



Original Paper

A review of mobile chatbot apps for anxiety and depression and their self-care features

Arfan Ahmed^a, Nashva Ali^a, Sarah Aziz^a, Alaa A Abd-alrazaq^a, Asmaa Hassan^a, Mohamed Khalifa^b, Bushra Elhusein^c, Maram Ahmed^c, Mohamed Ali Siddig Ahmed^c, Mowafa Househ^{a,*}

^a Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar

^b Centre for Health Informatics, Australian Institute of Health Innovation, Faculty of Medicine, Health and Human Sciences, Macquarie University, Sydney, Australia

^c Mental Health Services, Hamad Medical Corporation, Qatar

ARTICLE INFO

Keywords:

"Anxiety"
"Depression"
"Chatbots"
"Pandemics"
"COVID-19"

ABSTRACT

Background: Anxiety and depression rates are at an all-time high. Smartphone-based mental health chatbots can aid psychiatrists replacing some of the costly human based interaction providing a unique opportunity to expand the availability and quality of mental health intervention whilst providing an alternative approach to fill the much-needed self-care gap.

Objective: Assess the quality and characteristics of chatbots for anxiety and depression available on Android and iOS systems.

Methods: A search was performed in the App Store and Google Play Store following the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) protocol to identify existing chatbot apps for anxiety and depression. Eligibility was assessed by two individuals based on predefined eligibility criteria. Meta-data of the included chatbots and their characteristics were extracted from their description and post-installation by two reviewers. Information on quality was assessed by following the mHONcode principles.

Results: Only a handful (n=11) of chatbots were included from an initial search of 1000 that provide a substitute for human-human based interaction and clearly had a therapeutic human substitute goal in mind. The majority of reviewed apps had a high number of downloads indicating their popularity. The apps were also of a general high quality based on our assessment according to the mHONcode principles.

Conclusion: The general popularity of apps reviewed, and results of our quality assessment indicate chatbots have a promising future within the realm of anxiety and depression. Anxiety and depression chatbot apps have the potential to increase the capacity of mental health self-care providing much needed low-cost assistance to professionals.

1. Background

Mental health disorders such as anxiety and depression are at an all-time high, globally, over 264 million people of all ages suffer from depression alone [1]. Reports of anxiety amongst the population are also alarming, with anxiety figures reaching 3.76% of the global world population [2]. Anxiety and depression can have wider impacts on society leading to a decline in job performance and high rates of sick leave. Health care professionals are not immune to such disorders which is especially concerning knowing they are critical to the functioning of vital services, in times of a pandemic health professionals have been known to encounter severe stress (a sub syndrome of anxiety) [3-5]. When faced

with an overwhelming event such as a pandemic the emotional and behavioural reactions of health care professionals and patients is believed to be a normal and adaptive response to stress [6].

Psychological and pharmacological approaches are amongst some of the known effective treatments for moderate and severe depression. With limited evidence to suggest successful intervention for mental health disorders such as anxiety and depression, studies suggest interventions that have a combination of cognitive behavioural techniques (CBT) and work relaxation techniques can be effective [3].

Reports suggest developed countries have only about 9 psychiatrists per 100,000 people [7], whilst low-income countries have as little as 0.1 for every 1,000,000 people [8]. Since the outset of the COVID-19

* Corresponding author.

E-mail address: mhouseh@hbku.edu.qa (M. Househ).

pandemic, stress, anxiety, and depression cases have risen dramatically largely caused by quarantines, lockdowns, self-isolation, and a general feeling of fear [9, 10]. The huge effect of social stigma and mental illness shaming have prevented a considerable number of patients from seeking the proper medical help. This gap could be minimized through the usage of mental health chatbot apps assisting mental health professionals via self-care (self-management) approaches.

With the increase in access to mobile internet and smartphones, smartphone-based mental health (mHealth) apps represent a unique opportunity to expand the availability and quality of mental health self-care. In 2015 alone, according to a survey conducted by the World Health Organization (WHO) of 15,000 mHealth apps exist of which 29% focus on mental health diagnosis, treatment, or support [5]. The UK's National Health Service (NHS) and the U.S. National Institute of Mental Health (NIMH) are amongst some of the public health organisations that have pointed to mental health apps as cost-effective and scalable solutions to address the mental health treatment gap. Although the number of mHealth apps is promising, it should be considered with caution as very few have been thoroughly tested or are evidence-based [11]. Our study concentrates on anxiety and depression aiming to provide some insight on the features of chatbots available targeting these two disorders.

Several quality evaluation resources exist, including checklists, rating scales, and evaluation models with varying degrees of details and recommendations. They provide means of establishing the trustworthiness of an app to principles that guide clinicians on what to look out for when recommending an app to patients [17], ultimately inspiring users with more confidence for app usage.

The mHONcode principles aim at identifying the source information (medical evidence) and naming the medical team(s) involved in the development of mental health apps. The mHONcode also seeks to identify information age as outdated information could hinder the ability to make good health decisions, it only validates apps containing up to date information. The mHONcode has eight principles which include authority, complementarity, confidentiality, validity, justifiability and objectivity, user's practice, financial disclosure, and advertisement policy [17].

Using the Google Play Store [12] or Apple Store [13] on a smartphone to search words such as 'depression' returns a range of categories of apps containing chatbots that include diagnosis, mood tracking, and depression cure features. A user will be overwhelmed with the number of apps available. Health professionals and patients require clear guidance on which mobile app based chatbots to use or recommend, often individual developers or commercial mobile app companies whose primary purpose is to generate revenue are not always supported by scientific evidence or input from medical professionals.

