Lessons Learned in Selection and Review of Depression Apps for Primary Care Settings

Margaret R. Emerson[1] · Shinobu Watanabe-Galloway[2] · Danae Dinkel[3] · Suhasini Kotcherlakota[1] · Louis Fok[2]

Abstract

Depression is the leading cause of disability worldwide and is one of the most common mental health issues being addressed within primary care settings. Mobile apps, which can be used to help people manage their depressive symptoms, are rapidly developing. However, many challenges exist for clinicians and providers to simply select an appropriate app for use within target populations. The objectives of this article are as follows: (1) to describe the search processes that were used to identify depression-related mobile apps and (2) to describe the review process that was implemented to inform and evaluate the identified depression-related mobile health apps for use with our target population. A research team consisting of information technology researchers, primary and psychiatric care providers, and health care researchers completed two mobile app searches to identify depression-related apps which could be used for further exploration within an underserved integrated primary care setting. Sixteen mobile apps were narrowed down to 4 mobile apps, through a series of steps involving screening, collaboration of the interprofessional team, information technology expertise input, and mobile app evaluation tools. This article described the steps a research team used to search, screen, and assess mental health mobile apps for integrated primary care patients with depression. This step-by-step guide focused on depression-related apps; however, similar steps and principles identified in this guide can be applied to other health apps.

Keywords Depression · Mobile applications · Mental health · Information technology · Primary health care

Introduction

Depression is the leading cause of disability worldwide and is one of the most common mental health issues being addressed within primary care settings (The World Health Organization, 2017). Mobile apps, also known as "apps," which can be used to help people manage their depressive symptoms, are rapidly developing. In 2017, approximately 325,000 mental health–related mobile apps were available in the digital health care market (Research2Guidance, 2017; Schueller, et al., 2018). Despite the proliferation of apps, currently, there are no requirements for developers to demonstrate or publish data on the effectiveness and efficacy of apps before they put products on the market. Thus far, only a limited number of health-related mobile apps have been approved by the U.S. Food and Drug Administration (FDA) (U.S. Food and Drug Administration, 2019). Moreover, the applicability and effectiveness of many of these apps have not been tested in real-world settings resulting in the availability of minimally tested or unevaluated apps on the market (Carlo, et al., 2019).

It is imperative that providers and clinicians become aware of evidence-based apps to deliver effective patient education and health care services (Clark, 2018; Kayyali, et al., 2017). However, many challenges exist for clinicians and providers to simply select an appropriate app. First, developers are continually producing new apps and updating versions of existing apps, which makes it almost impossible to keep track of all of the mental health apps available at any given time. Second, information about the data confidentiality and effectiveness of the intervention delivered through the app is not always readily available. The Apple and Google Play market post-user reviews or average ratings but, in many cases, it is difficult to find objective assessments of the quality of the apps. Third, while studies may report the quality and usefulness of mental health apps, in some instances by the time the study results are

✉ Margaret R. Emerson
    Margaret.emerson@unmc.edu

1   University of Nebraska Medical Center College of Nursing, Omaha, NE, USA

2   University of Nebraska Medical Center College of Public Health, Omaha, NE, USA

3   School of Health & Kinesiology, University of Nebraska Omaha, Omaha, NE, USA

published, some of the apps that were tested are no longer available on the market.

Nevertheless, based on the existing research data, it is clear that mobile apps have many benefits for individuals with depression and other mental health problems (Vantola, 2014; Chandrashekar, 2018) and there is an urgent need for making up-to-date app evaluation information available for researchers, clinicians, and patients wanting to use mobile app technology in given patient populations. Various app screening models and evaluation tools have been published by researchers and/or professional organizations (Chan et al. 2015; Ferguson and Jackson 2017; Neary & Schueller, 2018; Nouri et al., 2018). For example, organizations, such as the American Psychiatric Association recommends that to evaluate an app, researchers should have enough app background information to thoroughly assess and explore features pertaining to, the privacy and security, potential benefits, evidence of use, engagement, and interoperability of an app (Torous et al., 2018; Henson et al., 2019). The Mobile App Rating Scale is a multidimensional mobile app quality–rating tool which provides an app total quality score in addition to four subscale scores for engagement, functionality, aesthetics, and information quality (Stoyanov et al., 2015). Despite the availability of these tools and models, the practice of evaluating mobile apps prior to use is not mandated in clinical practice and the uptake of these resources into clinical practice appears limited in the integrated care setting. While there have been studies outlining processes for selecting mobile apps in a clinical setting (Boudreaux et al., 2014; Chan, et al., 2015; Ferguson & Jackson, 2017; Neary & Schueller, 2018) and reviews of mental health apps (Marshall, Dunstan, & Bartik, 2020; O'Loughlin, Neary, Adkins, & Schueller, 2019; Powell et al., 2016; Shen et al., 2015), limited research has applied these processes and considered app selection specifically within an integrated primary care setting for use with a low-income, high disparate population who experience higher rates of mental health issues in comparison to the general population. The clinical utility of mobile apps within these target populations needs to be explored further, to enable the utilization of quality mobile apps to be seamlessly incorporated into these practice settings. This was a dilemma our research team faced when we were developing a pilot study to identify depression mobile apps for underserved patients being managed in an integrated primary care setting.

