



Evaluating the quality and safety of health-related apps and e-tools: Adapting the Mobile App Rating Scale and developing a quality assurance protocol

Anna E. Roberts, Tracey A. Davenport, Toby Wong, Hyei-Won Moon, Ian B. Hickie, Haley M. LaMonica*

Brain and Mind Centre, The University of Sydney, Sydney, NSW, Australia

ARTICLE INFO

Keywords:

Mental health
Mobile health
Technology
Apps
Quality assurance
Digital tools

ABSTRACT

Background: Whilst apps and e-tools have tremendous potential as low-cost, scalable mental health intervention and prevention tools, it is essential that consumers and health professionals have a means by which to evaluate their quality and safety.

Objective: This study aimed to: 1) adapt the original Mobile App Rating Scale (MARS) in order to be appropriate for the evaluation of both mobile phone applications as well as e-tools; 2) test the reliability of the revised scale; and 3) develop a quality assurance protocol for identifying and rating new apps and e-tools to determine appropriateness for use in clinical practice.

Methods: The MARS was adapted to include items specific to health-related apps and e-tools, such as the availability of resources, strategies for self-management, and quality information. The 41 apps and e-tools in the standard youth configuration of the InnoWell Platform, a digital tool designed to support or enhance mental health service delivery, were independently rated by two expert raters using the A-MARS. Cronbach's alpha was used to calculate the internal consistency and interclass correlation coefficients were used to calculate interrater reliability.

Results: The A-MARS was shown to be a reliable scale with acceptable to excellent internal consistency and moderate to excellent interrater reliability across the subscales. Given the ever-increasing number of health information technologies on the market, a protocol to identify and rate new apps and e-tools for potential clinical use is presented.

Conclusions: Whilst the A-MARS is a useful tool to guide health professionals as they explore available apps and e-tools for potential clinical use, the training, time, and skill required to use it effectively may be prohibitive. As such, health professionals and services are likely to benefit from including a digital navigator as part of the care team to assist in selecting and rating apps and e-tools, increasing the usability of the data, and technology troubleshooting. When selecting, evaluating and/or recommending apps and e-tools to consumers, it is important to consider: 1) the availability of explicit strategies to set, monitor and review SMART goals; 2) the accessibility of credible, user friendly information and resources from reputable sources; 3) evidence of effectiveness; and 4) interoperability with other health information technologies.

1. Introduction

1.1. Improving health care delivery using health information technologies

With traditional in-clinic and online mental health care services in high demand, there is increasing evidence that health information technologies (HITs) will play a vital role in health care delivery

(O'Connor et al., 2016). Furthermore, the disruption caused by the COVID-19 global pandemic has resulted in a greater need for and reliance on digital health models of care for screening, treatment and ongoing maintenance of health (Wind et al., 2020). To that end, health and wellbeing apps and e-tools (e.g. websites, web-based courses) have enormous potential for empowering self-management of chronic conditions (Yardley et al., 2016). Additionally, they offer an alternative for

* Corresponding author at: Brain and Mind Centre, 94 Mallett Street, Camperdown, NSW 2050, Australia.

E-mail address: haley.lamonica@sydney.edu.au (H.M. LaMonica).

<https://doi.org/10.1016/j.invent.2021.100379>

Received 15 July 2020; Received in revised form 10 February 2021; Accepted 23 February 2021

2214-7829/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

those who prefer or are required to use information and communications technologies (e.g. during the COVID-19 pandemic), those with geographical or physical constraints (Burns et al., 2010; Rowe et al., 2020), or those who may lack awareness of available services (Burns and Rapee, 2006). Specific to mental health and wellbeing, apps and e-tools have the potential to provide low-cost intervention and prevention tools that are designed specifically for mental health disorders, such as anxiety, depression and problematic health behaviours (e.g. alcohol, gambling and smoking). The use of self-directed apps and e-tools for the purposes of symptom monitoring and management may be sufficient for individuals with lower levels of clinical need (i.e. prevention, early intervention, ongoing symptom maintenance), thus improving availability of service-based care for those who need it the most and reducing overall burden on the mental health system (Burns et al., 2014).

1.2. Health-related apps and e-tools

In 2019, there were over 204 billion apps downloaded, reflecting an increase of approximately 5% compared to 2018 (App Annie, 2020). A similar pattern of growth is evident for health-related apps, with 318,000 apps available as of March 2020 and an additional 200 being added to the market daily (IQVIA, 2017). Given the constant development of new Web-based content, the quantification of e-tools is not possible. The clinical utility of apps and e-tools has great potential. For example, evidence shows that Headspace, a mindfulness app, is associated with improvements in several aspects of psychosocial wellbeing, including irritability, affect and stress (Economides et al., 2018). The Mind Spot Clinic, has been shown to be an effective e-tool, delivering health professional and technician-assisted, Web-based cognitive behavioural therapy (iCBT) programs that have resulted in clinically significant improvements as measured on self-report symptom scales for individuals with depression (Titov et al., 2010), anxiety disorders (Newby et al., 2020) and social phobia (Titov et al., 2009). However, due to the unregulated free market that exists within the digital landscape, apps and e-tools are often of uncertain quality and efficacy (Byambasuren et al., 2018). Beyond the star ratings presented in commercial app stores, there is little information about the quality and accuracy of apps and for e-tools there are no such ratings available. Whilst star ratings may be useful as an indicator of user satisfaction and sustained engagement, they do not necessarily equate to the safety and quality of an app. This is corroborated by Singh et al. (2016), who reported that star ratings correlated poorly with clinical utility and usability of health-related apps. Similarly, it has been found that the number of app installs and active minutes of use is not associated with long term usage of more popular apps (Baumel et al., 2019). Additionally, Australians, for example, are becoming less trusting of app stores and technology companies for their recommendations and have a desire for a credible and regulated rating system for health-related apps (Consumer Health Forum of Australia, 2018). Research from other countries demonstrates that this issue is global. For example, on their respective national health system websites, the United Kingdom (NHS Innovations South East, 2014) and France (Haute Autorité de Santé, 2016) are providing more information outside of commercial app stores for individuals regarding safe health-related app choice.

1.3. Evaluating the quality and safety of apps and e-tools

As the uptake of apps and e-tools increases, both by individuals for the purposes of self-management and by health professionals as a means to complement clinical services, it is essential for all potential users to have access to measures by which they can evaluate quality and safety. Failing to evaluate the accuracy and appropriateness of health-related apps and e-tools could compromise user health and safety (Lewis and Wyatt, 2014). Several studies have highlighted inaccuracies in and a lack of evidence-base for health-related apps. For example, apps designed to help with opioid conversion calculations (Haffey et al.,

2013) or melanoma detection (Wolf et al., 2013) do not consistently follow evidence-based guidelines and may provide inaccurate information with potential hazardous repercussions (e.g. drug overdose, incorrect diagnosis). Specific to apps supporting mental health and wellbeing, a recent review found that apps for bipolar disorder were cost effective and convenient, but the majority fail to provide information on all core psychoeducation principles and do not adhere to best practice guidelines (Nicholas et al., 2015). Similarly, a review of suicide prevention apps found that many were not supported by an evidence-base and, perhaps more strikingly, identified some apps as being more harmful than helpful (Larsen et al., 2016). In relation to the latter, Larsen et al. (2016) found some apps to include potentially extremely damaging content describing or facilitating access to lethal means, providing encouragement to people to end their life and portraying suicide in a fashionable manner.

Given the potential risk associated with the use of health-related apps, there is growing interest in evaluating the quality and safety of these digital tools. The Mental Health Commission of Canada (MHCC) and the American Psychological Association (APA) have introduced health-related app assessment frameworks. The MHCC framework was uniquely developed for the Canadian context and includes criteria specific to the available evidence-base, gender responsiveness and cultural appropriateness of apps; however, does not assign ratings to these criteria (MHCC, 2018). Additionally, the MHCC is yet to develop an empirical assessment tool by which to assess the criteria. The APA's framework is a step-based model designed to inform decision making about health-related apps (APA, 2020). More specifically, the framework guides users through a hierarchical review of four key features: 1) safety and privacy; 2) evidence and benefit; 3) user engagement and; 4) interoperability. Only if a criterion is satisfied, should the user move on to the next step in the evaluation framework. Whilst a useful rubric to assist users in choosing a high quality and safe app that suits their individual needs and preferences, it does not provide an explicit rating and relies on the individual to apply the logic for each app under consideration.

In 2018, Nouri and colleagues conducted a review of criteria for assessing the quality of health-related apps. Over 23 evaluation scales were identified, of which ten were developed for general purposes with no specific subject category. The authors consolidated the evaluation criteria into seven categories: design, information/content, usability, ethical issues, security and privacy, and user-perceived value. The scales included in this review varied in what criteria they used. For example, Scott et al. (2015) focussed solely on security and safety measures, whereas others solely looked at app usability (Zapata et al., 2015; Schnall et al., 2016; Brown 3rd et al., 2013).

Included in this review is the Mobile App Rating Scale (MARS) introduced by Stoyanov et al. (2015). The MARS is a reliable, simple, multidimensional scale that requires little training to be implemented. The 23-item scale has four subscales each with multiple items: engagement (5 items), functionality (4 items), aesthetics (3 items), and information quality (7 items); and one subjective quality scale (4 items). Each feature is rated on a scale from 1 ("inadequate") to 5 ("excellent"), with more specific descriptors for the response options for each question. Upon completion, seven scores are calculated, including the mean score for each subscale, a total mean score, a mean subjective quality score, and an app-specific subscale that assesses perceived impact of an app on the user's knowledge, attitudes, intentions to change and likelihood of change specific health behaviours. The MARS has been used globally, including for the evaluation of apps to support symptom monitoring and self-care management in diverse fields of medicine, such as cardiology (Creber et al., 2016), rheumatology (Knitza et al., 2019) and obstetrics (Tency et al., 2019). It has also been adapted into different languages including German (Messner et al., 2019) and Spanish (Payo et al., 2019).

