Review

Criteria for assessing the quality of mHealth apps: a systematic review

Rasool Nouri,[1] Sharareh R Niakan Kalhori,[1] Marjan Ghazisaeedi,[1] Guillaume Marchand,[2] and Mobin Yasini[2]

[1]Health Information Management Department, School of Allied Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran and [2]Research and Development Department, DMD Santé, Paris, France

Corresponding Author: Marjan Ghazisaeedi, PhD, Health Information Management Department, School of Allied Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran; tums.sdi@gmail.com

ABSTRACT

Objective: Review the existing studies including an assessment tool/method to assess the quality of mHealth apps; extract their criteria; and provide a classification of the collected criteria.

Methods: In accordance with the PRISMA statement, a literature search was conducted in MEDLINE, EMBase, ISI and Scopus for English language citations published from January 1, 2008 to December 22, 2016 for studies including tools or methods for quality assessment of mHealth apps. Two researchers screened the titles and abstracts of all retrieved citations against the inclusion and exclusion criteria. The full text of relevant papers was then individually examined by the same researchers. A senior researcher resolved eventual disagreements and confirmed the relevance of all included papers. The authors, date of publication, subject fields of target mHealth apps, development method, and assessment criteria were extracted from each paper. The extracted assessment criteria were then reviewed, compared, and classified by an expert panel of two medical informatics specialists and two health information management specialists.

Results: Twenty-three papers were included in the review. Thirty-eight main classes of assessment criteria were identified. These were reorganized by expert panel into 7 main classes (Design, Information/Content, Usability, Functionality, Ethical Issues, Security and Privacy, and User-perceived value) with 37 sub-classes of criteria.

Conclusions: There is a wide heterogeneity in assessment criteria for mHealth apps. It is necessary to define the exact meanings and degree of distinctness of each criterion. This will help to improve the existing tools and may lead to achieve a better comprehensive mHealth app assessment tool.

Key words: mobile applications, evaluation studies, mobile health

INTRODUCTION

Mobile health, or more commonly, mHealth, has been defined as "the use of wireless communication devices to support public health and clinical practice".[1] Mobile health applications have been defined as "soft wares that are incorporated into smartphones to improve health outcome, health research, and health care services."[2] For the general public, mHealth apps could empower patients to take a more active role in managing their own health.[3,4] They could more effectively engage patients,[5,6] positively influence their behavior, and potentially impact health outcomes.[7] Therefore, mHealth apps could help empower high-need and high-cost patients to self-manage their health.[8]

Management and control of diabetes, mental health, cardiovascular diseases, obesity, smoking cessation, cancer, pregnancy, birth, and child care are some examples of the targets of mHealth apps for

patients and the general public.[9–12] Healthcare professionals also use mHealth apps in performing important tasks including patient management, access to medical references and research, diagnosing medical conditions, access to health records, medical education and consulting, information gathering and processing, patient monitoring, and clinical decision-making.[13–19] With this diversity of use cases and the growth of the needs that could be addressed by mHealth apps, concerns arise about potential misinformation and the role of the health professional in recommending and using apps.

There is continuing worldwide growth in the number of these apps. Recent reports showed that there are more than 325 000 mHealth apps available on the primary app stores, and more than 84 000 mHealth app publishers have entered the market.[20] Despite the huge number of mHealth apps, a quarter of these apps are never used after installation. Many are of low quality, are not evidence-based,[21] and are developed without careful consideration of the characteristics of their intended user populations.[22] Users often pay little attention to the potential hazards and risks of mHealth apps.[23] This has led to an interest in better oversight and regulation of the information in these apps. Changes in mobile devices and software have been accompanied by a broadening discussion of quality and safety that has involved clinicians, policy groups, and, more recently, regulators. Compared to 2011, there is greater understanding of potential risks and more resources targeted towards medical app developers, aiming to improve the quality of medical apps.[24]

The decision to recommend an app to a patient can have serious consequences if its content is inaccurate or if the app is ineffective or even harmful. Healthcare providers and healthcare organizations are in a quandary: increasingly, patients are using existing health apps, but providers and organizations have quality and validity concerns, and do not know which ones to recommend.[25] Although there are various tools to assess the quality of health-related web sites, there is limited information and methods describing how to assess and evaluate mHealth apps.[26]

The issues concerning mHealth apps are wide-ranging, including location and selection of appropriate apps, privacy and security issues, the lack of evaluation standards for the apps, limited quality control, and the pressure to move into the mainstream of healthcare.[27]

Important reported limitations of mHealth apps include lack of evidence of clinical effectiveness, lack of integration with the health care delivery system, lack of formal evaluation and review, and potential threats to safety and privacy.[28]

Creating a comprehensive set of criteria that covers every aspect of mHealth app quality, seems to be a complex task. In this study, we aimed to review the existing papers that included an assessment tool or method to assess the quality of mHealth apps, to extract and summarize their criteria, if possible, and to provide a classification of the collected criteria.