Although this is not always the case, fortunately some evidence-based apps containing chatbots do exist (Ada, Replika, Wysa and Youper being among some of the most popular ones that we reviewed). Nonetheless, detailed reviews of such apps are needed to compare the kind of features on offer. Evaluation of every mental health related mobile based chatbot app would be cumbersome, and time consuming. Earlier studies have focussed on mental health apps in general or on one specific disorder such as anxiety [14]. There are several mobile app reviews, but they have a few shortcomings and/or limitations worth mentioning.

We did not come across academic reviews for chatbot apps available free at point of download, we noted subjective information is available normally affiliated to the app itself based on user reviews. To the best of our knowledge no recent review has been conducted following quality evaluation resources such as mHONcode principles. Furthermore, we did not come across recent reviews that focus on self-management (self-care) and treatment-related functionality and characteristics of highly rated anxiety and depression apps containing chatbots. Our study aims to identify the characteristics of mobile mental health apps, we assess the quality of chatbots specialized to assist individuals with anxiety and depression. We attempt to use objective reporting using an adapted ver-

sion of existing guidelines, which will help healthcare professionals as well as individuals to make a more informed decision about which apps to use or recommend based known measures of quality as opposed to subjective reviews.

2. Methods

We conducted a review of anxiety and depression apps containing chatbots available on the Google Play Store (Android) and Apple Store (iOS). We followed the guidelines set out by the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) group [15]. The use of PRISMA ensures clarity and transparency of reporting as well as permitting readers to replicate reviews. Furthermore, we adapted the mHONcode guidelines (Fig. 1), which amongst other objectives aims to identify the source information (medical evidence) and naming the medical team(s) involved in the development of mental health apps. We added further categories to the mHONcode principles to objectively test the quality of the included apps.

2.1. Search strategy

2.1.1. Search sources

The Google Play Store (Android) and Apple Store (iOS) were the two sources used for identifying the chatbots related to anxiety and depression. Google Play and the Apple Store are the two most popular app repositories with over 45,791 health care apps available in the Apple Store [13] and 46,360 in the Google Play Store [12]. Our search was conducted between the 13th and 22nd of October 2020. When apps are viewed in the app stores, the app stores recommend further apps to a user. These apps were also checked and included if they met our inclusion criteria. We also performed a simple Google search as a means of decreasing the likelihood of missed apps that the app search engines overlooked. We analyzed the first three pages of results as google automatically arranges the most relevant results in ascending order.

2.1.2. Search terms

This review investigates anxiety and depression related health care apps containing chatbots available on Google Play Store (Android) and Apple Store. Therefore, the words "anxiety" and "depression" were used as the search terms for both app repositories. The terms "anxiety" and "depression" were automatically applied to both titles and descriptions of Android and iOS apps [16]. The Google web search tool was also used for applying the search term "anxiety and depression chatbots" as one word. Analysing the first three pages of Google search engine increased likelihood that we included all apps even if the app store searches overlooked these. This was not part of our main search strategy rather a way of being thorough in our search.

2.2. Study eligibility criteria

For apps to be eligible as part of this study they had to contain chatbots for the purpose of anxiety or depression. Below is the detailed outline of the inclusion and exclusion criteria that was developed prior to beginning the searches.

Inclusion Criteria (IC):

- IC1: The first inclusion criteria were that it should be an app related to either anxiety or depression or both.
- IC2: Apps should contain a chatbot feature as opposed to human interaction, as the main chatting agent. Operationally this is when the user is speaking to a virtual agent via text, as opposed to a real live person conversation.
- IC3: The apps should be free of cost at point of download. Free-to-download apps with in-app purchase features were included, if their full chatbot features were included in the free version.
- IC4: Apps were included only if their rating was above four stars as a sign of user satisfaction.



Fig. 1. mHONcode principles.

- IC5: Based on consultations with mobile app experts and consensus of authors, apps that had a total number of five raters or less were excluded. Relatively new apps with only a few users, for example 5-star feedback can falsely be perceived as good apps. This can be misleading compared to apps that have 1000s of ratings but a lower average score.

Exclusion Criteria (EC):

- EC1: Filtering and excluding apps that do not contain a clear chatbot feature.
- EC2: Apps that do not have English as their Primary or Secondary Language.

2.3. Apps selection

The app repositories were searched separately by NA and SA. NA searched for the apps in the Apple Store and SA in the Play Store. In the first step, the apps were selected based on the information provided in their respective description in the Google Play Store and Apple Store. The first step served almost like an initial screening stage, for example if by reading the description an app suggests it has a chatbot feature we included it in our study. Secondly, the set of selected apps were individually downloaded on a smartphone device and assessed, and all the recommended apps were screened. If an app did not meet other inclusion/exclusion criteria, it would then be excluded at this stage. Once apps were downloaded, recommendations appear for further apps within the app store. These apps were also assessed and included or excluded according to our exclusion/inclusion criteria.

2.4. Data extraction

We developed a data extraction form (Appendix A) where data about each app was captured. The extraction process was carried out in two phases. In phase one, features were extracted from the descriptions of

each respective app. In phase two, we downloaded all included apps and further assessed their functionality thus extracting the required features. An iPhoneX (IOS 13) and Samsung A10 (Android 10) were used to assess the apps and their functionalities and preplanned features were recorded in the data extraction sheet. Two reviewers carried out the feature extraction process independently on the respective devices.