1.4. Objectives

At the present time, the MARS is one of the most widely used and internationally recognised app rating tools; however, it remains limited in its utility as it has only been designed to assess apps with modifications are needed in order to inform the quality rating of e-tools (Stoyanov et al., 2015). Furthermore, Stoyanov et al. (2015) highlight that research is needed to evaluate the safety of health-related apps specifically, both in terms of accuracy of information and privacy and security of user information (Lewis and Wyatt, 2014). Therefore, in order to evaluate not only mobile phone applications but e-tools as well, the objective of this study was to adapt the MARS to consider features and functionality that are of particular importance for the quality and safety of health-related apps and e-tools, henceforth referred to as the Adapted MARS (A-MARS), and then to test the reliability of the revised scale. Finally, this paper presents a quality assurance protocol for identifying and rating new apps and e-tools to determine appropriateness for potential use in clinical practice.

2. Material and methods

2.1. Adaptation of the MARS for use with health-related apps and e-tools

As described above, the original MARS was adapted to be appropriate for health-related apps and e-tools. As such, all questions and responses were reworded to refer to both apps and e-tools (i.e. “Do you feel engaged enough to complete the e-tool program or use the app on multiple occasions?”). The ‘Engagement’ section of the original MARS was also expanded specifically for e-tools, taking into account e-tool program completion, return use and engagement in strategies from the program. In relation to more specific changes, ‘Entertainment (Q1)’ was relabelled as ‘Engagement’ in order to better capture user engagement, including the likelihood of completing the e-tool program or using the app repeatedly, noting that health-related apps and e-tools are not necessarily designed to be fun or entertaining. The description of ‘Customisation (Q3)’ was broadened to assess whether customising the app or e-tool improves the ease of use. Modifications also included the evaluation of ‘Interoperability (Q4)’ or the ability to exchange data between other apps, e-tools, or wearables. ‘Performance (Q6)’ was expanded to specifically enquire about program errors or glitches experienced by users. Details related to the login process, the utility of the help function, and frequently asked questions were added to ‘Ease of Use (Q7)’. ‘Gestural design (Q9)’ was relabelled as ‘Design’ in order to better capture design elements of both apps and e-tools, such as popup windows and flash images, as well as to assess the consistency of the theme throughout the tool. The ‘Accuracy’ question from the original MARS which assessed the accuracy of the description in the app store was removed as it was deemed irrelevant for this tool.

A new section was also added to the A-MARS with questions of particular relevance for health-related apps and e-tools, including ‘Additional resources available (Q23)’ which evaluates whether the app or e-tool provides current and relevant resources. ‘Strategies (Q24)’ was added to determine if the app or e-tool recommends strategies linked to the target area of concern, and ‘Solutions (Q25)’ was added to assess if the app or e-tool provided one or more solutions to address the identified symptom(s). To evaluate the scope of the app or e-tool, ‘Multiple health issues/symptoms (Q26)’ was included to determine how many symptoms or health issues are addressed. The ability to use the app or e-tool in real time (i.e. real-time data tracking) was included as ‘Real time tracking (Q27)’. ‘Access to help (Q28)’ was also added to assess the ease with which help or support can be accessed via the app or e-tool. Finally, a ‘Not applicable’ option was included for all items included in the health-related subscale. No other substantive changes were made to the remainder of the questions from the original MARS.

As with the original MARS, each feature in the A-MARS is rated on a scale from 1 (“inadequate”) to 5 (“excellent”), with more specific

descriptors for the response options for each question. Upon completion, eight scores are calculated, including the mean score for each subscale (i.e. engagement, functionality, aesthetics, information, subjective quality, health-related quality), a mean quality score based on the engagement, functionality, aesthetics and information subscales, and a mean total score. The A-MARS is provided as Appendix A.

2.2. Testing of the A-MARS

Apps and e-tools to be rated were chosen based on their inclusion in the youth configuration of the InnoWell Platform (Hickie et al., 2019a). As described by LaMonica et al. (2019), the InnoWell Platform is a co-designed digital tool that is embedded within traditional in-clinic and Web-based mental health services. The InnoWell Platform was developed through Project Synergy (a \$30 M Australian-Government funded initiative delivered by InnoWell Pty Ltd.; a joint venture between the University of Sydney and PwC [Australia]; Hickie et al., 2019b; InnoWell, 2018) to collect, store, score and report clinical data back to a consumer and their health professional. Within the InnoWell Platform, a range of different care options exist, commonly known as interventions, to help the participant manage areas of health (i.e. psychological distress, sleep, physical activity). Care options are divided into two types: clinical and non-clinical. Clinical care options require a health professional’s involvement, such as individual therapy and group therapy. In contrast, a participant can immediately access and begin using non-clinical care options (see Fig. 1), such as apps and e-tools, without the support of a health professional (Davenport et al., 2019). During the co-design process, care options are tailored to the consumer population, in this case young people receiving care through primary youth mental health services.

2.3. Statistical analysis

There were 41 apps and e-tools included in the youth configuration of the InnoWell Platform at the time of writing this paper. Through participatory design workshops as well as word of mouth (Hickie et al., 2019a), these apps and e-tools are iteratively suggested by young people and their supportive others and/or recommended by health professionals for use in clinical practice. As such, the apps and e-tools included within the InnoWell Platform are continuously reviewed and updated to ensure quality and safety. All apps and e-tools were rated using the A-MARS. There were four expert raters: 1) a senior research fellow with a PhD in Clinical Psychology and three years’ experience in the design and evaluation of HITs; 2) a senior research assistant with a Master’s degrees in Exercise Physiology and Brain and Mind Sciences and two years’ experience in the design and evaluation of HITs; a 3) research affiliate with a Bachelor’s Degree in Psychology with two years’ experience working in mental health support services in a university accommodation setting, focused on engagement with culturally and linguistically diverse individuals; and 4) a research affiliate with a Bachelor’s degree in Psychology and experience in consulting for a non-profit organisation specialised in providing technological support for people with disabilities.

The first four apps and e-tools were used for training purposes. After independently rating the apps and e-tools, the raters met to compare and review the results of the pilot test and to resolve discrepancies in ratings. To reach consensus, all raters reviewed the scale in depth in order to improve the alignment of app ratings. Additionally, the meaning, purpose and descriptors of *goals, quality of information, quantity of information, credibility of source, and evidence base* were reviewed in detail to address disagreement between raters on these items. The remaining 37 apps and e-tools from the InnoWell Platform standard youth configuration were then independently rated by two raters. Using methodology previously described by Stoyanov, each rater trialled the apps and e-tools for a minimum of 10 min and then independently rated their quality using the revised A-MARS.

What you can do now:

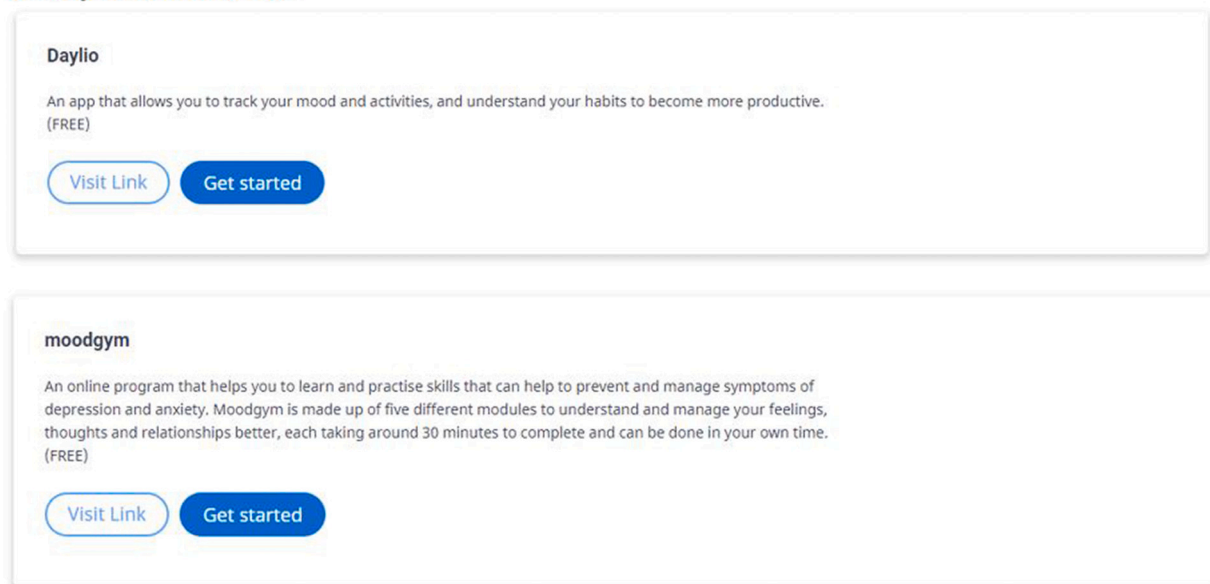


Fig. 1. Sample apps and e-tools available in the InnoWell Platform.

Cronbach's alpha was used to calculate the internal consistency of the A-MARS including the mean scores for the engagement, functionality, aesthetics and information subscales and the mean quality score reflecting the average of these subscales, the mean subjective quality and health-related quality scores, and finally the mean total score. These scores reflect the internal consistency of the scale or the degree to which the questions are measuring the same construct. Alphas were interpreted as excellent (≥ 0.90), good (0.80–0.89), acceptable (0.70–0.79), questionable (0.60–0.69), poor (0.50–0.59) and unacceptable (<0.50) (George and Mallery, 2003). Interrater reliability of the A-MARS subscales, the mean quality score and the mean total score was evaluated using the intraclass correlation coefficient (ICC). This descriptive statistic evaluates the similarity between ratings. A two-way mixed effects, average measures model with absolute agreement was used. ICCs were interpreted as excellent (≥ 0.90), good (0.76–0.89), moderate (0.51–0.75) and poor (≤ 0.50) (Portney and Watkins, 2009). All analyses were conducted using IBM SPSS (Version 26).

