggs.org/jbi-approach.html#tabbed-nav=Levels-of-Evidence> [Accessed 24 Oct 2018].
5. European Commission. Report of the Working Group on mHealth assessment guidelines | digital single market, 2017. Available: <https://ec.europa.eu/digital-single-market/en/news/report-working-group-mhealth-assessment-guidelines> [Accessed 7 Apr 2018].
6. Working Group on mHealth Assessment Guidelines. Report of the Working Group on mHealth assessment guidelines; 2017: 1–10. <https://ec.europa.eu/digital-single-market/en/news/report-working-group-mhealth-assessment-guidelines>
7. Betton V, Craven M, Davies B, et al. Framework for the effectiveness evaluation of MOBILE (mental) health tools, 2017. Available: <https://s3-eu-west-1.amazonaws.com/assets.waremhabitat.com/Framework-for-the-effectiveness-evaluation-of-mobile-mental-health-tools.pdf> [Accessed 7 Apr 2018].
8. Stoyanov SR, Hides L, Kavanagh DJ, et al. Mobile app rating scale: a new tool for assessing the quality of health mobile Apps. *JMIR mHealth uHealth* 2015;3:e27.
9. BSI. Health and wellness apps - Quality criteria across the life cycle - Code of practice (PAS 277:2015) 2015;36.
10. Public Health England. Health APP development: the APP assessment process. gov.uk, 2017. Available: <https://www.gov.uk/government/collections/health-app-development-creating-apps-for-health-professionals> [Accessed 10 Feb 2018].
11. NHS Digital. Digital Assessment Questionnaire V2.1, 2018. Available: <https://www.england.nhs.uk/nhsid> [Accessed 6 Feb 2019].
12. NICE. Evidence standards framework for digital health technologies 2019.
13. Bennion MR, Hardy G, Moore RK, et al. E-therapies in England for stress, anxiety or depression: what is being used in the NHS? A survey of mental health services. *BMJ Open* 2017;7:e014844.
14. O'Connor P, Byrne D, Butt M, et al. Interns and their smartphones: use for clinical practice. *Postgrad Med J* 2014;90:75–9.
15. Barton AJ. The regulation of mobile health applications. *BMC Med* 2012;10.
16. Hamilton AD, Brady RRW. Medical professional involvement in smartphone 'apps' in dermatology. *Br J Dermatol* 2012;167:220–1.
17. O'Neill S, Brady RRW. Colorectal smartphone apps: opportunities and risks. *Color Dis* 2012;14:e530–4.
18. Proudfoot J, Swain S, Widmer S, et al. The development and beta-test of a computer-therapy program for anxiety and depression: hurdles and lessons. *Computers in Human Behavior* 2003;19:277–89.
19. Cuadrado F, Dueñas J. Mobile application stores: success factors, existing approaches, and future developments. *IEEE Commun. Mag.* 2012;50:160–7.
20. Lean T. Mediating the microcomputer: the educational character of the 1980s British popular computing boom. *Public Underst Sci* 2013;22:546–58.
21. Chandrashekar P. Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. *Mhealth* 2018;4.
22. Hill C, Martin JL, Thomson S, et al. Navigating the challenges of digital health innovation: considerations and solutions in developing Online and smartphone-application-based interventions for mental health disorders. *Br J Psychiatry* 2017;211:65–9.
23. Cummings E, Borycki EM, Roehrer E. Issues and considerations for healthcare consumers using mobile applications. *Stud Health Technol Inform* 2013;183:227–31.
24. Donker T, Petrie K, Proudfoot J, et al. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res* 2013;15:e247.
25. Bauer J, Murphy R. Apps library is advance for a digital NHS. NHS Engl, 2017. Available: <https://www.england.nhs.uk/blog/apps-library-is-advance-for-a-digital-nhs/> [Accessed 10 Feb 2018].
26. National Institute for Health and Care Excellence. Depression in adults: recognition and management. CG90, 200963. Available: <https://www.nice.org.uk/guidance/cg90/> [Accessed 9 Sep 2016].
27. Postelnicu L, MobiHealthNews. More than 100 digital health and care tools evaluated against key standards for NHS Apps library | mobihealthnews. Available: <https://www.mobihealthnews.com/content/more-100-digital-health-and-care-tools-evaluated-against-key-standards-nhs-apps-library> [Accessed 9 May 2019].
28. Zhang MWB, Ho RCM, Loh A, et al. Current status of postnatal depression smartphone applications available on application stores: an information quality analysis. *BMJ Open* 2017;7:e015655.