2.5. Data synthesis

We used a narrative approach for data synthesis. We assessed the characteristics of chatbots according to the technology used, usability, disease focus and other general features to provide recommendations for future research (Appendix A). We assessed the quality of the chatbots according to the mHONcode principles (Fig. 1) adapting it to our own needs by removing and adding categories.

2.6. Quality assessment framework

To evaluate the quality of apps, we used the mHONcode principles, providing users who download our reviewed apps a certain level of confidence. The mHONcode authors offers developers of apps an opportunity to gain certification of health mobile applications if its principles are proven to have been adhered to, the principles were clearly outlined, easy to follow and met some of the aims we wanted to achieve from our study. Consequently, the layout of the results section for this study follows the mHONcode eight principles albeit partially as shown in Fig. 1.

Evaluating the eight principles of the mHONcode model, we found that some of the categories were not relevant for our study. Therefore, we removed, renamed, and added some categories according to the objectives of this study. In addition, we incorporated more features within the existing categories and left out some we did not see fit for our purpose.

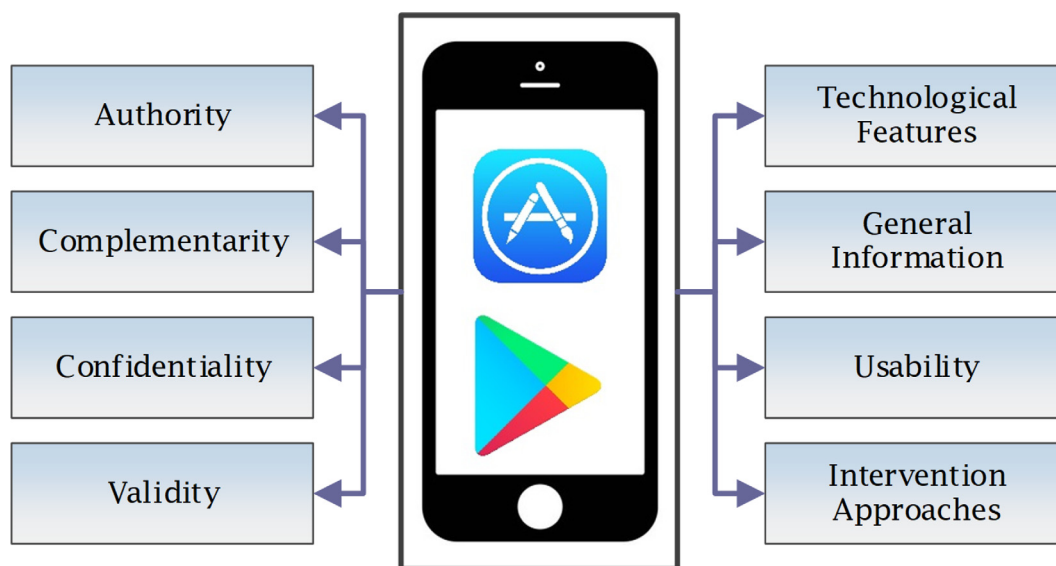


Fig. 2. 8 Principal Model (Adapted version of mHONcode with additional categories and some removed creating our own model).

Only one category, Advertisement Policy, was removed because it was not relevant to our study. Some other categories were updated and modified based on authors’ consensus and according to the objectives of the study. For example, the User Practices category includes user friendliness and clarify of mission. Accordingly, we renamed the User Practices category to Usability, to make the category easier to understand by readers.

We added a Technological Features category, as this was very crucial to the aims of our study. This category aims to identify what technologies and levels of intelligence are used for the purpose of the chatbot, for example, is it AI driven or not?

The Technological Features category was added to reflect the nature of this study as it was not one of the original mHONcode categories; it investigates the technological features about the apps. Similarly, our added General Information category covered both the Objectivity and Financial Disclosure categories of the original model. The modified model is shown in Fig. 2.

Upon downloading all 11 apps, two reviewers noted their experience on the usability of the apps, though we realize this is highly subjective we felt we could extract useful information conducting such a process. Two independent adult users tested the apps on an iPhoneX and Samsung A10.

3. Results

3.1. Search results

The app stores were initially searched using the two keywords “anxiety” or “depression” and returned a total of 1,069 apps, 550 from Google Play Store and 519 from the Apple Store. Many apps were returned with both ‘anxiety’ and the ‘depression’ keyword; these apps were considered as duplicates. Descriptions of the apps were read which are noted in each app store providing information to users before download. The descriptions allowed initial features to be extracted allowing application of the inclusion and exclusion criteria. Only 89 apps, 42 from Google Play Store and 47 from Apple Store were shortlisted. A total of eight apps satisfied the inclusion and exclusion criteria and were included in the review. Upon downloading the apps and analyzing chatbot functionality a further three apps were discovered as recommendations within the app stores. Eventually, 11 unique apps were selected for the purpose of this review (Fig. 3).

Table 1

Names and author given reference codes of the 11 reviewed Apps.