3. Results

3.1. App and e-tool ratings

Independent A-MARS ratings on the total score for the 37 apps and e-tools showed the scale to have an excellent level of internal consistency (Cronbach $\alpha = 0.938$) and interrater reliability (2-way mixed ICC = 0.920, 95% CI 0.797–0.987). Similarly, the independent A-MARS ratings showed the mean quality score, excluding the subjective quality and health-related quality subscales, to have an excellent level of internal consistency (Cronbach $\alpha = 0.908$) and good interrater reliability (2-way mixed ICC = 0.895, 95% CI 0.731–0.983). Internal consistencies of the A-MARS subscales were also high, ranging from acceptable to excellent (Cronbach $\alpha = 0.721$ –0.920, median = 0.824), and their interrater reliabilities were moderate to excellent (ICC = 0.687–0.910 median = 0.711), with the engagement and information subscales having the highest and lowest interrater reliability, respectively.

Examination of the corrected item-total correlations indicate that items 15 (quantity of information, $r = -0.163$), 16 (visual information, $r = 0.240$), and 17 (credibility of source, $r = 0.225$) did not correlate well with the overall information subscale; however, removal of these

items did not markedly improve the reliability of the subscale (Cronbach α if item 15 deleted = 0.737; Cronbach α if item 16 deleted = 0.715; Cronbach α if item 17 deleted = 0.728). Similarly, item 27 (real-time tracking, $r = 0.198$) was also noted to have a weak correlation with the health-related subscale; however, again, removal of the item did not notably impact the reliability (Cronbach α if item 16 deleted = 0.796).

Detailed item and subscale statistics are presented in Table 1. Additionally, a full list of the apps and e-tools rated as part of this study, including their mean scores on the A-MARS, is provided in Appendix B.

3.2. Protocol for identifying and rating new apps and e-tools

Whilst our project relied on information about apps and e-tools collected via participatory design workshops, a more real-world approach is likely to be appropriate for most health professionals and services. This approach should include: 1) a broad exploration for appropriate apps and e-tools; 2) shortlisting of apps and e-tools based off the consumer, health professional or service needs; 3) evaluation using the A-MARS; and 4) review of A-MARS scores relative to established service-specific criteria (e.g. minimum A-MARS total score, requirement of a University or Government-based source) to determine appropriateness for recommendation. This approach can be seen in more detail in Fig. 2.

4. Discussion

4.1. The need to evaluate the quality and safety of apps and e-tools

The field of HITs is evolving rapidly in order to meet the needs of consumers and health professionals as well as health services and systems of care, with the aim of driving more efficient use of resources, promoting coordination rather than fragmentation of care, and facilitating information sharing to improve shared and informed decision making. Based on a recent community consultation process conducted by the Australian Digital Health Agency, 45% of 3100 participants noted that they had difficulty accessing health care due to cost, travel distance, or the unavailability of appointments (Australian Digital Health Agency, 2018). Due to such limitations in the current health care system, consumers and health professionals alike are turning to HITs, including apps

Table 1

Interrater reliability and internal consistency of the A-MARS items and subscale scores, and corrected item-total correlations and descriptive statistics of items, based on independent ratings.

Item number	Subscale/item	Corrected item-total correlation	Mean	SD
Engagement alpha = 0.920, ICC = 0.910 (95% CI 0.859–0.948)				
1	Engagement	0.739	3.595	0.794
2	Interest	0.721	3.607	0.852
3	Customization	0.660	3.323	0.968
4	Interactivity	0.804	3.365	1.058
5	Target group	0.627	3.931	0.873
Functionality alpha = 0.785, ICC = 0.735 (95% CI 0.582–0.846)				
6	Performance	0.564	4.176	0.649
7	Ease of use	0.365	3.742	0.735
8	Navigation	0.460	3.703	0.652
9	Design	0.607	3.634	0.798
Aesthetics alpha = 0.862, ICC = 0.830 (95% CI 0.719–0.904)				
10	Layout	0.576	3.715	0.638
11	Graphics	0.707	3.715	0.920
12	Visual appeal	0.696	3.676	0.822
Information alpha = 0.721, ICC = 0.704 (95% CI 0.264–0.939)				
13	Goals	0.681	3.856	0.866
14	Quality of information	0.463	4.070	0.921
15	Quantity of information	-0.027	4.213	0.681
16	Visual information	0.240	4.000	0.576
17	Credibility of source	0.225	3.429	1.192
18	Evidence base	0.659	3.429	0.880
Subjective quality alpha = 0.889, ICC = 0.811 (95% CI 0.662–0.898)				
19	Would you recommend this app/e-tool to people who might benefit from it?	0.756	3.540	1.100
20	How many times do you think app/e-tool in the next 12-months if it was relevant to you?	0.653	3.080	1.094
21	Would you pay for this app/e-tool?	0.509	2.026	1.003
22	What is your overall star rating of the app/e-tool?	0.769	3.431	0.994
Health-related alpha = 0.786, ICC = 0.767 (95% CI 0.626–0.872)				
23	Additional resources available	0.619	3.467	0.961
24	Strategies	0.748	3.582	0.852
25	Solutions	0.475	3.532	1.039
26	Multiple health issues/symptoms addressed	0.343	3.282	1.193
27	Real-time tracking	0.198	2.782	1.253
28	Access to help	0.353	3.517	1.175

and e-tools, and for good reason. A recent meta-analysis found that apps were superior to control conditions in improving stress levels and quality of life as well as depressive and generalised anxiety symptoms, with no marked difference relative to active interventions, including in-clinic treatment (Linardon et al., 2019). Furthermore, our group has found that technology is an important tool for mental health promotion and prevention activities for young people (Burns et al., 2010). Whilst apps and e-tools hold great promise as a way of delivering self-management strategies and clinical interventions for mental ill health and maintenance of wellbeing, it is essential that consumers and health professionals have the appropriate tools by which to evaluate the quality and safety of such technologies.

4.2. The adapted Mobile App Rating Scale

This study sought to adapt the MARS to be appropriate for e-tools as well as to include items specific to health-related apps and e-tools, including details related to the availability of resources, strategies for self-management of mental ill health and wellbeing, and quality information as well as contact details to access help. Furthermore, item 4

(interactivity) was expanded to include details related to the interoperability of apps and e-tools as part of the rating. Consistent with the original validation study of the MARS, our analyses found the A-MARS total score to have excellent internal consistency and interrater reliability (Stoyanov et al., 2015). Furthermore, the median internal consistency of all subscales was good, aligning with previous studies (Domnich et al., 2016; Stoyanov et al., 2015). Our results confirm that the A-MARS is a reliable measure of apps and e-tools, including those designed specifically for health-related purposes, and is suitable for use by any relevant stakeholder, including health professionals and developers of HITs.

Interrater reliability ranged from moderate to excellent across the subscales. Whilst these levels are acceptable and consistent with those of the original validation study (Stoyanov et al., 2015), it is recommended that all raters attend a training session with an expert rater to thoroughly review the response options for each item, clarifying any ambiguities. The A-MARS should be pilot tested on three to five apps and e-tools, then reviewed until an appropriate level of interrater reliability or consensus is reached. When conducting the ratings, all apps or e-tools should be used for a suitable period so as to gain a complete understanding of the informational content, functionalities and features; a minimum of 10 min of use is recommended.

4.3. Important considerations for developers of apps and e-tools

As referenced above, as a reliable tool, the A-MARS is appropriate for use by relevant stakeholders, which is important as consumers, health professionals and services become increasingly reliant on HITs to deliver, support, or enhance care. Individual items and subscales may be particularly relevant when determining whether an app or e-tool is effective, engaging and appropriate for prospective consumer use.

4.3.1. Goal setting

Our data demonstrated that the majority (62.3%, 23/37) of health-related apps and e-tools rated as part of this study implicitly addressed goal setting by providing opportunities for the user to track outcomes through regular assessment coupled with practical strategies to improve aspects of health and wellbeing. However, five (13.5%) apps and e-tools did not provide an opportunity for the consumer to establish tangible goals, consistently indicating that this criterion was 'not applicable.' For an additional nine (24.2%) apps and e-tools, there was disagreement between raters as to whether consumers were able to set goals, indicating that the process may not be readily apparent and, in turn, unlikely to be used or effective in driving behavior change.

Whilst technology is commonly used to support health and wellbeing goals, apps in particular often fail to address key components of SMART (specific, measurable, achievable, relevant, time-based) goal setting. For example, a recent review and content analysis found that 95% (38/40) of a selection of physical activity apps had functionality for setting specific and measurable goals, but lacked the other features of SMART goals and did not allow for the re-evaluation goals, all of which are considered key to an effective goal setting strategy (Baretta et al., 2019). Importantly, the process behind planning and setting customisable goals within an app or e-tool may contribute to greater sustained engagement as well as more robust clinical outcomes. These findings align with previous research supporting the use of consumer-centred and goal-directed design approaches (Vaghefi and Tulu, 2019; Williams, 2009). Additionally, they are consistent with the outcomes from traditional in-clinic services where goal setting has been shown to be associated with higher engagement with the service whereas the absence of goals was correlated with service disengagement (Cairns et al., 2019). More robust research is now required to determine if SMART goals can be effectively set, monitored and achieved using HITs.