Various stakeholders including app developers, citizens/patients, policy makers, health business owners, assessment bodies/regulators, clinicians/health professionals, authorities/public administration, funders/health insurance, and academic departments may find a consolidated view of the existing criteria helpful.

METHODS

This study was conducted according to the PRISMA[29] (Preferred Reporting Items for Systematic reviews and Meta-Analyses).

Search Strategy

A comprehensive literature search was conducted in MEDLINE (OVID), EMBase, ISI web of science and Scopus for English language citations published from January 1, 2008 (considering that the first mobile phone app store was started in mid-2008) to December 22, 2016. A researcher with a library and information science degree (RN) developed and carried out a Boolean search strategy using key words related to "mobile health applications" (eg mHealth apps OR mobile health applications OR mobile medical applications OR medical smartphone applications) and keywords related to quality assessment or scaling the mobile health applications (eg evaluation OR assessment OR measurement OR scaling OR scoring OR criteria). We used MeSH terms in Medline and EM Tree terms in EMBase and also truncation, wildcard and Proximity operators to strengthen the search (Supplementary Appendix 1).

Inclusion Criteria

The inclusion criteria were studies in English that provide a tool or method to evaluate mHealth apps and published from January 1, 2008 to December 22, 2016.

Exclusion Criteria

Studies in a language other than English, studies on mobile apps that were not related to medicine or health area, and those that contained a tool without presenting a scientific method for developing it, were excluded. Papers that focused only on the design and development of mHealth apps were also excluded.

Study Selection

After duplicates were removed, the titles and abstracts were screened by two researchers (SRNK, RN) according to the inclusion/exclusion criteria. The full texts of potentially relevant papers were then individually assessed by the same 2 researchers. Disagreements were resolved in consultation with a senior researcher (MY, the lead author), who also examined and confirmed the relevance of all included papers.

Data Extraction and Classification

Data elements extracted from selected articles, included authors, date of publication, subject fields of target mHealth apps, method used to develop the assessment tool, and the assessment criteria. The extracted criteria were then reviewed by an expert panel including two medical informatics specialists and two health information management specialists. The panel compared similar criteria from different studies and created categories that grouped all criteria relevant to a specific concept of evaluation (eg, usability). We analyzed each criterion and tried to find or create classes and subclasses that could encompass all criteria. Whenever a criterion did not match an already existing class, a new class was added. The criteria found in each of the included studies, the classes and subclasses were then listed incrementally, as they were discovered.

RESULTS

We retrieved 1057 records by searching the previously mentioned databases. After removing duplicates, 851 articles remained. Based on the review of titles and abstracts, 72 were found to have met the initial selection criteria. After examination of the full texts, 23 articles met the inclusion criteria and were included in the final review. Two of these articles were found by reviewing the references

(Figure 1). Most of the 23 articles were published in 2016 (39.1%) or 2015 (34.8%).

Some of the assessment criteria and tools described in the included studies[30–42] were developed based on a literature review. The reliability and validity of these tools were then evaluated in the respective studies. Others[43–52] were adapted versions of a selected array of existing tools for software or website evaluation. Many of the mobile apps evaluation criteria, which were reviewed in this investigation, were originally developed as website assessment criteria; Silberg,[37,44,47,51] HON code,[37] Kim Model,[30,33,38] Brief DISCERN,[37] HRWEF (Health-Related Website Evaluation Form)[36] and Abbott's scale[37] are examples of these instruments and scales.