App Code	App Name
A1	Ada – your health companion
A2	InnerHour Self-Care Therapy - Anxiety & Depression
A3	Mindspa: Self Help 4UR Mental Wellbeing & Wellness
A4	Replika: My AI Friend
A5	Serenity: Guided Mental Health
A6	Woebot: your self-care expert in CBT & mindfulness
A7	Wysa: stress, sleep & mindfulness therapy chatbot
A8	Youper - Anxiety & Depression
A9	MindDoc: Depression & Anxiety
A10	Dr. Sila - your smart health assistant
A11	Pocketcoach - Anxiety Helper

All eleven apps were given reference codes (Table 1). Table 1 shows a list of the 11 apps with serial codes from A1 to A11. Having conducted a Google search and analyzing the results of three pages, we found the returned results were either already included in our study or were web-based apps which we already decided did not meet our inclusion criteria (Fig. 3).

3.2. Quality of the apps

3.2.1. Authority

The authority category aims to establish whether the editorial and team involved in the app development are named and if information on these are available to potential users. All the eleven included apps have clearly mentioned those involved in developing the apps either directly or otherwise through their associated websites. This category allows users to establish the trustworthiness and look further into who the developers of the apps are.

3.2.2. Complementarity

The complementarity category aims to establish information about the medical aspect of the app. This is achieved by including clear information about the target disease, the management approach, and the limitations of the app from a medical aspect. All the apps included in this study clearly stated that their apps do not replace the need of a professional and can only be used as an initial tool to understand oneself. They have explicitly agreed that these apps are not meant for diagnosis, cure, or treatment. This statement is often mentioned more than once in the user agreement section.

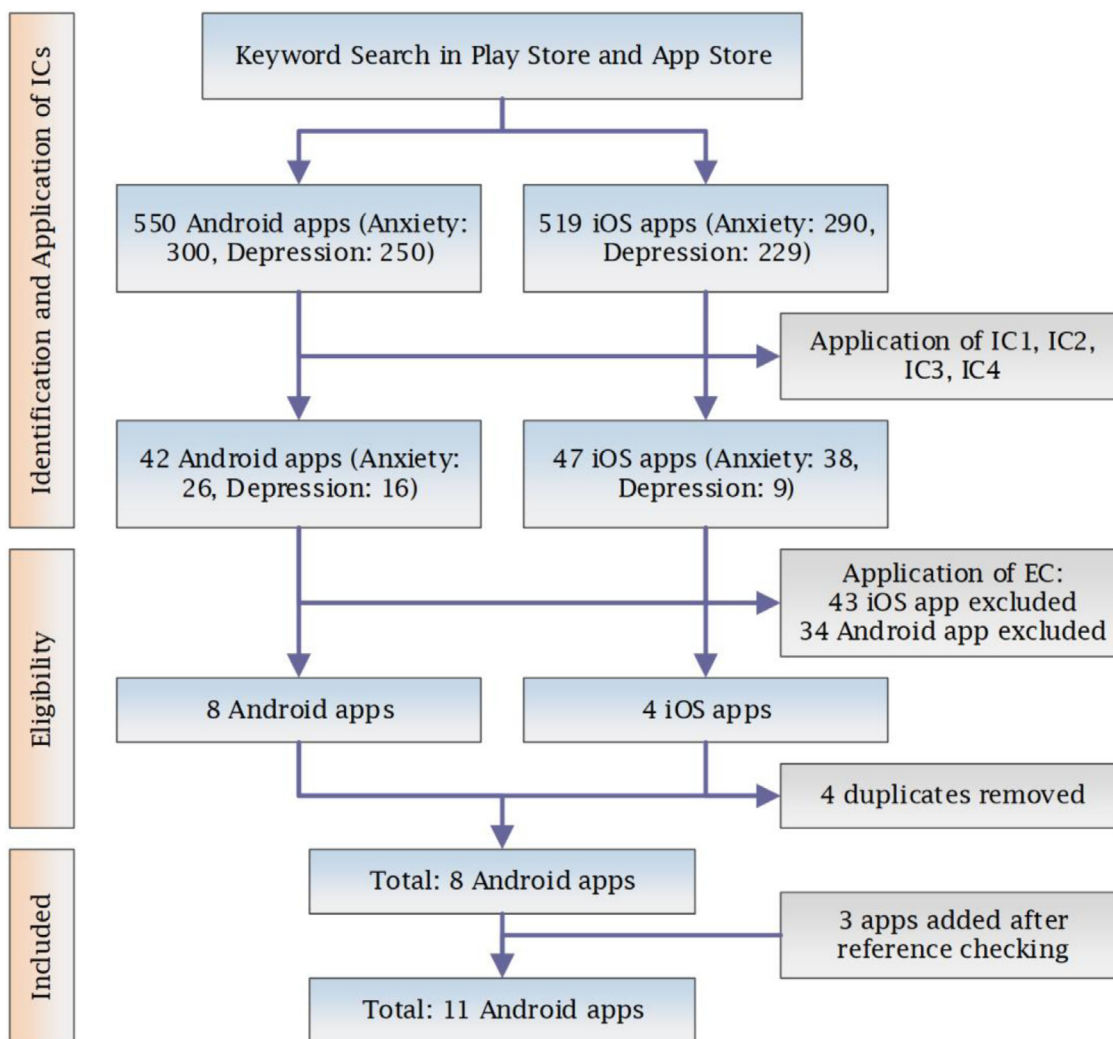


Fig. 3. Flow chart of the study selection process.

Table 2 Target mental health disorders.