4.3.2. Availability of credible information from a reputable source

Health care quality improvement projects have highlighted factors

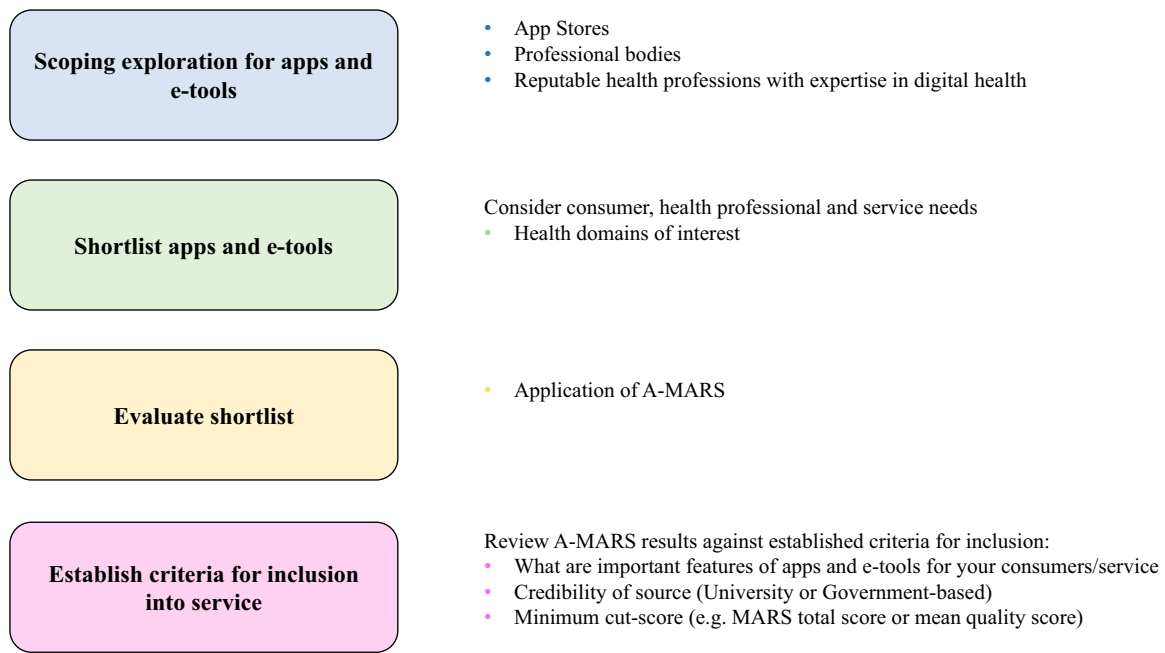


Fig. 2. Protocol for identifying and evaluating apps and e-tools.

that are most important to consumers in relation to their experience of care, including access to up to date, user friendly information and resources from reputable sources. For example, the [National College of Physicians \(UK\) \(2012\)](#) recommends that health professionals provide consumers with information in an accessible format, taking into account their preferences for the level and type of information, and advise them as to where to find additional, reliable, and credible information to support their care. Furthermore, the results of a consumer survey undertaken at four Australian hospitals indicated that the provision of high-quality information at both the point of admission and discharge was valued by consumers ([Rapport et al., 2019](#)). From our group's participatory design work, we know that the determination of the quality of available health information is directly tied to the credibility of the source, with health services and Universities being viewed as reputable ([LaMonica et al., 2021](#)). Importantly, providing details about such trustworthy sources is likely to increase uptake of and engagement with HITs and impact on health-related decision making.

Whilst quality information is an important part of a positive consumer experience, our results highlight considerable variability in the availability of quality information, ranging from 'not applicable: no available information' and '2: poor or barely relevant, appropriate, coherent, or incorrect' through to '5: highly relevant, appropriate, coherent, and correct.' Though the majority (59.4%) of apps and e-tools were rated as a 4 or 5 (mean = 4.17, standard deviation = 0.88), this remains an area for improvement for some apps and e-tools as well as an important consideration in the design and development of new HITs. Furthermore, ratings of credibility of source ranged from '1: source identified but legitimacy/trustworthiness of source is questionable' to '5: developed using nationally competitive government or research funding,' with a mean rating of 3.32 (standard deviation = 1.26). As apps and e-tools are critical mediums by which to deliver information about health risks and associated risk reduction strategies as well as interventions for symptoms of mental ill health, clear documentation of the source is an essential design element so that consumers and health professionals can accurately evaluate the trustworthiness of the information. Whilst commercially developed apps and e-tools are not inherently flawed or ineffective, it is important to recognise that the credibility of the source is important to consumers as well as health professionals. As such, commercial companies may seek to partner with academic research

teams with the aim of demonstrating a commitment to positive outcomes for consumers as well as to develop the evidence base by which to support their product.

4.3.3. Evidence base

Despite the potential of HITs to support and maintain mental health and wellbeing, the majority continue to have no scientific evidence to support their use. Based on the A-MARS, an app or e-tool is deemed to have an emerging or established evidence-base if it has been trialled and found to be effective in one or more randomised clinical trials (RCT), the gold standard for effectiveness research ([Hariton and Locascio, 2018](#)). Consistent with previous research ([Alyami et al., 2017](#); [Larsen et al., 2019](#); [Sucala et al., 2017](#); [Van Ameringen et al., 2017](#)), the majority of apps and e-tools rated in this study did not meet this standard (mean = 3.43, standard deviation = 0.87). In fact, 64.9% (24/37) have never been evaluated through a research trial. Of the remaining apps and e-tools, 21.6% (8/37) had been found to be effective in at least one RCT and 13.5% (5/37) were found to have positive or partially positive outcomes in studies of acceptability, usability and satisfaction.

The dearth of evidence of the effectiveness of apps and e-tools may relate to the iterative nature of technology development. Traditional clinical science approaches to the development and implementation of interventions rely on a linear approach, including basic science, intervention creation or adaptation, efficacy testing in both research and clinical settings, effectiveness research in community settings, and finally dissemination ([Onken et al., 2014](#)). Whilst the outcomes of each step in this process are indeed valuable, this progressive staged model can result in delays of up to 17 years for research translation into clinical practice ([Balas and Boren, 2000](#)). In light of these extended timelines as well as the standardisation requirements of RCTs, there is a high likelihood that an app or e-tool would be obsolete by the time results were published ([Kumar et al., 2013](#)). As such, developers may be inclined to move more quickly from a pilot study to dissemination ([Kumar et al., 2013](#)). Consideration of a new model for the evaluation of HITs may be necessary to streamline the identification of effective apps and e-tools. Based on a recent review, depression and anxiety apps without an evidence base were viewed to be less beneficial by consumers and had lower consumer ratings compared to evidence-based apps ([Baumel et al., 2020](#)). As such, evidence of effectiveness has the potential to

promote uptake and engagement, thus leading to enhanced outcomes.

4.3.4. Interoperability

Interoperability is the ability of HITs to exchange information with and use information from other technologies, such as apps and wearables. Interoperability is considered a fundamental requirement of HIT innovation (Lehne et al., 2019), underpinning the potential of artificial intelligence and big data analytics to improve diagnostic precision, personalised interventions and disease prevention (Insel, 2017). Furthermore, the exchange of information between electronic medical records and data from personal health apps and e-tools has the potential both to reduce documentation burden for health professionals, allowing them to spend more time focusing on care, as well as to empower and inform consumers so they can actively manage their own health and wellbeing (Lehne et al., 2019). In other words, interoperability can enhance data-driven care, including better monitoring of health and wellbeing, delivery of effective and personalised clinical care, and personalised feedback to consumers (Burns et al., 2014).

There was considerable variability in the interactivity/interoperability ratings of the apps and e-tools in this study, ranging from 1 to 5 (see Appendix A for a full description of the ratings), with a mean score of 3.36 (standard deviation = 1.06). Evidence from focus groups indicates that consumers support data sharing both between health professionals as well as with consumers to facilitate care, noting the importance of data privacy and security (Pew Charitable Trusts, 2020; LaMonica et al., 2021). Increasingly, consumers are being provided with access to their own data through personal health records; however, further consumer-driven integration between health systems and services as well as with HITs is now required to realise the full potential of interoperability, including improved transparency, efficiency and coordination of data-driven care.

4.4. Supporting health professionals to evaluate apps and e-tools for integration in practice

Despite recognising the potential benefits of using technology as part of their work, health professionals often do not have the time or appropriate resources to explore available apps and e-tools to determine their appropriateness for their consumer base (LaMonica et al., 2020). The A-MARS offers a streamlined solution to guide health professionals in the evaluation of HITs, keeping their own clinical context in mind. Whilst some health professionals may see this as an opportunity to upskill and develop digital health literacy and competence, others simply may not have the time or skill to do so. Given the complexities of evaluating such tools as well as the time required to keep up to date with what is available, including technology requirements and clinical utility of data, it may not be practical for this responsibility to fall to health professionals. Additionally, the reliability of the A-MARS as used independently by health professionals is yet to be investigated in order to determine what training requirements or materials might be required. Health services are encouraged to consider a digital navigator as an integral team member, serving to bridge the gap between HITs and in-clinic care (Wisniewski and Torous, 2020). The digital navigator would review and rate apps and e-tools ensuring only those that are safe and effective are recommended, help with technology troubleshooting for consumers and health professionals, and summarize digital data to facilitate the delivery of clinical care (Wisniewski and Torous, 2020). The digital navigator could also provide training, guidance and instruction for health professionals who are interested in developing their own skills in evaluating apps and e-tools using tools such as the A-MARS. The integration of a digital navigator within a traditional care team will serve to increase confidence and trust in the use of HITs by both health professionals and consumers, thus promoting engagement. This may, in turn, have broader implications for promoting the uptake of self-management strategies and decreasing burden on the health system in general, which is particularly relevant given the increased reliance on

HITs.

5. Conclusions

The A-MARS was shown to be a reliable scale for the purposes of evaluating the quality of health-related apps and e-tools, with moderate to excellent interrater reliability across the subscales. Specific items and subscales may be particularly important to consider when selecting, evaluating and/or recommending apps and e-tools to consumers, including: 1) the availability of explicit strategies to set, monitor and review SMART goals; 2) the accessibility of credible, user friendly information and resources from reputable sources; 3) documentation of evidence of effectiveness; and 4) interoperability of the app or e-tool with other HITs, including personal health records and electronic medical records.