Objective

Our research team consisting of information technology researchers, primary and psychiatric care providers, and health care researchers developed a process for screening and evaluating mental health mobile apps for use in the underserved and health disparaging target population based on existing recommendations (Boudreaux et al., 2014; Chan, et al., 2015; Ferguson & Jackson, 2017; Neary & Schueller, 2018; Nouri et al., 2018). Given the limited availability of disseminated standardized screening and evaluation processes for our intended target population which includes low-income patients with high rates of depression, the ever-increasing availability of health-related apps, and the inability of the typical academic research timeline for communicating results, this information can provide much-needed guidance for clinicians and researchers working within an integrated primary care setting. The overall purpose of the study was to document the application of two existing methods for screening and evaluating apps for depression to generate a list of recommended apps for the integrated primary care clinic setting, which are suitable for the target population. Specifically, the objectives of this paper are as follows: (1) to describe the search processes that were used to identify depression-related mobile apps and (2) to describe the review process that was implemented to inform and evaluate the identified depression-related mobile health apps for use with our target population. While the focus of this study was on depression-related apps, similar steps and principles can be applied to other health apps.

Methods

Mobile App Search Processes

First Search

The main objective of the research project was to explore the use of depression-related mobile apps in underserved populations being cared for in an integrated primary care setting. A literature search was completed to identify relevant studies that examined the effectiveness and usefulness of mental health apps for further exploration within our intended study population. Several sources were identified; however, there was limited information available which specifically evaluated the apps with consistent methods and/or which explored the app use in populations which were congruent with our population focus. Therefore, the decision was made to use eight depression and cognitive behavioral therapy apps identified from the most current available systematic review available at the time which explored apps in alignment with the goals of our study (Huguet et al., 2016).

Second Search

During the 4 months, between our first search and the initiation of the study, critical events were encountered prompting a second search to be conducted. These events included (1) individual app-accessibility issues, (2) dissemination of guidance for incorporating apps into mental health treatment from

professional organizations, and (3) identification of previous studies (Lee & Kim, 2018, 2019) indicating that app searches should mimic what patients and clinicians may do to identify potential apps for their use. For example, evidence suggests that most users do not look beyond the first ten apps identified or even download past the top 5 (O'Loughlin et al., 2019). Furthermore, apps are commonly identified by searching the Internet, Google, and Apple app stores, in addition to identifying apps within other apps themselves (Tiongson, 2015). Therefore, the decision was made to identify mobile apps utilizing a similar approach, along with incorporating the guidance from a professional organization which resulted in a list of 16 applications for the review process (American Psychiatric Association: App Evaluation Model, n.d.; Ferguson & Jackson, 2017).

Screening and Assessment of Apps

Development of Screening Check List

An information technology (IT) specialist reviewed the literature and identified several studies providing additional methods for evaluating and selecting apps. Additionally, the IT specialist contacted the FDA for guidance. The FDA responded via email to indicate that they evaluate the claims and associated performance data only for mobile apps that are actively regulated medical devices (S. Kotcherlakota, personal communication, August 29, 2019). In summary, unless the mobile apps are part of a regulated medical device per Section 201 (h) of the Food, Drug, and Cosmetic Act, general wellness use only apps are considered to be low risk to the public. Therefore, the FDA referred the IT specialist to general guidance documents (U.S. Food and Drug Administration, 2017, 2019) for the evaluation of mobile apps. Therefore, a table of criteria was created by the IT specialist to tabulate information available for each health app to allow for key comparisons to be made among the apps.

Screening of Initially Selected 16 Apps

The IT specialist used the screening checklist to document details of the 16 apps selected. Based on the screening results and other processes (Boudreaux et al., 2014; Ferguson & Jackson; 2017), the minimum criteria deemed to be acceptable to the IT specialist, and all three research team members were as follows:

1. Minimum number of reviews (min of 1000 on Google or Android App Store reviews)
2. App rating minimum

   a. Minimum of 4 out of 5 stars on the Android App Store
   b. Minimum of 4 out of 5 stars on Google App Store

3. Updated within the last 6 months
4. Evidence of experts involved in the development of the app
5. Evidence of any cited research in the development or use of the app

The above criteria were applied to the 16 apps. Three of the apps were eliminated due to the limited number of reviews, 2 apps were eliminated because no updates had been completed within the last 6 months, 2 were eliminated because there was no evidence of research and/or expert involvement, and 2 were eliminated due to star ratings being below 4. This resulted in 7 apps which can be found in Table 1. These apps were ordered by (1) apps that specifically targeted depression and used cognitive behavioral strategies in their app description and (2) provided evidence of research and/or expert involvement either within the app or on their website with higher priority given to those with research since our app was to be used as part of a research study.