Target Mental health issue	App ID	Total
Anxiety	A1, A2, A4, A5, A6, A7, A8, A9, A11	9
Depression	A1, A2, A4, A5, A6, A7, A8, A9	8
Stress	A2, A6, A8, A9, A11	5
Sleep	A2, A5, A7	3
Others	A1, A2, A3, A5, A6, A7, A9, A10, A11	9

Table 3 Management approaches.

Approach	App ID	Total
Mindfulness	A2, A3, A5, A6, A7, A8, A11	7
Mood tracking	A2, A6, A8, A9	4
Meditation	A3, A7	2
Assessment tools	A1, A2, A3	3

Although we targeted anxiety and depression apps only, most of the apps would also cover other disorders such as stress, as outlined in Table 2. Therefore, we included this information in our data extraction. Given that our inclusion criteria were to select apps for anxiety and depression, naturally all the included apps targeted either anxiety or depression through their services. Additionally, five apps targeted stress related problems, which often come with anxiety or is a subcategory of anxiety. Three of the apps incorporated solutions for sleep problems and nine apps included solutions for other mental health conditions, such as social relationship problems, procrastination, loneliness, grief, addiction, pain grief, guilt, distrust, jealousy, envy, shyness, apathy, and other complex feelings management, even stomach problems and headaches. We included them in our findings but reserved detail data extraction for anxiety and depression related features only. Where solutions are mentioned, they are reported in the apps that we reviewed

in the form of therapy, such as mindfulness, CBT, mood trackers and other assessment tools. None of the apps involved any direct advice on medical interventions.

From the aspect of disease management seven of the apps deal with mindfulness (Table 3). Mindfulness is a kind of meditation wherein the focus of the person is to become strongly aware of what they are feeling now without being interrupted or judgmental [18]. While four apps also have daily moods trackers, three apps also can be used for assessing daily behaviors. Meditation features were only seen in two of the apps.

3.2.3. Confidentiality

The confidentiality category explores whether there are statements explaining the legal requirements about the confidentiality of personal data. All apps that were included in this review explain this issue in their privacy policy section. To prevent unauthorized access to the personal data, these apps have made use of the industry standard safeguards (as

Table 4
General features of included apps (N=11).

Features	Features	Apps	Total
Category	Health & Fitness	A2, A3, A4, A5, A7, A11	6
	Medical	A1, A6, A8, A9, A10	5
Cost	Free	A3, A5, A6, A10	4
	In-app Purchase	A1, A2, A4, A7, A8, A9, A11	7
Language	English	A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11	11
	French	A1, A5	2
	German	A1, A5, A9, A11	4
	Portuguese	A1, A5	2
	Romanian	A1	1
	Spanish	A1, A5	2
	Italian	A3	1
	Russian	A3	1
	Arabic	A10	1

Table 5
General information about the apps in both systems.

App	Rating		No. of Raters		Size (MB)		System versions		Age rating		No. of Installations Android Only
	iOS	Android	iOS	Android	iOS	Android	iOS	Android	iOS	Android	
A1	4.7	4.7	103	284619	33.2	9.3	9.3+	4.3 +	17+	Everyone	5,000,000+
A2	4.0	4.4	21	5994	81.2	15	12.0 +	5.0 +	17+	3+	500,000+
A3	5.0	4.9	16	1143	127.3	Varies	10.3 +	5.0 +	17+	3+	100,000+
A4	4.7	4.5	110.1K	280055	164.3	54	11.2 +	5.0 +	12+	16+	5,000,000+
A5	4.6	3.7	14	132	77.3	N/A	10.0 +	N/A	12+	3+	10,000+
A6	4.8	4.7	3.7K	8449	44.1	43	11.0 +	6.0+	12+	3+	100,000+
A7	4.8	4.7	3.5K	61888	72.9	8.3	10.0 +	4.1+	12+	Everyone	1,000,000+
A8	4.9	4.9	14K	49031	124.7	83	11.0 +	7.0 +	12+	N/A	1,000,000+
A9	4.6	4.6	2.1K	34340	20.8	44	13.0 +	6.0 +	12+	Everyone	1,000,000+
A10	-	5	-	8	37.9	35	12.0 +	5.0 +	17+	3+	50+
A11	4.6	4.1	52	175	57.9	32	9.0 +	5.0 +	12+	Everyone	10,000+

per google or apple play policies) and the encrypted data is securely stored in virtual private cloud servers in a physical location, no further details are provided for most of the apps

3.2.4. Validity

This category aims to establish whether the health content within the app is updated periodically. We observed that all 11 apps have been updated in the ear the study was conducted (2020). However, although some apps mention the date of latest update, others fail to mention the exact date.

3.2.5. Technological features

As mentioned in the introduction one of the main objectives of this review was to identify the apps that contain some sort of a chatbot feature related to anxiety or depression, surprisingly very few apps that initially fulfilled our keyword search criteria had a true non-human chatbot feature. Eventually this was narrowed down to 11 apps. Of the 11 apps reviewed, five explicitly mention AI or intelligence as a key driving force of the chatbot while the remaining six were also marketed as chatbots or chat companions (without explicitly mentioning AI). Only three of these apps’ highlights AI (Artificial Intelligent) as part of their intervention approach (Replica, Wysa and Youper).