Although the A-MARS is a useful tool to guide health professionals as they explore available apps and e-tools for potential clinical use, the training required to be able to use the scale effectively may be prohibitive. Additionally, health professionals may not have the time or skill set to engage in the evaluation process. The inclusion of a digital navigator as part of the care team may mitigate this barrier to identifying and using HITs in clinical practice; however, further research is required to evaluate the impact of this role on the uptake of and engagement with HITs by consumers and health professionals as well as the associated clinical outcomes. Additionally, it will be important to further evaluate how the A-MARS scores impact the selection of apps and e-tools by the digital navigator and how these scores are associated with systematically measured consumer feedback (as opposed to star ratings). It will also be important to evaluate the cost-effectiveness and return on investment of this new team member.

Finally, utilising strategies to enhance community and consumer uptake of and sustained engagement with HITs, such as apps and e-tools, is a priority in the health, medical and research sectors internationally (Australian Government Department of Health and Ageing, 2012; UK NHS, 2014). To that end, co-design methodologies, including participatory design and user testing, are widely recognised as key to ensuring the quality, usability, and acceptability of HITs. It is likely that the A-MARS can inform this co-design process, by highlighting key areas to be explored with potential end users both in the co-creation and testing phase of product development.

Funding

This research was conducted on behalf of The Australian Government Department of Health (DOH) as part of Project Synergy (2017–20). InnoWell has been formed by the University of Sydney and PwC (Australia) to deliver the \$30 M Australian Government-funded Project Synergy.

Declaration of competing interest

Professor Ian Hickie was an inaugural Commissioner on Australia's National Mental Health Commission (2012–18). He is the Co-Director, Health and Policy at the Brain and Mind Centre (BMC) University of Sydney. The BMC operates an early-intervention youth service at Camperdown under contract to headspace. He is the Chief Scientific Advisor to, and a 5% equity shareholder in, InnoWell Pty Ltd. InnoWell was formed by the University of Sydney (45% equity) and PwC (Australia; 45% equity) to deliver the \$30 M Australian Government-funded Project Synergy (2017–20; a three-year program for the transformation of mental health services) and to lead transformation of mental health services internationally through the use of innovative technologies.

Tracey Davenport is now Director (Research and Evaluation), Design and Strategy Division, Australian Digital Health Agency. The other authors have nothing to disclose.

The source of funding does not entail any potential conflict of interest for the other members of the Project Synergy Research and Development Team.

Acknowledgements

We would like to acknowledge and thank all of the participants from our co-design workshops who helped to identify apps and e-tools for use in clinical practice by and with young people.

Appendix A. Adapted Mobile App Rating Scale (A-MARS)

Adapted Mobile App Rating Scale (A-MARS)

APP / ETOOL NAME: _____

SECTION A

Engagement – fun, interesting, customisable, interactive, has prompts (e.g. sends alerts, messages, reminders, feedback, enables sharing)

1. Engagement: Is the app/e-tool engaging for the user? Do you feel engaged enough to complete the e-tool program or use the app on multiple occasions? Does it have components that make you want to use it more than similar apps/e-tools?

1 Dull, not engaging, not encouraged to start using app/e-tool

2 Mostly boring, would start program but never finish OR would download app and only use once or twice

3 OK, engaging enough to use app for a brief time (<5 min), would finish up to half an e-tool program

4 Moderately engaging, would use app/e-tool for some time; most likely complete the full program

5 Highly engaging, would stimulate repeat use of app/e-tool or would use for 5–10 min, OR would finish the full program, complete additional programs and/or return to use program again and engaged in tools/strategies given/learnt

2. Interest: Is the app/e-tool interesting to use? Does it present information in an interesting way compared to other apps/e-tools or offline/traditional tools?

1 Not interesting at all,

2 Mostly uninteresting

3 OK, neither interesting nor uninteresting; appears the same as offline/traditional tools/similar apps

4 Moderately interesting; would engage user for some time, somewhat more interesting than offline/traditional tools/similar apps

5 Very interesting, would engage user in repeat use, more engaging than traditional/offline tools/similar apps

3. Customisation: Does the app/e-tool need to be customised to make it more user friendly for you to use? Can you change settings such as sound, content, notifications, email/SMS reminders, display more to your liking?

1 App/e-tool is not user friendly; has no customisation options or requires setting to be input every time

2 App/e-tool allows little customisation; App/e-tool could be improved with more customisation options

3 Basic customisation to function adequately and/or can use app/e-tool without customization

4 Allows numerous options for customisation and/or easy to use app/e-tool; customisation somewhat unnecessary

5 Does not need any customisation for me to use the app/e-tool effectively; Allows complete tailoring the user's characteristics/preferences, remembers all settings

4. Interactivity/Interoperability: Does it allow user input, provide feedback, contain prompts (reminders, sharing options, notifications, etc.)? Does the app/e-tool adapt based on user input? Does it allow exchange of data with other apps, e-tools or wearable devices (if applicable)?

1 No interactive and features and/or no response to user input; does not adapt based off user information; has no function for exchange of data with other apps, e-tools or wearables (if applicable)

2 Some, but not enough interactive and/or interoperability features which limits app/e-tool functions; some adaptability

3 Basic interactive features to function adequately; has some capacity for exchanging data with other apps, e-tools or wearables (if applicable)

4 Offers a variety of interactive features, feedback and user input options, app/e-tool adapts somewhat to user input; with some effort can exchange data with multiple different apps, e-tools and wearables (if applicable)

5 Very high level of responsiveness through interactive features, feedback and user input options; app/e-tool adapts as user inputs; allows easy exchange of data with apps, e-tools or wearable devices (if applicable)

5. Target group: Is the content (visuals, language, design) appropriate for the target audience?

1 Completely inappropriate, unclear or confusing

2 Mostly inappropriate, unclear or confusing

3 Acceptable but not specifically designed for the target audience. May be inappropriate/ unclear/confusing at times

4 Designed for the target audience, with minor issues

5 Designed specifically for the target audience, no issues found

A. Engagement mean score = _____

SECTION B

Functionality –app/e-tool functioning, easy to learn, navigation, flow logic, and intuitive design of app/e-tool

6. Performance: How accurately/fast does the app/e-tool run (functions) and do all components with the app/e-tool (buttons/menus) work? Are there any error messages, glitches, crashes?

1 App/e-tool is broken; no/insufficient/inaccurate response (e.g. crashes/bugs/broken features, etc.) 2 Some functions work, but lagging or contains major technical problems

3 App/e-tool works overall. Some technical problems need fixing, or is slow at times

4 Mostly functional with minor/negligible problems

5 Perfect/timely response; no technical bugs found, or contains a 'loading time left' indicator (if relevant)

7. Ease of use: How easy is it to learn how to use the app/e-tool; how clear are the menu labels, icons and instructions? Is the sign-up process quick and/or simple? Are there relevant help buttons/FAQ's?

- 1 No/limited instructions; menu labels, icons are confusing; complicated; sign up process is complicated with no help buttons/FAQ's
 - 2 Takes a lot of time or effort, sign up process is somewhat complicated and/or asks for too much information and/or offers little help
 - 3 Takes some time or effort
 - 4 Easy to learn (or has clear instructions); sign up process relatively simple; some help/FAQ's
 - 5 Able to use app/e-tool immediately; intuitive; simple (no instructions needed); relevant support is obvious and helpful
- 8. Navigation:** Does moving between screens make sense; Is it easy to move from one section of the app/e-tool to another? Does the app/e-tool

provide all necessary links between screens?

- 1 No logical connection between screens at all/navigation is difficult
- 2 Understandable after a lot of time/effort
- 3 Understandable after some time/effort
- 4 Easy to understand/navigate
- 5 Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts

9. Design: Are there intuitive popup boxes, videos, animations, audio clips, flash images etc within the e-tool or are there consistent taps/swipes, pinches/scrolls within the app/e-tool? Are these relevant/accurate/make sense and in theme with the rest of the app/e-tool?

- 1 Completely confusing/inconsistent, information lacks relevance or is inaccurate/unnecessary
- 2 Often confusing/inconsistent, information of little relevance or contains some unnecessary/incorrect
- 3 Okay, some confusing and/or unnecessary information or some inconsistencies
- 4 Mostly intuitive, with negligible problems with majority of information is accurate/necessary
- 5 Perfectly consistent and intuitive, information is accurate/necessary

B. Functionality mean score = _____

SECTION C

Aesthetics – graphic design, overall visual appeal, colour scheme, and stylistic consistency

10. Layout: Is arrangement and size of buttons, icons, menus and content on the screen appropriate?

- 1 Very bad design, cluttered, some options impossible to select, locate, see or read
- 2 Bad design, random, unclear, some options difficult to select/locate/see/read
- 3 Satisfactory, few problems with selecting/locating/seeing/reading items
- 4 Mostly clear, able to select/locate/see/read items
- 5 Professional, simple, clear, orderly, logically organised

11. Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus and content?

- 1 Graphics appear amateur, very poor visual design - disproportionate, stylistically inconsistent
- 2 Low quality/low resolution graphics; low quality visual design – disproportionate
- 3 Moderate quality graphics and visual design (generally consistent in style)
- 4 High quality/resolution graphics and visual design – mostly proportionate, consistent in style
- 5 Very high quality/resolution graphics and visual design - proportionate, consistent in style throughout

12. Visual appeal: How good does the app/e-tool look?

- 1 Ugly, unpleasant to look at, poorly designed, clashing, mismatched colours
- 2 Bad – poorly designed, bad use of colour, visually boring
- 3 OK – average, neither pleasant, nor unpleasant
- 4 Pleasant – seamless graphics – consistent and professionally designed
- 5 Beautiful – very attractive, memorable, stands out; use of colour enhances app/e-tool features/menus

C. Aesthetics mean score = _____

SECTION D

Information – Contains high quality information (e.g. text, feedback, measures, references) from a credible source

13. Goals: Does app/e-tool have specific, measurable and achievable goals (are these goals specified/obvious within the app/e-tool)?