In-depth Assessment of 7 Apps Using MARS and APA Forms

The next step was to consider the quality of the mobile health apps. Two tools were selected: the Mobile App Rating Scale (MARS) and the APA App Evaluation Form (App Evaluation Form, n.d.; Stoyanov et al., 2015). Using two tools—one which has been established in the literature to indicate app quality and another recommended by the leading organization in the USA—to inform our exploration of the apps was deemed important since there is no current standardization of app evaluation within the integrated primary care setting.

The MARS is a 23-item scale used for trialing, classifying, and rating the quality of mobile health apps (Stoyanov et al., 2015). The scale assesses engagement, functionality, aesthetics, information, and quality. The MARS has shown excellent internal consistency (alpha = 0.90) and interrater reliability (ICC = 0.79). The research team deemed the MARS tool to be an excellent indicator of app quality from the view of both the clinician and the patient, as the developed categories were based on an extensive search of publications and resources encompassing a variety of factors indicating app quality including engagement (Stoyanov et al., 2015). The MARS has been used to evaluate apps in studies seeking to increase self-management. Promoting self-management is often a feature within integrated care practices suggesting that although the MARS had not been used in our target population to our awareness, it did have application in studies with similar focus and had been suggested in other mobile app search processes (Masterson Creber et al., 2016; Neary & Schueller, 2018). Furthermore, since the development of the MARs, the User Version of the Mobile App Rating Scale (uMARS) tool has been created (Stoyanov et al., 2016), which enables end-users to assess the quality of mHealth apps as

Table 1 App characteristics from IT specialist

App name	Moodpath	Sanvello	Youper	Wysa	Woebot	Calm	Headspace
Platform	Both	Both	Both	Both	Both	Both	Both
App rating (average)	4.6	4.8	5	4.6	4.7	4.7	4.5
User reviews (average)	11.7k	11.5k	36k	18k	1.1k	293k	102k
Features	Assess (daily questions, bi-weekly health assessment, share with mental health professionals). Track and reflect (mood journal, quick ref. of emotional states, reflection, receive supportive insights), Choose from a large library of audio and written exercises, based on the methods of cognitive behavioral therapy (CBT) to understand the psychology behind your moods, improve your sleep, reduce brooding, learn effective mindfulness techniques, acquire skills in self-compassion.	An all-new "Feeling Better"-guided journey with new original videos, audio exercises, and activities. New engaging educational content based on the principles of cognitive behavioral therapy (CBT) New home base with dynamic content, allowing users to select from a number of immersive experiences Daily mood tracking as first view each day. Friendly weekly assessments to measure progress, next step reminders, and recommendations based on progress Refreshed user interface with new color palette, fonts, icons, and background videos	Quick conversations that can change your day Beautiful mood tracker Emotional and behavioral insights Personality test Symptom monitoring Integration with Google Fit makes self-help and self-care easy	Vent and talk through things or just reflect on your day Practice CBT (cognitive behavioral therapy) and DBT techniques to build resilience in a fun way Use one of 40 conversational coaching tools which helps in dealing with stress, anxiety, depression, panic attacks, worry, loss, or conflict	Think through situations with step-by-step guidance from chatbot using tools from cognitive behavioral therapy (CBT) Learn about yourself with intelligent mood tracking Master skills to reduce stress and live happier through over 100+ evidence–based stories from our clinical team	Calming anxiety, managing stress, deep sleep, focus and concentration, relationships, breaking habits, happiness, gratitude, self-esteem, body scan, loving-kindness, forgiveness, non-judgment, Commuting to work or school. mindfulness at college mindfulness at work, walking meditation, calm kids 100+ sleep stories: adult bedtime stories guaranteed to lull you to sleep 7- and 21-day programs for both beginner and advanced users Life-changing Calm Masterclasses taught by world-renowned experts Breathing exercises to help you relax Exclusive music engineered to help you focus, relax or sleep Unguided timed meditation Open-ended meditation 30+ soothing nature sounds and scenes to use during meditation, yoga or to help you sleep Track your progress with the following: daily streaks, time spent meditating	Courses, stress & anxiety meditation courses include the follwoing: managing anxiety, letting go of stress, restlessness, navigating change Falling asleep & waking up meditation courses include the following: sleep, restlessness Personal growth meditation courses include the following: relationships, self-esteem, kindness, happiness, acceptance, appreciation, balance, patience, generosity Work & productivity meditation courses include the following: finding focus, prioritization, productivity, creativity Life challenges meditation courses include: transforming anger, handling sadness, dealing with regret, self-esteem, grieving Movement and sport meditation courses

Table 1 (continued)