3.2.6. General information

Table 4 and Table 5 present the general information about the apps. Table 4 illustrates some of the general features which are common in both stores such as the app category, cost, language. Six out of the eleven included apps belonged to the Health & Fitness category while the remaining five belonged to the Medical category. Many of the apps (7/11, 63.6%) offer some in-app purchases whilst the remainder (4/11, 36%) were found to be completely free. Often app developers are using in-app purchases in order to generate income, it is difficult to come across any highly rated popular app that is completely free from that perspective, some of the apps offer in-app purchases whilst others have full

functionality without any payment. All the apps are available in English language. However, six of the apps we reviewed had alternative languages as an option although these were not analyzed in the alternative language only in the English language version.

All except one of the apps were available in both Google Play Store and Apple Store we combined the characteristics from both, as in some cases information was available in the play store but not in app store. Table 5 highlights information on the differences observed between iOS and Android platforms including the rating, number of raters, size and system versions needed for the apps and age ratings, some features such as number of installations was available only on play store. The average user rating of the apps in the Apple Store was 4.67/5 and for the Play Store 4.56/5. The Serenity app has a rating of 3.7/5 in the Google Play Store having been released recently to the market (January 2020), while in the Apple Store has a rating of 4.6, therefore it was included in this study on that basis.

The number of raters ranged from 14 to 110.1K in the Apple Store, this range was from 8 to 284619 raters in the Google Play Store. The size of the included apps varies from 20.8 MB to 164.3 Mb for the iOS versions and 8.3 MB to 83 MB for the Android versions. For the system versions in the iOS apps, it varied from iOS 9.0 to iOS 13.0, while in the Android apps it ranged from Android 4.0 to Android 7.0 or later. For the age ratings in the iOS apps, there were seven 12+ rated apps and four 17+ apps. In the Android apps age ratings, there were five 3+ rated apps, one app with 16+ age rating and three apps that were available for everyone. The number of installations information is available only on the Google Play Store, this varied from 50+ installations to 5,000,000+ installations with the median value of 500,000.

3.2.7. Usability

This section describes the way in which a user interacts with the chatbot. The main purpose of all the included chatbots was to provide some form of therapy to the users, one of the apps also could be used for assessment purposes. Regarding the input and output modality, all the

chatbots main interaction method was written text with the Wysa app having an additional option for voice chats. Four of the apps showed recommendations for responses by users as opposed to free-text and if the questions were short or multiple-choice type questions, nine among the eleven apps showed various options to choose from as a response. However, two apps Wysa and Replika always had their keyboard always displayed whilst chatting.

In terms of embodiment and to enhance the user experience of the app seven chatbots had some fictional character or avatar, this is often intended to form some virtual friend type feeling as well as removing barrier so the user does not feel they are talking to a real person.

The Ada chatbot focusses on symptom assessments through their chats, the Dr. Sila app can assess any symptom whether mental or physical and provides recommendations or choices. Chatbots like Replika, Inner Hour, Woebot, Wysa, Youper, MindDoc and Pocketcoach can be used for any sort of conversations allowing the users to engage in small talk. These chatbots prompt for details about the user's feelings and emotions and the user in turn can converse about anything. The Replika chatbot is an AI chatbot that is learning on the go. Mindspa chatbot is designed as an emergency only chatbot and not for light conversation. The Wysa chatbot has an option to add a therapist in the chat itself. Woebot, Wysa, Youper, MindDoc and Pocketcoach can be used even without creating an account in their app while all others require the user to create an account before use.

3.2.8. Intervention approaches

Within our study, only one app claimed a science backed approach while another claimed to be using a therapy-based approach, the remainder ten apps made no mention of what the development techniques of the app was based on. Six out of the eleven apps [A2, A5, A6, A7, A8, A9] explicitly revealed cognitive behavioral therapy (CBT) as an intervention technique. CBT is a therapy that tries to modify and change individuals' dysfunctional cognitive distortions like thoughts and beliefs. CBT is a therapy that tries to modify and change individuals' dysfunctional cognitive distortions like thoughts and beliefs [19]. Intervention approaches of two of the included apps are mentioned as Dialectical Behavior Therapy (DBT). While CBT identifies negative thinking patterns and promotes positive behavioral changes, DBT can be used to identify and treat self-destructive behaviors such as self-harm, drug abuse and suicide ideation [20].

Although few apps claimed scientific backing, six apps mentioned involvement of a mental health expert, therapist, or doctor. Finally, one app offered a depression test feature and one featured guidance from a real human coach.

4. Discussion

In this review, mental health related apps, that incorporate chatbots or conversational agents particularly for anxiety and depression, were outlined. These include apps available on Google Play Store (Android) and Apple Store. However, some of these chatbots also targeted other mental health issues such as sleep and stress. A total of 11 chatbot containing apps were included in this review collected from the Google Play and Apple Store in which (6/11) were related to health and fitness whilst (5/11) were medical-related apps. The main self-help (management) interventions focused on in this review include mindfulness, mood tracking, meditation, and functional assessments. The most common mental conditions amongst the chatbot apps were anxiety and depression (9/11) and (8/11), respectively, in line with our search strategy. The most common management approach amongst the apps was mindfulness (7/11), as illustrated in Table 3.

Moreover, the study observed some general characteristics of mental health apps in Table 4 and only included apps with rating 4+ in this review in an attempt to not include lowly rated apps. The app Mindspa received the highest rating (4.9/5), however, with fewer installations (100, 000+) compared to some of the others. This could be down to

the fact it is completely free of charge therefore with less funding revenue. The mindspa app contains an emergency reporting based chatbot feature helping individuals with anxiety and depression as opposed to the personalized avatar style bots seen in some of the other apps. The targeted age group of majority of the apps reviewed was 12 years and above, outlining possible additional challenges faced when providing such services to children.