N/A Description does not list goals, or app/e-tool goals are irrelevant to research goal (e.g. using a game for educational purposes)

- 1 App/e-tool has no chance of achieving its stated goals
- 2 Description lists some goals, but app/e-tool has very little chance of achieving them
- 3 OK. App/e-tool has clear goals, which may be achievable.
- 4 App/e-tool has clearly specified goals, which are measurable and achievable
- 5 App/e-tool has specific and measurable goals, which are highly likely to be achieved

14. Quality of information: Is the content within the app/e-tool correct (including description in app store – if an app)? Is app/e-tool up to date with current research, well written, and relevant to the goal/topic of the app/e-tool?

N/A There is no information within the app/e-tool

- 1 Irrelevant/inappropriate/incoherent/incorrect
- 2 Poor. Barely relevant/appropriate/coherent/may be incorrect
- 3 Moderately relevant/appropriate/coherent/and appears correct
- 4 Relevant/appropriate/coherent/correct
- 5 Highly relevant, appropriate, coherent, and correct

15. Quantity of information: Is the information within the app/e-tool comprehensive and/or relevant but concise?

N/A There is no information within the app/e-tool

- 1 Minimal or overwhelming
- 2 Insufficient or possibly overwhelming
- 3 OK but not comprehensive or concise
- 4 Offers a broad range of information, has some gaps or unnecessary detail; or has no links to more information and resources

5 Comprehensive and concise; contains links to more information and resources

16. Visual information: Is visual explanation of concepts – through charts/graphs/images/videos, etc. – clear, logical, correct?

N/A There is no visual information within the app/e-tool (e.g. it only contains audio, or text)

1 Completely unclear/confusing/wrong or necessary but missing

2 Mostly unclear/confusing/wrong

3 OK but often unclear/confusing/wrong

4 Mostly clear/logical/correct with negligible issues

5 Perfectly clear/logical/correct

17. Credibility of source: does the information within the app/e-tool seem to come from a credible source?

1 Source identified but legitimacy/trustworthiness of source is questionable (e.g. commercial business with vested interest)

2 Appears to come from a legitimate source, but it cannot be verified (e.g. has no webpage)

3 Developed by small NGO/institution (hospital/centre, etc.) /specialised commercial business, funding body

4 Developed by government, university or as above but larger in scale

5 Developed using nationally competitive government or research funding (e.g. Australian Research Council, NHMRC)

18. Evidence base: Has the app/e-tool been trialled/tested; must be verified by evidence (in published scientific literature)?

N/A It has not been trialled/tested

1 The evidence suggests the app/e-tool does not work

2 App/e-tool has been trialled (e.g., acceptability, usability, satisfaction ratings) and has partially positive outcomes in studies that are not randomised controlled trials (RCTs), or there is little or no contradictory evidence.

3 App/e-tool has been trialled (e.g., acceptability, usability, satisfaction ratings) and has positive outcomes in studies that are not RCTs, and there is no contradictory evidence.

4 App/e-tool has been trialled and outcome tested in 1-2 RCTs indicating positive results

5 App/e-tool has been trialled and outcome tested in > 3 high quality RCTs indicating positive results

D. Information mean score = _____*

* Exclude questions rated as “N/A” from the mean score calculation

SECTION E

App/e-tool subjective quality rating

19. Would you recommend this app/e-tool to people who might benefit from it?

1 Not at all - I would not recommend this app/e-tool to anyone

2 There are very few people I would recommend this app/e-tool to

3 Maybe - There are several people whom I would recommend it to

4 There are many people I would recommend this app/e-tool to

5 Definitely - I would recommend this app/e-tool to everyone

20. How many times do you think you would use this app/e-tool in the next 12 months if it was relevant to you?

1 None

2 1-2

3 3-10

4 10-50

5 >50

21. Would you pay for this app/e-tool?

1 No

3 Maybe

5 Yes

22. What is your overall star rating of the app/e-tool?

1 ★ One of the worst apps/e-tools I've used

2 ★★

3 ★★★Average

4 ★★★★

5 ★★★★★ One of the best apps/e-tools I've used

E. Subjective mean score = _____

SECTION F

Supplement Health Related App/e-tools: Questions to consider when using a health-related app/e-tool

23. Additional resources available. Does the app/e-tool provide up to date relevant offline/online resources to support the information presented?

1 No – provides no further resources

2 Provides few online OR offline resources

3 Somewhat – provides some offline and/or online resources, may be outdated

4 Provides adequate online and/or offline resources

5 Yes – provides abundant and up to date offline and online resources

24. Strategies: Does the app/e-tool recommend strategies that are non-tech based and linked to the problems you have reported?

1 No – none

2

3 Somewhat – some information; may be too much, or too little resources

- 4
 5 Yes – adequate/plentiful but not overbearing
25. Solutions: Does it offer multiple solutions for one issue?
 1 No – offers one solution to address one issue/health symptom
 2
 3 Some – offers some solutions for the one issue/health symptom; offers solutions but they indirectly address the issue
 4
 5 Yes – offers various related solutions to directly address the issue
26. Multiple health issues/symptoms: Does it address more than one symptom or health issue?
 1 No – addresses one symptom/health issue only
 2
 3 Some – addresses some symptoms/health issues; or considers many but only partly address them
 4
 5 Yes – considers multiple symptoms/health issues and related ones, and sufficiently addresses them
27. Real time tracking: Can you use the app/e-tool in real time, as you’re experiencing a health issue?
 1 No – the app/e-tool is mainly useful for prevention or recovery
 2
 3 The app/e-tool is useful for prevention, management and/or recovery of the health issue(s)
 4
 5 Yes – the app/e-tool is useful for prevention, management and recovery of the health issue(s)
28. Access to help: Easy/obvious to access health related help when needed?
 1 No – Difficult to navigate or find related health information when needed
 2 Can find needed information after a lot of time/effort
 3 Can find needed information after some time/effort
 4 Easy to understand/navigate needed information
 5 Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts to needed health information. Offline options are available.

F. Health-related information mean score = _____

Scoring

App/e-tool quality scores:

SECTION

A: Engagement Mean Score = _____

B: Functionality Mean Score = _____

C: Aesthetics Mean Score = _____

D: Information Mean Score = _____

Quality mean Score = _____

E. Subjective quality Score = _____

F. Health-related quality Score = _____

Appendix B. Mean objective and subjective app and e-tool quality scores

App	Engagement	Functionality	Aesthetics	Information ^a	Quality ^a	Subjective quality	Health-related quality ^a	Total mean ^a
Consensus rated								
Daylio	3.20	4.00	4.00	2.67	3.47	3.00	1.33	3.04
Rise Up Recover	3.20	3.25	3.00	2.75	3.05	3.00	3.17	3.07
eCouch	3.80	3.00	3.00	4.20	3.50	1.75	4.17	3.63
Eclipse	4.00	4.00	4.00	4.50	4.13	3.25	4.17	4.13
Independent ratings for reliability testing								
Beacon2.0	3.10	3.00	2.67	3.80	3.14	3.00	2.92	3.08
Beyond blue forums	3.50	4.00	3.83	3.38	3.68	3.38	3.58	3.61
Beyond now safety plan	3.60	4.25	3.83	3.80	3.87	3.00	3.33	3.64
Black dog bite back	4.20	4.50	4.67	4.42	4.45	3.75	4.58	4.35
Black dog my compass	3.70	3.75	3.67	3.75	3.72	3.75	4.00	3.77
Black dog snapshot	2.20	3.00	2.00	2.40	2.40	1.25	2.42	2.21
BRAVE-Online	4.30	4.38	4.33	4.42	4.36	4.00	4.58	4.33
Butterfly Foundation	3.80	3.88	4.00	3.50	3.79	3.38	3.50	3.68
Calm Harm	4.00	4.13	4.33	3.40	3.96	3.13	3.58	3.76
CCI ^b building body acceptance	1.70	3.13	2.83	3.60	2.81	3.13	2.92	2.88
CCI ^b disordered eating	1.70	3.50	2.83	3.70	2.93	3.13	2.92	2.96
Christopher bot	2.50	3.13	2.50	1.50	2.41	1.38	1.00	2.00
Counselling ONLINE	4.00	4.00	4.33	4.10	4.11	3.25	3.75	3.91
Daisy	2.80	4.13	3.67	3.83	3.61	2.25	2.92	3.27
Daybreak	3.90	3.88	3.67	3.42	3.71	2.50	3.33	3.45
HabitBull	4.40	4.25	4.33	3.90	4.22	3.63	N/A	4.10
LoveSmart	3.40	3.50	3.33	2.60	3.21	2.13	N/A	2.99
MindMax	3.10	3.38	3.67	3.42	3.39	2.63	3.00	3.20
MindShift	3.90	4.25	4.67	2.50	3.83	4.50	2.67	3.75
Moodgym	4.00	3.50	3.50	3.50	3.63	3.63	4.08	3.70

(continued on next page)

(continued)

App	Engagement	Functionality	Aesthetics	Information ^a	Quality ^a	Subjective quality	Health-related quality ^a	Total mean ^a
MoodMission	4.30	4.25	4.00	3.83	4.10	3.88	4.25	4.08
MyFitnessPal	4.60	3.75	4.17	3.42	3.98	3.75	3.50	3.86
My QuitBuddy	3.60	3.38	3.17	3.30	3.36	2.50	3.08	3.17
My Study Life	3.80	3.88	3.17	1.50	3.09	2.13	1.00	2.58
Nike Training Club	4.20	4.38	4.33	3.50	4.10	4.00	N/A	4.08
Positive Choices	3.60	3.00	3.50	3.92	3.50	3.50	3.50	3.50
PTSD Coach Australia	3.80	3.75	3.17	3.80	3.63	2.63	3.83	3.50
ReachOut Breathe	2.80	4.50	4.17	2.00	3.37	2.50	2.50	3.08
ReachOut WorryTime	2.80	3.75	3.50	1.60	2.91	2.00	3.00	2.78
Recharge	3.90	3.50	4.00	3.40	3.70	2.00	3.08	3.31
Recovery Record	3.60	3.50	3.50	3.10	3.43	2.50	3.33	3.26
Schizophrenia Health Storylines	3.10	3.50	3.17	2.10	2.97	1.75	2.25	2.64
Smiling Mind	3.60	4.13	4.00	3.50	3.81	4.13	3.25	3.77
Smoke Free	4.00	4.13	4.33	3.75	4.05	3.00	3.00	3.70
SuperBetter	4.20	4.13	4.00	4.17	4.12	3.50	3.83	3.97
THIS WAY UP	4.10	4.00	4.00	4.25	4.09	3.63	3.83	3.97
Zombies, Run!	4.10	4.13	4.17	2.50	3.72	3.63	N/A	3.71

^a Mean scores calculated based on number of variables rated, excluding not applicable items (e.g. goals, evidence-base).^b CCI: Centre for Clinical Interventions.