App name	Moodpath	Sanvello	Youper	Wysa	Woebot	Calm	Headspace
							include: motivation, focus, training, competition, communication, analysis, recovery, rehab, concentration Physical health meditation courses include: mindful eating, coping with cravings, pain management, pregnancy, coping with cancer Student meditation courses include: dealing with distractions, leaving home, sleep casts, sleep music, wind down exercises, an app for kids, eyes open exercises
Experts involved?	Yes	Yes	Yes	Yes	Yes	No	Yes
Seller	Address available	Address available	Address available	Address available	Address available	Address available	Address available
Age rating	Everyone	Teen	Everyone	Everyone	Everyone	Everyone	Everyone
Price	Free	Free	Free	Free	Free	Free	Free
In-app purchases	$6.49–$84.99 per item	$3.99–$199	$0.99–$400.00 per item	$1.49–$154.26 per item		$0.99–$399.99 per item	$6.49–$119.99 per item
Last updated	11 Jun 2019	26 Jun 2019	25 Jun 2019	13 Jun 2019	20 Jun 2019	20 Jun 2019	24 Jun 2019
Size	26 M	41 M	62 M	6.5 M	25 M	Varies with device	32 M
Installs	500 K+	1 M+	500 K+	500 K+	100 K+	10 M+	10 M+
Permissions							
Access location: GPS and network-based	x	x	x				
Photos/media/files: read, modify, delete USB storage	x	x	x		x		
Receive data from the Internet	x	x	x			x	x

Table 1 (continued)

App name	Moodpath	Sanvello	Youper	Wysa	Woebot	Calm	Headspace
View network connections	x	x	x	x	x		
Prevent device from sleeping	x	x	x	x	x	x	x
Control vibration	x	x			x	x	x
Run at startup	x	x	x			x	
Access contacts		x			x	x	x
Camera: take pictures and videos		x					x
Phone: read phone status and identity		x					
Wi-Fi connection information, view Wi-Fi connections		x	x	x			x
Microphone-record audio		x	x			x	x
Device ID & call information		x	x	x	x		
Read phone status and identity		x	x	x	x		x
Use accounts on the device		x					
Change audio settings		x	x				x
Full network access	x	x	x				
Pair with Bluetooth devices		x		x	x	x	x
Wearable sensors/activity data body sensors (like heart rate monitors)			x		x	x	
Receive data from the Internet			x	x	x		
Read Google service configuration			x	x		x	
Draw over other apps					x		
Google play license						x	
Calendar, add or modify calendar events and send							x

Table 1 (continued)

App name	Moodpath	Sanvello	Youper	Wysa	Woebot	Calm	Headspace
	email to guests without owners' knowledge read calendar events plus confidential information						
Research	Available	Available	Available	Available	Available	Available	Available
Main areas of app promotion	Stress depression and anxiety	Stress/anxiety depression help	Emotional health assistant	AI coach CBT DBT chatbot-better mindset and mental health	CBT, DBT & mindfulness skills self-care expert	Meditation and sleep stories	Meditation and sleep

well. This adaptation to the MARS was an important reason for selecting this tool as our team felt it was essential to incorporate a quality tool which patients in the target population could use in the future to evaluate intended apps. Two different team members using different mobile operating systems (iOS vs Android) independently applied the MARS to the 7 apps. Due to a required cost to access a majority or all features of 2 apps, the MARS was only fully applied to 5 apps. The two MARS scores were calculated and averaged for each app and the results can be found in Table 2.

The APA App Evaluation Form was a second means used to explore each of the remaining apps. The App Evaluation Tool/Form is derived from the APA Evaluation Model which is arranged strategically to prioritize the divisions of the tool as follows: (1) Safety/Privacy, (2) Evidence (i.e., effectiveness), (3) Ease of Use, and (4) Interoperability (App Evaluation Model, n.d.). The 4 categories of the App Evaluation Tool/Form are further explored using a series of questions that pertain to that particular category. Questions can be answered as Yes, No, or Unsure. At the end of each category, there is a final question, which asks the evaluator to rank the overall concern level with the following options: Major Concerns, Some Concerns, and No Concerns. The APA (APA App Evaluation Form, nd) places a higher emphasis on the first two categories in terms of the selection process. The evaluation tool/form was completed online and submitted directly to the APA. Two research team members collaborated to complete the APA tool/form for all of the apps which did not have an associated cost. Once the forms were completed, a member of the digital APA team was contacted to determine how the results and information our team entered would be communicated. Correspondence from the digital team member indicated a mobile app APA team was being developed with the goal of evaluating submitted apps. Since that time, there has been a call for interested parties to submit applications to serve on the APA App Advisor Expert Panel (App Advisor Expert Panel, n.d.); it is presumed results will be communicated with the research team once the processes of this panel have been established. Nonetheless, while the results have not been communicated, the act of completing the forms allowed the team to further critically evaluate app selection. The results of the APA form tool completion can be found in Table 3.