It is important for apps claiming to provide therapy in whichever form that they highlight the evidence for their techniques and their therapeutic approach, which has previously been dismissed in many applications as a cause for causing harm for users [21]. Using such models as the mHONcode should ensure such principles are adhered to. A recent review estimated 293 mental health apps providing evidence-based practice framework involving treatment and comprehensive therapeutic intervention specifically for anxiety and depression. The management approaches involved in other studies include CBT, Mindfulness (MIND), Positive psychology (POS), Dialectical behavior therapy (DBT), and Acceptance and commitment therapy (ACT) [21]. However, chatbot features involved in this review were promoted as a basic feature that could help individuals with anxiety and depression.

Similarly, most of our findings reveal cognitive behavioral therapy to be the leading framework for e-mental health programs as CBT has been effective in lowering stress and anxiety for the past two decades [22]. Previous studies have reviewed chatbots for mental health conditions including autism, stress, anxiety, and depression while targeting standalone software [23] and various studies examined the characteristics of highly rated applications. Some of the mental health conditions covered in such studies included PTSD (Post Traumatic Stress Disorder), OCD (Obsessive Compulsive Disorder), suicide, self-harm, and ADHD. In addition to the high ratings in Google Play Store and Apple Store, the high number of installations suggests that the overall quality of these apps that contain a chatbot feature remains high. This study adds value to those planning to offer similar evidence-based services to anxiety and depression sufferers by aiding their decision-making process when looking for such apps in a crowded marketplace. To the best of our knowledge, there are no research publications linked to chatbots and their management features for anxiety and depression combined with a study of management approaches and usability features for well-known applications such as Wysa, Woebot, MindDoc, Youper, and Replika.

4.1. Strengths and limitations

This review provides valuable insights into chatbot applications available on Google Play Store and Apple Store that provide therapeutic interventions and could help individuals suffering from anxiety and depression, providing a real insight into the best available tools by feature breakdowns and user ratings. We add value by using known measures from the mHONcode principles as opposed to half hazard analysis based on user subjective reviews. Some scoping reviews concentrate on apps reported in literature only, often such scoping reviews include proposals and web-based applications reported in literature only. This study concentrates on apps available to install and use free at point of download to anyone with access to Google Play Store and Apple Store running mobile devices. This review is dedicated to facilitating entities and stakeholders who are concerned about chatbots with outlined management features such as mindfulness, mood tracking and meditation. We incorporated most of the principles of mHONcode app certification into the 11 included applications along with categories we felt important to readers; this should give some strength over previous studies and a level of objectivity to the study.

As with most studies of this nature a degree of caution must be taken. First, the app selection, particularly the search criteria, was limited to the keyword's "anxiety" and "depression" only, which could hinder extracting some apps that fall within our inclusion criteria. The search keywords used returned apps containing features such as stress-related therapies, in a similar manner no doubt other search terms may have re-

turned chatbot apps with anxiety or depression related features which we potentially overlooked. Secondly, the search was limited to the English language only; we observed for example, that only one of our returned apps was in the Arabic language, this may highlight a huge gap in the market for platforms dedicated to Arabic-speaking populations, which excludes a large part of the world population. Thirdly, many of the apps we found from research papers or through backward reference checking were not found in either app store or play store and were either outdated or unavailable for unknown reasons. Despite following some of the mHONcode principles to provide some sort of objectivity to the study coupled with our incorporated principles, some of our findings remain highly subjective based on the individual reviewer's experiences.

4.2. Practical implications

Further work is needed in terms of mental health services for non-English speaking populations. We might find mental health-related apps for anxiety and depression. However, few involve chatbots as a therapeutic intervention on a multi-lingual level, at least within the methodology used in the search for this review. This review also provides users and chatbot developers with the major dimensions that should guide developing better chatbots and their corresponding features that can gain wide acceptance among users and achieve better results for patients, such as the authority of apps developers, complementarity, the confidentiality of user information, validity requirements, critical technological features, practical usability, and effective intervention approaches.

4.3. Research implications

Multiple gaps were noted in relation to chatbot technologies dedicated to anxiety and depression. Few applications attempted to outline the evidence-based framework behind the proposed intervention. Future studies on how useful these apps are in reducing the psychological impact of depression and anxiety should also be explored. This will help recognise applications and features that will be useful in the future. Furthermore, a full thematic analysis of the review comments available within each of the app stores could provide further insight in to the positive and negative features of the chatbots reviewed in this study and their related management features. Including such a thematic study was beyond the scope of this study but the authors plan to build upon this review's findings in upcoming studies.

5. Conclusion

Although many mental health disorder related apps exist, we only found 11 chatbot related apps that met our inclusion/exclusion criteria available on Google Play Store and Apple Store. These 11 apps are targeting anxiety and depression disorders. All the apps containing these chatbots were popular, based on the high number of installations. Moreover, we found them to be of general high quality and popularity based on the fact they all met most of the mHONcode principles and are among the highly rated apps. The apps we reviewed were built with various self-care interventions and other types of interventions, such as CBT and mindfulness. Moreover, few applications attempted to outline the evidence-based framework behind the proposed intervention. Anxiety and depression chatbot apps have the potential to increase the capacity of mental health self-care providing much needed low-cost assistance to professionals. Future developers of chatbots need to focus on evidence-based interventions to allow professionals and users to have greater confidence in the technology.