References

- Alyami, M., Giri, B., Alyami, H., Sundram, F., 2017. Social anxiety apps: a systematic review and assessment of app descriptors across mobile store platforms. *Evidence-Based Mental Health* 20 (3), 65–70. <https://doi.org/10.1136/eb-2017-102664>.
- American Psychiatric Association, 2020. Available at: <https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model>. (Accessed 15 July 2020).
- App Annie, 2020. The State of Mobile in 2020: The Key Stats You Need to Know. Available at: <https://www.appannie.com/en/>. (Accessed 15 July 2020).
- Australian Digital Health Agency, 2018. Australia's National Digital Health Strategy. Available at: https://conversation.digitalhealth.gov.au/sites/default/files/adha-strategy-doc-2ndaug_0_1.pdf. (Accessed 15 July 2020).
- Australian Government Department of Health and Ageing, 2012. E-Mental Health Strategy for Australia. DoHA, Canberra.
- Balas, E., Boren, S., 2000. Managing clinical knowledge for healthcare improvements. In: van Bemmel, J.H., McCray, A.T. (Eds.), *Yearbook of Medical Informatics*. Schattauer, Stuttgart, pp. 65–70.
- Baretta, D., Bondaronek, P., Direito, A., Steca, P., 2019. Implementation of the goal-setting components in popular physical activity apps: review and content analysis. *Digit. Health* 5. <https://doi.org/10.1177/2055207619862706>, 2055207619862706.
- Baumel, A., Muench, F., Edan, S., Kane, J.M., 2019. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J. Med. Internet Res.* 21 (9), e14567 <https://doi.org/10.2196/14567>.
- Baumel, A., Torous, J., Edan, S., Kane, J.M., 2020. There is a non-evidence-based app for that: a systematic review and mixed methods analysis of depression-and anxiety-related apps that incorporate unrecognized techniques. *J. Affect. Disord.* 273, 410–421. <https://doi.org/10.1016/j.jad.2020.05.011>.
- Brown 3rd, W., Yen, P.-Y., Rojas, M., Schnall, R., 2013. Assessment of the health IT usability evaluation model (health-ITUEM) for evaluating mobile health (mHealth) technology. *J. Biomed. Inform.* 46 (6), 1080–1087. <https://doi.org/10.1016/j.jbi.2013.08.001>.
- Burns, J.R., Rapee, R.M., 2006. Adolescent mental health literacy: young people's knowledge of depression and help seeking. *J. Adolesc.* 29 (2), 225–239. <https://doi.org/10.1016/j.adolescence.2005.05.004>.
- Burns, J.M., Davenport, T.A., Durkin, L.A., Luscombe, G.M., Hickie, I.B., 2010. The internet as a setting for mental health service utilisation by young people. *MJA* 192 (11), S22–S26. <https://doi.org/10.5694/j.1326-5377.2010.tb03688.x>.
- Burns, J., Hickie, I., Christensen, H., 2014. Strategies for Adopting and Strengthening e-Mental Health. A review of the evidence, Mental Health Commission of New South Wales. Available at: <https://nswmentalhealthcommission.com.au/sites/default/files/assets/File/Report%20-%20The%20Sax%20Institute%20E-Mental%20Health%20Evidence%20Review%20cover%20page.pdf>. (Accessed 15 July 2020).
- Byambasuren, O., Sanders, S., Beller, E., Glasziou, P., 2018. Prescribable mHealth apps identified from an overview of systematic reviews. *NPJ Digital Medicine* 1 (1), 1–12. <https://doi.org/10.1038/s41746-018-0021-9>.
- Cairns, A.J., Kavanagh, D.J., Dark, F., McPhail, S.M., 2019. Goal setting improves retention in youth mental health: a cross-sectional analysis. *Child Adolesc. Psychiatry Ment. Health* 13 (1), 31. <https://doi.org/10.1186/s13034-019-0288-x>.
- Consumers Health Forum of Australia, 2018. Results of Australia's Health Panel Survey on Recommendations and Regulations of Smartphone Apps for Health and Wellness. Author, Canberra, Australia.
- Creber, R.M.M., Maurer, M.S., Reading, M., Hiraldo, G., Hickey, K.T., Iribarren, S., 2016. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the Mobile Application Rating Scale (MARS). *JMIR Mhealth and Uhealth* 4 (2), e74 <https://doi.org/10.2196/mhealth.5882>.
- Davenport, T.A., LaMonica, H.M., Whittle, L., English, A., Iorfino, F., Cross, S., Hickie, I.B., 2019. Validation of the InnoWell Platform: protocol for a Clinical Trial. *JMIR Res. Protoc.* 8 (5), e13955 <https://doi.org/10.2196/13955>.
- Domnich, A., Arata, L., Amicizia, D., Signori, A., Patrick, B., Stoyanov, S., Hides, L., Gasparini, R., Panatto, D., 2016. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. *BMC Med. Inform. Decis. Mak.* 16 (1), 83. <https://doi.org/10.1186/s12911-016-0323-2>.
- Economides, M., Martman, J., Bell, M.J., Sanderson, B., 2018. Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone app: a randomized controlled trial. *Mindfulness* 9 (5), 1584–1593. <https://doi.org/10.1007/s12671-018-0905-4>.
- George, D., Mallery, P., 2003. SPSS for Windows step by step: A simple guide and reference. 11.0 update, 4th ed. Allyn & Bacon, Boston.
- Haffey, F., Brady, R.R.W., Maxwell, S., 2013. A comparison of the reliability of smartphone apps for opioid conversion. *Drug Saf.* 36 (2), 111–117. <https://doi.org/10.1007/s40264-013-0015-0>.
- Hariton, E., Locascio, J.J., 2018. Randomised controlled trials—the gold standard for effectiveness research. *BJOG* 125 (13), 1716. <https://doi.org/10.1111/1471-0528.15199>.
- Haute Autorité de Santé, 2016. Good Practice Guidelines on Health Apps and Smart Devices (Mobile Health or mHealth). Available at: https://www.has-sante.fr/upload/docs/application/pdf/2017-03/dir1/good_practice_guidelines_on_health_apps_and_smart_devices_mobile_health_or_mhealth.pdf. (Accessed 15 July 2020).
- Hickie, I.B., Scott, E.M., Cross, S.P., Iorfino, F., Davenport, T.A., Guastella, A.J., Naismith, S.L., Carpenter, J.S., Rohleder, C., Crouse, J.J., Hermens, D.F., 2019a. Right care, first time: a highly personalised and measurement-based care model to manage youth mental health. *Med. J. Aust.* 211 (9 Suppl).
- Hickie, I.B., Davenport, T.A., Burns, J.M., 2019b. Project synergy: co-designing technology-enabled solutions for Australian mental health services reform. *Med. J. Aust.* 211 (7 Suppl).
- InnoWell, 2018. What is project synergy? Available at: <https://synergy.innowell.org/>. (Accessed 15 July 2020).
- Insel, T.R., 2017. Digital phenotyping: technology for a new science of behaviour. *JAMA* 318 (13), 1215–1216. <https://doi.org/10.1001/jama.2017.11295>.
- IQVIA, 2017. The Growing Value of Digital Health: Evidence and Impact on Human Health and the Healthcare System Institute Report. IQVIA. Available at: <https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health>. (Accessed 16 July 2020).
- Knitza, J., Tascilar, K., Messner, E.-M., Meyer, M., Vossen, D., Pulla, A., Bosch, P., Kittler, J., Kleyer, A., Sewerin, P., Mucke, J., Hasse, I., Simon, D., Krusche, M., 2019. German mobile apps in rheumatology: review and analysis using the mobile application rating scale (MARS). *JMIR Mhealth and Uhealth* 7 (8), e14991 <https://doi.org/10.2196/14991>.
- Kumar, S., Nilsen, W.J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., Riley, W.T., Shar, A., Spring, B., Spruijt-Metz, D., Hedeker, D., Honavar, V., Kravitz, R., Lefebvre, R.C., Mohr, D.C., Murphy, S.A., Quinn, C., Shusterman, V., Swendeman, D., 2013. Mobile health technology evaluation: the mHealth evidence workshop. *Am. J. Prev. Med.* 45 (2), 228–236. <https://doi.org/10.1016/j.amepre.2013.03.017>.
- LaMonica, H.M., Davenport, T.A., Braunstein, K., Ottavio, A., Piper, S., Martin, C., Hickie, I.B., Cross, S.P., 2019. Technology-enabled person-centred mental health service reform: strategy for implementation science. *JMIR Ment Health* 6 (9), e14719 <https://doi.org/10.2196/14719>.
- LaMonica, H.M., Milton, A., Braunstein, K., Rowe, S.C., Ottavio, A., Jackson, T., Easton, M.A., Hambleton, A., Hickie, I.B., Davenport, T.A., 2020. Technology-enabled solutions for Australian mental health services reform: impact evaluation. *JMIR Formative Research* 4 (11), e18759.
- LaMonica, H.M., Davenport, T.A., Roberts, A., Hickie, I.B., 2021. Understanding technology preferences and requirements for health information technologies