Findings and Results

Finalized App List Considering Target Population

Our finalized list of apps needed to take into consideration the population in which the apps were intending to be used. After discovering through the application of the MARS and APA form that two of the apps had limited features available

Table 2   MARS results

App name	Engagement	Functionality	Aesthetics	Information	MARS mean	Subjective	Perceived impact
Reviewer 1							
Sanvello	3.40	4.50	3.67	3.57	3.78	2.50	3.00
Woebot	4.40	4.50	4.00	3.57	4.12	3.00	3.67
Wysa	1.80	4.00	3.33	2.60	2.93	1.50	2.00
Moodpath	2.20	4.50	3.33	3.20	3.31	2.75	2.83
Youper	2.40	2.00	2.67	2.40	2.37	1.50	2.50
Reviewer 2							
Sanvello	4.00	4.75	5.00	4.14	4.47	4.25	4.00
Woebot	4.00	4.75	4.00	3.86	4.15	3.25	3.50
Wysa	3.60	4.25	4.00	4.00	3.96	3.50	4.00
Moodpath	4.00	5.00	4.33	4.00	4.33	4.00	4.00
App 7	3.80	3.75	4.33	3.57	3.86	3.25	4.00
Mean scores							
Sanvello	3.70	4.63	4.33	3.86	4.13	3.38	3.50
Woebot	4.20	4.63	4.00	3.71	4.13	3.13	3.58
Wysa	2.70	4.13	3.67	3.30	3.45	2.50	3.00
Moodpath	3.10	4.75	3.83	3.60	3.82	3.38	3.42
Youper	3.00	2.88	3.50	2.99	3.12	2.38	3.25

without cost, the team consulted clinical partners. In consultation with our clinical partners who work directly with patients, it was determined that any app that asks for payment would discourage long-term use by the patient population. Thus, the decision was made to exclude any apps which appeared to have costs associated with some of the desired features and/or if the app required a fee to download the content to complete the evaluation tools. By initiating the process of completing the evaluation tools, the research team was able to see how cost played into the features of each app. In some instances, the app would no longer be free after the trial period was over. In other instances, desirable features of the app were not accessible unless payment/subscription was received. The MARS and APA forms were not completed in their entirety for any apps that had an associated cost that could impede app utilization; this resulted in only 5 of the 7 apps being fully evaluated.

In summary, we used the following steps for searching and reviewing apps:

1. Identified mobile apps based on patients typical methods for identifying mobile apps
2. Used these strategies along with lessons learned from literature to identify a preliminary list of mobile apps
3. Directed a team of interprofessionals to search for apps using patient strategies and expertise to compile a list of no more than 10 mobile apps per professional
4. Compiled all lists from each team members and arranged in order of commonality

5. Sought guidance from IT specialist to suggest additional factors from the literature to consider
6. Established minimum criteria for the apps for further evaluation based on information compiled by IT specialist
7. Eliminated any apps which did not meet minimum or inclusion criteria
8. Evaluated remaining apps using the MARS and APA mobile evaluation tools
9. Prioritized and finalized apps based on mean scores of the MARS, features illuminated in completing the APA Evaluation Form, and target population considerations including the prohibition of cost

Based on this information, Table 4 outlines recommended steps for individuals/teams taking a similar approach to mobile app selection for use within an integrated primary care clinic working with an underserved population.

Conclusions

This article described the steps our research team used to search, screen, and assess mental health mobile apps for underserved patients with depression in an integrated primary care setting. Because the reasons for identifying apps may vary widely depending on the research questions or clinical settings, we are not advocating for using the exact steps outlined above. However, researchers or clinicians should consider the following points before they initiate the app

Table 3   APA app evaluation forms

Yes-3, No-1, Unsure-2

App Name	Youper	Youper	Moodpath	Moodpath	Wysa	Wysa	Sanvello	Sanvello	Woebot	Woebot
App Version, if known	6.11.0 00	7.00.0 01	3.1.2	3.1.4	4.6.5	4.6.1	8.1.0	8.0.2	3.7	3.6.1
Release Date	16-Jul-19	8-Aug-19	Jul-19	Aug-20	8-Aug-19	2017, Unable to determine exact date from release history	1-Aug-19	14-Nov-16	23-Aug-19	2017, was released for apple store 1/25/2018
PRIVACY AND SAFETY										
Is there a Privacy Policy?	3	3	3	3	3	3	3	3	3	3
Is Personal Data collected?	3	3	3	3	1	3	3	3	3	3
Is Personal Data de-identified?	3	3	2	2	3	3	2	3	2	3
Can you opt-out of data collection?	3	3	1	1	2	2	1	3	3	3
Can you delete data?	3	3	3	3	2	3	2	3	2	3
Are Cookies Placed?	3	3	3	1	2	3	3	3	3	3
Is data shared with outside parties or groups?	3	3	3	3	1	3	3	3	1	2
Is Data encrypted on the device?	2	1	2	3	3	3	2	3	1	3
Is data encrypted on the server (if stored there)?	3	3	3	3	2	3	2	3	3	3
Does it claim it meets HIPAA?	1	2	1	2	3	2	3	3	1	1
Average of each reviewer	2.7	2.7	2.4	2.4	2.2	2.8	2.4	3	2.2	2.7
Average of both reviewers	2.7		2.4		2.5		2.7		2.5	
OVERALL PRIVACY/SAFETY EVALUATION	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	No Concerns	Some Concerns	No Concerns
EVIDENCE										
Does the app appears to do what it claims?	3	3	3	3	3	3	3	3	3	3
Is there peer reviewed, published evidence about the app, or the science behind it?	2	3	3	3	3	3	3	3	3	3
Is there any feedback from users t support claims (app store, website, review sites, etc.)?	2	3	2	3	3	3	3	3	3	3
Does the content appear of at least reasonable value?	3	3	3	3	3	3	3	3	3	3
Average of each reviewer	2.5	3	2.75	3	3	3	3	3	3	3
Average of both reviewers	2.8		2.9		3.0		3.0		3.0	
OVERALL EVIDENCE EVALUATION	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	No Concerns	Some Concerns	No Concerns
USABILITY										
Is the application easy to access for the patient?	3	3	3	3	3	3	3	3	3	3
Will it be easy to use on a long term basis?	3	3	3	3	3	2	3	1	3	2