Declaration of competing interest

The authors have no competing interests to declare.

Acknowledgments

This study was a part of a project funded by the Qatar National Research Fund (NPRP12S-0303-190204). The project title is "A Personalized and Intelligent Digital Mental Health Platform for Qatar and the Arab world".

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpbup.2021.100012.

References

- [1] Organisation, W.H. Depression – Key facts. [cited 2020 October 13th]; Available from: <https://www.who.int/en/news-room/fact-sheets/detail/depression>.
- [2] Elflein, J. Share of the Population Worldwide Who Suffered From Anxiety Disorders From 1990 to 2017. 2019 [cited 2020 13th October]; Available from: <https://www.statista.com/statistics/1035153/percentage-of-people-with-anxiety-worldwide/>.
- [3] V. Breninkmeijer, et al., Predicting the Effectiveness of Work-Focused CBT for Common Mental Disorders: The Influence of Baseline Self-Efficacy, Depression and Anxiety, *J. Occup. Rehabil.* 29 (1) (2019) 31–41.
- [4] D Chang, H Xu, A Rebaza, L Sharma, CS Dela Cruz, Protecting HCWs from subclinical coronavirus infection, *Lancet Respir Med* 8 (2020) e13.
- [5] SM Lee, WS Kang, A-R Cho, T Kim, JK. Park, Psychological Impact of the 2015 MERS outbreak on hospital workers and quarantined hemodialysis patients, *Compr Psychiatry* 87 (2018) 123–127.
- [6] R Maunder, J Hunter, L Vincent, J Bennett, N Peladeau, M Leszcz, et al., The immediate psychological and occupational impact of the 2003 SARS outbreak in a teaching hospital, *CMAJ* 168 (2003) 1245–1251.
- [7] C.J. Murray, et al., Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010, *Lancet North Am. Ed.* 380 (9859) (2012) 2197–2223.
- [8] B.D. Oladeji, O. Gureje, Brain drain: a challenge to global mental health, *BJPsych international* 13 (3) (2016) 61–63.
- [9] S.K. Brooks, et al., The psychological impact of quarantine and how to reduce it: rapid review of the evidence, *Lancet North Am. Ed.* (2020).
- [10] Shihabuddin, L. How to Manage Stress and Anxiety from Coronavirus (COVID-19). 2020 [cited 2020 13th October]; Available from: <https://www.rwjbh.org/blog/2020/march/how-to-manage-stress-and-anxiety-from-coronavirus/>.
- [11] E. Anthes, Mental health: There's an app for that, *Nature* 532 (7597) (2016) 20–23.
- [12] Statista, Number of mHealth Apps Available in the Apple App Store from 1st Quarter 2015 to 2nd Quarter 2020, (2020) (accessed October 2020), <https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/>.
- [13] Statista, Number of mHealth Apps Available in the Google Play Store from 1st Quarter 2015 to 2nd Quarter 2020, (2020) (accessed October 2020), <https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/>.
- [14] N. Drissi, et al., An analysis on self-management and treatment-related functionality and characteristics of highly rated anxiety apps, *Int J Med Inform* 141 (2020) 104243.
- [15] AC Tricco, E Lillie, W Zarin, KK O'Brien, H Colquhoun, D Levac, et al., PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation, *Ann Intern Med* 169 (7) (2018) 467–473 Oct 02 [doi: 10.7326/M18-0850] [Medline: 30178033].
- [16] 19 N. Drissi, S. Ouhbi, M.A.J. Idrissi, M Ghogho, An analysis on self-management and treatment-related functionality and characteristics of highly rated anxiety apps, *Int. J. Med. Inf.* (2020) 104243.
- [17] K. Anderson, O. Burford, L. Emmerton, App chronic disease checklist: protocol to evaluate mobile apps for chronic disease self-management, *JMIR research protocols* 5 (4) (2016) e204.
- [18] Langer, E. J. (1989). *Mindfulness*. Addison-Wesley/Addison Wesley Longman.
- [19] G.H. Eifert, M. Heffner, The effects of acceptance versus control contexts on avoidance of panic-related symptoms, *J. Behav. Ther. Exp. Psychiatry* 34 (3-4) (2003) 293–312.
- [20] L. Mehlum, A.J. Tørmøen, M. Ramberg, E. Haga, L.M. Diep, S. Laberg, ..., B. Grøholt, Dialectical behavior therapy for adolescents with repeated suicidal and self-harming behavior: a randomized trial, *J. Am. Acad. Child Adolesc Psychiatry* 53 (10) (2014) 1082–1091.
- [21] DD Ebert, A Zarski, H Christensen, Y Stikkelbroek, P Cuijpers, M Berking, et al., Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: A meta-analysis of randomized controlled outcome trials, *PLoS One* 10 (3) (2015) e0119895 [FREE Full text] [doi: 10.1371/journal.pone.0119895].
- [22] Jamie M. Marshall, Debra A. Dunstan, Warren Bartik, Apps with maps—anxiety and depression Mobile apps with evidence-based frameworks: systematic search of major app stores, *JMIR mental health* 7 (6) (2020) e16525.
- [23] Alaa A. Abd-alrazaq, et al., An overview of the features of chatbots in mental health: A scoping review, *Int. J. Med. Inf.* 132 (2019) 103978.