- designed to improve mental health and maintain wellbeing for older persons: a participatory design study. *JMIR Aging* 4 (1), e21461. <https://doi.org/10.2196/21461>.
- Larsen, M.E., Nicholas, J., Christensen, H., 2016. A systematic assessment of smartphone tools for suicide prevention. *PLoS One* 11 (4), e0152285. <https://doi.org/10.1371/journal.pone.0152285>.
- Larsen, M.E., Huckvale, K., Nicholas, J., Torous, J., Birrell, L., Li, E., Reda, B., 2019. Using science to sell apps: evaluation of mental health app store quality claims. *NPJ Digital Medicine* 2 (1), 1–6. <https://doi.org/10.1038/s41746-019-0093-1>.
- Lehne, M., Sass, J., Essenwanger, A., Schepers, J., Thun, S., 2019. Why digital medicine depends on interoperability. *NPJ Digit Med.* 2, 79. <https://doi.org/10.1038/s41746-019-0158-1>.
- Lewis, T.L., Wyatt, J.C., 2014. mHealth and mobile medical apps: a framework to assess risk and promote safer use. *J. Med. Internet Res.* 16 (9), e210 <https://doi.org/10.2196/jmir.3133>.
- Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., Fuller-Tyszkiewicz, M., 2019. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry* 18 (3), 325–336. <https://doi.org/10.1002/wps.20673>.
- Mental Health Commission of Canada, 2018. Mental Health Apps: How to Make an Informed Choice. Available at: https://www.mentalhealthcommission.ca/sites/default/files/2018-01/eMH_app_eng.pdf. (Accessed 15 July 2020).
- Messner, E.-M., Terhorst, Y., Barke, A., Baumeister, H., Stoyanov, S., Hides, L., Kavanagh, D.J., Pryss, R., Sander, L., Probst, T., 2019. Development and validation of the German version of the Mobile Application Rating Scale (MARS-G). *JMIR Mhealth and Uhealth.* 14479 <https://doi.org/10.2196/14479>.
- National College of Physicians (UK). Patient Experience in Adult NHS Services: Improving the Experience of Care for People Using Adult NHS Services: Patient Experience in Generic Terms. 2012. London: Royal College of Physicians.
- Newby, J.M., Haskelberg, H., Hobbs, M.J., Mahoney, A.E., Mason, E., Andrews, G., 2020. The effectiveness of internet-delivered cognitive behavioural therapy for health anxiety in routine care. *J. Affect. Disord.* 264, 535–542. <https://doi.org/10.1016/j.jad.2019.11.087>.
- NHS Innovations South East, 2014. App Development: An NHS Guide for Developing Mobile Healthcare Applications. Available at: https://citcentoolkit.files.wordpress.com/2018/04/98533_nhs_inn_appdevroad.pdf. (Accessed 15 July 2020).
- Nicholas, J., Larsen, M.E., Proudfoot, J., Christensen, H., 2015. Mobile apps for bipolar disorder: a systematic review of features and content quality. *J. Med. Internet Res.* 17 (8), e198 <https://doi.org/10.2196/jmir.4581>.
- Nouri, R., Kalhori, S.R.N., Ghazisaeedi, M., Marchand, G., Yasini, M., 2018. Criteria for assessing the quality of mHealth apps: a systematic review. *J. Am. Med. Inform. Assn.* 25 (8), 1089–1098. <https://doi.org/10.1093/jamia/ocy050>.
- O'Connor, S., Hanlon, P., O'Donnell, C.A., Garcia, S., Glanville, J., Mair, F.S., 2016. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. *BMC Med. Inform. Decis. Mak.* 16, 120. <https://doi.org/10.1186/s12911-016-0359-3>.
- Onken, L.S., Carroll, K.M., Shoham, V., Cuthbert, B.N., Riddle, M., 2014. Reenvisioning clinical science: unifying the discipline to improve the public health. *Clin Psychol Sci.* 2 (1), 22–34. <https://doi.org/10.1177/2167702613497932>.
- Payo, R.M., Álvarez, M.F., Díaz, M.B., Izquierdo, M.C., Stoyanov, S.R., Suárez, E.L., 2019. Spanish adaptation and validation of the Mobile Application Rating Scale questionnaire. *Int. J. Med. Inform.* 129, 95–99. <https://doi.org/10.1016/j.ijmedinf.2019.06.005>.
- Pew Charitable Trusts, 2020 March. Patients seek better exchange of health data among their care providers. https://www.pewtrusts.org/-/media/assets/2020/03/health_data_patientbrief.pdf accessed 30/06/2020.
- Portney, L.G., Watkins, M.P., 2009. *Foundations of Clinical Research: Applications to Practice*. Pearson/Prentice Hall, Upper Saddle River, NJ.
- Rapport, F., Hibbert, P., Baysari, M., Long, J.C., Seah, R., Zheng, W.Y., Jones, C., Preece, K., Braithwaite, J., 2019. What do patients really want? An in-depth examination of patient experience in four Australian hospitals. *BMC Health Serv. Res.* 19 (38) <https://doi.org/10.1186/s12913-019-3881-z>.
- Rowe, S.C., Davenport, T.A., Easton, M.A., Jackson, T.A., Melsness, J., Ottavio, A., Sinclair, J., Hickie, I.B., 2020. Co-designing the InnoWell platform to deliver the right mental health care first time to regional youth. *Aust. J. Rural Health* 28 (2), 190–194. <https://doi.org/10.1111/ajr.12617>.
- Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Dieguez, A., Carry, M., Gelaude, D., Mosley, J.P., Travers, J., 2016. A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J. Biomed. Inform.* 60, 243–251. <https://doi.org/10.1016/j.jbi.2016.02.002>.
- Scott, K., Richards, D., Adhikari, R., 2015. A review and comparative analysis of security risks and safety measures of mobile health apps. *Australasian Journal of Information Systems* 19. <https://doi.org/10.3127/ajis.v19i0.1210>.
- Singh, K., Drouin, K., Newmark, L.P., Lee, J., Faxvaag, A., Rozenblum, R., Pabo, E.A., Landman, A., Klinger, E., Bates, D.W., 2016. Many mobile health apps target high-need, high-cost populations, but gaps remain. *Health Aff.* 35 (12), 2310–2318. <https://doi.org/10.1377/hlthaff.2016.0578>.
- Stoyanov, S.R., Hides, L., Kavanagh, D.J., Zelenko, O., Tjondronegoro, D., Mani, M., 2015. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR MHealth and Uhealth.* 3 (1), e27 <https://doi.org/10.2196/mhealth.3422>.
- Sucala, M., Cuijpers, P., Muench, F., Cardoso, R., Soflau, R., Dobrea, A., Achimas-Cadariu, P., David, D., 2017. Anxiety: there is an app for that. A systematic review of anxiety apps. *Depression and anxiety* 34 (6), 518–525. <https://doi.org/10.1002/da.22654>.
- Tency, I., Van Hecke, A., Coorevits, P., 2019. Assessing the Quality of Pregnancy Apps through Development and Validation of the Dutch Version of the Mobile Application Rating Scale (MARS). Presentation at European Congress of Intrapartum Care. Location: Turijn.
- Titov, N., Andrews, G., Choi, I., Schwencke, G., Johnston, L., 2009. Randomized controlled trial of web-based treatment of social phobia without clinician guidance. *Aust. N.Z. J. Psychiat.* 43 (10), 913–919. <https://doi.org/10.1080/00048670903179160>.
- Titov, N., Andrews, G., Davies, M., McIntyre, K., Robinson, E., Solley, K., 2010. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 5 (6), e10939. <https://doi.org/10.1371/journal.pone.0010939>.
- United Kingdom National Health Service, October 2014. Five Year Forward View. NHS England, Leeds.
- Vaghefi, I., Tulu, B., 2019. The continued use of mobile health apps: insights from a longitudinal study. *JMIR Mhealth and Uhealth.* 7 (8), e12983 <https://doi.org/10.2196/12983>.
- Van Ameringen, M., Turna, J., Khalesi, Z., Pullia, K., Patterson, B., 2017. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. *Depression and anxiety* 34 (6), 526–539. <https://doi.org/10.1002/da.22657>.
- Williams, A., 2009. User-centered design, activity-centered design, and goal-directed design: a review of three methods for designing web applications. In: *Proceedings of the 27th ACM International Conference on Design of Communication*, pp. 1–8.
- Wind, T.R., Rijkeboer, M., Andersson, G., Riper, H., 2020. The COVID-19 pandemic: the 'black swan' for mental health care and a turning point for e-health. *Internet Interv.* 20, 100317 <https://doi.org/10.1016/j.invent.2020.100317>.
- Wisniewski, H., Torous, J., 2020. Digital navigators to implement smartphone and digital tools in care. *Acta Psychiatr. Scand.* 141 (4), 350–355. <https://doi.org/10.1111/acps.13149>.
- Wolf, J.A., Moreau, J.F., Akilov, O., Patton, T., English, J.C., Ho, J., Ferris, L.K., 2013. Diagnostic inaccuracy of smartphone applications for melanoma detection. *JAMA Dermatol.* 149 (4), 422–426. <https://doi.org/10.1001/jamadermatol.2013.2382>.
- Yardley, L., Choudhury, T., Patrick, K., Michie, S., 2016. Current issues and future directions for research into digital behavior change interventions. *Am. J. Prev. Med.* 51 (5), 814–815. <https://doi.org/10.1016/j.amepre.2016.07.019>.
- Zapata, B.C., Fernández-Alemán, J.L., Idri, A., Toval, A., 2015. Empirical studies on usability of mHealth apps: a systematic literature review. *J. Med. Syst* 39 (2), 1. <https://doi.org/10.1007/s10916-014-0182-2>.