Table 3   (continued)

	Youper	Youper	Moodpath	Moodpath	Wysa	Wysa	Sanvello	Sanvello	Woebot	Woebot
Is it customizable?	1	1	2	1	2	3	3	3	2	2
Does it need the internet to work?	3	3	3	3	3	2	3	3	3	2
Is it accessible for those with impaired vision or other disabilities?	2	1	2	1	2	2	2	3	2	2
Average of each reviewer	2.4	2.2	2.6	2.2	2.6	2.4	2.8	2.6	2.6	2.2
Average of both reviewers	2.3		2.4		2.5		2.7		2.4	
OVERALL USABILITY EVALUATION	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns
DATA SHARING										
Can it share data with an EMR?	2	2	2	2	2	2	2	1	2	2
Can you print out our data?	1	1	1	1	1	2	1	3	1	2
Can you export/download your data?	1	1	1	1	1	2	1	3	3	2
Can you share data with other user tools? (AppleHealth Kit, FitBit)	3	3	1	1	2	3	3	3	2	2
Average of each reviewer	1.8	1.8	1.3	1.3	1.5	2.3	1.8	2.5	2.0	2.0
Average of both reviewers	1.8		1.3		1.9		2.1		2.0	
OVERALL DATA SHARING EVALUATION	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	Some Concerns	No Concerns	Some Concerns	Some Concerns

selection process. Firstly, it is important to specify the patient population (e.g., age, socioeconomic status), clinical setting (e.g., integrated primary care), and the health issue(s) (e.g., depression). Conduct a brief review of existing mobile app search processes to identify which ones fit within your setting and purpose (Boudreaux et al., 2014; Chan, et al., 2015; Ferguson & Jackson, 2017; Neary & Schueller, 2018). Additionally, if using the MARS, the developers recommend the following prior to use of this tool: (1) the raters undergo a training exercise before commencing use, (2) the raters have a common understanding of the population in which the app is intended for use, (3) the raters clarify any items in the tool that are unclear, and (4) the appropriate fit is determined for the MARS within the specified health concern (Stoyanov et al., 2015).

Secondly, it is important to identify initial criteria to screen a large number of apps that are available for consumers. Several criteria were used to narrow the top apps identified by the research team. Consideration was given to the number of reviews and overall star rating, as Martin et al. (2017) identified a positive correlation to the number of installs and these factors. Yet, as information continues to explore these areas, research suggests these factors play a minimal role in the usability and clinical application for specific apps (Singh et al., 2016) suggesting that using these criteria to narrow down our initial list could eliminate suitable apps. However, it is

Table 4 Recommended steps for research studies utilizing mental health mobile apps

Steps	
1. Team formation	Form a group of 3–5 members including: • Mental health care provider • IT expert • Mental health researcher
2. App identification	• Using relevant keywords (e.g., depression, anxiety) identify potential apps through an Internet search and in Apple and Android marketplaces • Conduct a literature review of peer-reviewed articles published in the last 2 years to identify potential apps • Team members conduct searches individually and compile a list of all potential apps
3. App documentation	Based on information obtained from the developer website and marketplaces, document information about apps (including but not limited to): • Developer • Year developed, a most recent update • Intended uses, target audience • Cost • Privacy, security • Research evidence • Marketplace rating, number of downloads/users At least two members do documentation of any given app and the team consolidates information to create a master table
4. App evaluation	• At least two members evaluate any given app using MARS, APA assessment form, or any other form(s) • Take an average score of two reviewers to create a final evaluation sheet
5. Final selection	Based on information gathered in Step 3 (app documentation) and Step 4 (app evaluation), select apps that meet the needs and expectations of the target audience/users

important to consider patients' perspectives and likelihood of continued use as the APA research team noted, app ratings, and reviews are highly influential when patients are searching for apps to use themselves (Torous et al., 2019). Therefore, taking reviewer ratings and the number of users or downloads as an elimination tool could be helpful for reducing the number of potential apps. Additionally, if working with a low-income underserved population, cost should also be considered early on in the process (i.e., select free apps only).

Thirdly, we recommend using a reviewer team comprised of individuals from appropriate backgrounds. Our team included a practicing clinician who specializes in the psychiatric care of integrated primary care patients, a researcher with experience of qualitative assessment of mobile technology, a researcher specializes in psychiatric epidemiology, and an IT specialist. This approach provided an opportunity to vet a wide array of apps from a series of different perspectives and vantage points. Ultimately, many of the apps that were identified by one team member were also identified by other members of the team suggesting that these apps would be suitable for a variety of settings and implementation efforts. Additionally, our results reflect details for apps given what was identifiable or able to be located at the time by a member of the research team. One could speculate that an individual's technology knowledge base likely plays a role in how apps are explored; thus, we felt it was important to include an IT specialist to assist in leveraging expertise and accommodate for various knowledge differences.

Our team would also recommend choosing evaluation tools that allow for a thorough assessment of potential apps. Our research team chose to use two different evaluation tools to provide more than one lens with which to evaluate the mobile apps, with a prominent difference between the two tools being a static score. Teams need to consider whether it is preferable to have a score versus a structured inquiry, which does not result in a score. In some ways, the APA App Evaluation Form is reflective of the ever-changing environment that exists within the use of technology, so perhaps there are advantages in not having a static score. Yet, not having a score makes app-to-app comparisons a little more challenging.

Importantly, although our team has outlined a process to be followed, the critical steps to consider early on when seeking to use mobile apps within the clinical practice are organizational compliance and privacy-related factors which can impede the integration of mobile app data into patient care. Organizational use of technology adoption, specifically that of mobile technology, into clinical practice, is not universal. Subsequently, there can be privacy and confidentiality concerns that limit a clinician or researchers' ability to actually use the mobile app in clinical practice. Therefore, we recommend that intended providers, researchers, and clinicians work with their organizations to ensure that the use of the app meets those compliance and privacy standards very early on in the process.

In conclusion, our results capture apps that were explored in one moment of time and, in no way, are an endorsement of

52

J. technol. behav. sci. (2021) 6:42–53

any one mobile application. In looking towards the future, it is essential to consider how provider perspectives can be incorporated into the app selection search process highlighting their perspectives regarding the appropriateness of these apps for an intended population.

Compliance with Ethical Standards

On behalf of all the authors, the corresponding author states that the work completed within this submission was congruent with our affiliated compliance and ethical standards. We obtained IRB approval for the purposes of our funded grant proposal; however, the findings provided in this manuscript do not contain the human subject data collected with informed consent; therefore, it is our assumption that the IRB and consent forms are not applicable to the content submitted.

Conflict of Interest   The authors declare that they have no conflict of interest.

References

App Advisor Expert Panel. (n.d.) Retrieved August 1, 2019 from https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-advisor-expert-panel

App Evaluation Model. (n.d.). Retrieved February 14, 2020 from https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model

App Mobile Evaluation Form. (n.d.). Retrieved July 20, 2019 from https://app.smartsheet.com/b/form/0f6b39c460ca421081a249ed4d6f03d0

Boudreaux, E. D., Waring, M. E., Hayes, R. B., Sadasivam, R. S., Mullen, S. M., & Pagoto, S. P. (2014). Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Translational Behavioral Medicine, 4*(4), 363–371 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4286553/. Accessed 15 Oct 2019.

Carlo, A. D., Ghomi, R. H., Renn, B. N., & Areán, P. A. (2019). By the numbers: ratings and utilization of behavioral health mobile applications. *NPJ Digital Medicine, 2*(54) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6572775/, 54. Accessed 20 May 2020.

Chan, S., Torous, J., Hinton, L., & Yellowlees, P. (2015). Towards a framework for evaluating mobile mental health apps. Telemedicine and e-Health, 21(12), 1038–1041.

Chandrashekar, P. (2018). Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. *mHealth, 4*(6). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5897664/. Accessed 5 May 2019.

Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: the IAPT program. *Annual Reviews, 14*, 159–183 https://www.annualreviews.org/doi/full/10.1146/annurev-clinpsy-050817-084833.

Ferguson, C., & Jackson, D. (2017). Selecting, appraising, recommending and using mobile applications(apps) in nursing. *Journal of Clinical Nursing.* https://doi.org/10.1111/jocn.13834.

Henson, P., David, G., Albright, K., & Torous, J. (2019). Deriving a practical framework for the evaluation of health apps. *The Lancet Digital Health, 1*(2), e52–e54.

Huguet, A., Rao, S., McGrath, P., Wozney, L., Wheaton, M., Conrod, J., & Rozario, S. (2016). A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLoS One, 11*(5), e0154248. https://doi.org/10.1371/journal.pone.0154248.

Kayyali, R., Peletidi, A., Ismail, M., Hashim, Z., Bandeira, P., & Bonnah, J. (2017). Awareness and use of mHealth apps: a study from England. *Pharmacy (Basel), 5*(2), 33–46. https://doi.org/10.3390/pharmacy5020033.

Lee, J., & Kim, J. (2018). Method of app selection for healthcare providers based on consumer needs. *CIN: Computers, Informatics, Nursing, 36*(1), 45–54. https://doi.org/10.1097/CIN.0000000000000399.

Lee, J., & Kim, J. (2019). Can menstrual health apps selected based on users' needs change health-related factors? A double-blind randomized controlled trial. *Journal of the American Medical Informatics Association, 26*(7), 655–666. https://doi.org/10.1093/jamia/ocz019.

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Apps with maps—anxiety and depression Mobile apps with evidence-based frameworks: systematic search of major app stores. *JMIR Mental Health, 7*(6), e16525.

Martin, W., Sarro, F., Jia, Y., Zhang, Y., & Harman, M. (2017). A survey of app store analysis for software engineering. *IEEE Transactions on Software Engineering, 43*(9), 817–847. https://doi.org/10.1109/TSE.2016.2630689.

Masterson Creber, R. 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://mhealth.jmir.org/2016/2/e74.

Neary, M., & Schueller, S. M. (2018). State of the field of mental health apps. *Cognitive and Behavioral Practice, 25*(4), 531–537.

Nouri, R., R Niakan Kalhori, S., Ghazisaeedi, M., Marchand, G., & Yasini, M. (2018). Criteria for assessing the quality of mHealth apps: a systematic review. *Journal of the American Medical Informatics Association, 25*(8), 1089–1098.

O'Loughlin, K., Neary, M., Adkins, E., & Schueller, S. (2019). Reviewing the data security and privacy policies of mobile apps for depression. *Internet Interventions, 15*, 110–115. https://doi.org/10.1016/j.invent.2018.12.001.

Powell, A. C., Torous, J., Chan, S., Raynor, G. S., Shwarts, E., Shanahan, M., & Landman, A. B. (2016). Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. *JMIR mHealth and uHealth, 4*(1), e15.

Research2Guidance (2017). *MHealth economics 2017 – current status and future trends in mobile Health* (pp. 2–26). Retrieved from https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/

Schueller, S., Neary, M., O'Loughlin, K., & Adkins, E. (2018). Discovery of and interest in health apps among those with mental health needs: survey and focus group study. *Journal of Medical Internet Research, 20*(6), e10141. https://doi.org/10.2196/10141.

Shen, N., Levitan, M. J., Johnson, A., Bender, J. L., Hamilton-Page, M., Jadad, A. A. R., & Wiljer, D. (2015). Finding a depression app: a review and content analysis of the depression app marketplace. *JMIR mHealth and uHealth, 3*(1), e16.

Singh, K., Drouin, K., Newmark, L., Lee, J., Faxvaag, A., Rozenblum, R., … Bates, D. (2016). Many mobile health apps target high-need, high-cost populations, but gaps remain. Health Affairs, 35(12), 2310–2318. https://doi.org/10.1377/hlthaff.2016.0578.

Stoyanov, S., Hlides, L., Kavanaugh, D., 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.

Stoyanov, S. R., Hides, L., Kavanagh, D. J., & Wilson, H. (2016). Development and validation of the user version of the Mobile

Application Rating Scale (uMARS). *MIR Mhealth Uhealth, 4*(2), e72. https://mhealth.jmir.org/2016/2/e72. https://doi.org/10.2196/mhealth.5849.

Tiongson, J. (2015). *Mobile app marketing insights: how consumers really find and use your apps*. Retrieved from https://www.thinkwithgoogle.com/consumer-insights/mobile-app-marketing-insights/

Torous, J. B., Chan, S. R., Gipson, S. Y. M. T., Kim, J. W., Nguyen, T. Q., Luo, J. et al. (2018). A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. Psychiatric Services, 69(5), 498–500.

Torous, J., Kumaravel, A., Myrick, K. J., Hoffman, L., & Druss, B. G. (2019). *Hands-on with smartphone apps for serious mental illness: an interactive tutorial for selecting, downloading, discussing, and engaging in apps*. San Francisco: Symposium conducted at the American Psychiatric Association Annual Meeting.

U.S. Food and Drug Administration. (2017) Office of Medical Products and Tobacco, Center for Devices and Radiological Health. Software as a medical device (SAMD): clinical evaluation - guidance for industry and Food and Drug Administration Staff (Docket Number: FDA-2016-D-2483). Retrieved from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/software-medical-device-samd-clinical-evaluation-guidance-industry-and-food-and-drug-administration. Accessed 29 Aug 2019.

U.S. Food and Drug Administration. (2019). *Device software functions including mobile medical applications*. Retrieved from https://www.fda.gov/medical-devices/digital-health/device-software-functions-including-mobile-medical-applications. Accessed 29 Aug 2019.

Vantola, C. L. (2014). Mobile devices and apps for health care professionals: uses and benefits. *Pharmacy and Therapeutics, 39*(5), 356–364 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4029126/. Accessed 5 May 2019.

World Health Organization. (2017). "Depression: let's talk" says WHO, as depression tops list of causes of ill health. World Health Organization. https://www.who.int/news-room/detail/30-03-2017%2D%2Ddepression-let-s-talk-says-who-as-depression-tops-list-of-causes-of-ill-health. Accessed 5 May 2019.