There is great variety in the mHealth apps assessment criteria and their classifications in the articles reviewed. Jin and Kim[30] developed an evaluation tool for mHealth apps that contains 7 main criteria including accuracy, understandability, objectivity, consistency, suitability of design, accuracy of wording, and security. Stoyanov et al.[33] identified 5 main categories of criteria including 4 objective quality scales: engagement, aesthetics, functionality, and information quality, and one subjective quality scale, with 23 sub-items. Anderson et al.[34] presented a protocol for evaluating self-management mobile apps for the patients suffering from chronic diseases; 4 main groups of criteria including engagement, functionality, ease of use and information management were addressed. Taki et al.[36] presented 9 main groups of criteria: currency, author, design, navigation, content, accessibility, security, interactivity and connectivity, and software issues. Loy et al.[41] developed a tool for quality assessment of apps that targets medication-related problems. Their criteria were classified in 4 main sections including reliability, usability, privacy, and appropriateness to serve intended function. Reynoldson et al.[52] proposed another classification of criteria which included 4 main classes of criteria: product description, development team, clinical content, and ease of use. Therefore, there were not homogeneous definitions across the different sets of criteria we reviewed. This was due to not clearly defined or non-existent documented definitions for the assessment criteria in each set of criteria. However, the problem is not raised within each set of criteria and individual studies were consistent independently according to their context.

Most of the tools found in the included articles[30–33,39,42–45] were developed for no special subject category of mHealth apps. Other tools were developed for specific subject categories of mHealth apps including self-management of asthma,[46] cardiovascular diseases apps,[47] chronic diseases,[34] depression management and smoking cessation,[35] HIV prevention,[48] infant feeding,[36] medication-related problems,[41] mobile patient health records (PHRs),[40] self-management of pain,[52] panic disorder,[37] self-management of heart diseases,[49] prevention and treatment of tuberculosis[50] and weight loss.[51] Table 1 summarizes the main characteristics of each studied paper.

There were 38 main classes of criteria in the 23 papers reviewed: Accessibility, Accuracy, Advertising, Policy, Aesthetics, Appearance, Attribution, Authority, Availability, Complementarity, Consistency, credibility, Currency, design, disclosure, Ease of Use, Engagement, Ethical Issues, Financial disclosure, Functionality, Information/content, Interactivity and connectivity, Justifiability, learning, legal consistency, navigation, Objectivity, performance, Precision, privacy, reliability, Safety, Security, User-perceived value, Transparency, Understandability, Usability, usefulness, and Wording Accuracy. The criteria, the articles in which they appeared, and the details from each article (criteria sub-classes, descriptions and questions) are

included in Supplementary Appendix 2. These criteria were reclassified by our expert panel. In this process, some classes of criteria were merged, and the sub-class criteria were rearranged under the main class to provide a consolidated classification of evaluation criteria for mHealth apps. The consolidated classification contains 7 main classes of criteria including Design, Information/Content, Usability, Functionality, Ethical Issues, Security and Privacy and User-perceived value. These 7 main classes contain a total of 37 sub-classes of criteria (Figure 2). More details are presented in Supplementary Appendix 3.

Each of the main classes in the consolidated classification was mentioned in several studies, either directly or indirectly. "Design" was found in 9 different articles[30,31,33,34,36,38,48,49,52] "Information/Content" (using different terms) observed in 15 studies[30,33,34,36–38,41,42,44,45,47,49,51,52]"Usability" in 14 articles,[30,31,34,36,38–43,47–49,52] "Functionality" in 7 studies,[31,33,34,38,43,49,50] "Ethical Issues" in 1 study,[42] "Security and Privacy" in 8 articles[30–32,36,41,42,49,52] and finally "User-perceived value" in 4 studies.[31,33,38,49]

DISCUSSION

We conducted a systematic review of the studies that applied a type of evaluation method or an assessment approach for mHealth apps. After reviewing all these studies, we have classified the extracted criteria in 7 main groups including: design, information/content, usability, functionality, ethical issues, security and privacy, and User-perceived value. Each of these classes was divided into various sub-classes. In total, we identified 37 sub-classes of criteria.

The development and evaluation of tools for assessing mHealth tools is an area of active research interest. More than a third of the papers included in this review were published in 2016, our last year of coverage, and therefore not included in previous systematic reviews on the topic. BinDhim et al.[26] conducted a systematic review and tried to summarize the criteria used to assess the quality of mHealth apps and analyze their strengths and limitations. Their study was limited to the mHealth apps designed for consumers. More recently, Grundy et al.[60] conducted a systematic review to find and explain emerging and common methods for evaluating the quality of mHealth apps. They also provide a framework for assessing the quality of mHealth apps.

There are great differences in the way assessment criteria are defined and classified in the studies reviewed. For example, "usability" was treated in many different ways. Zapata, et al.[39] divided "Usability" in four sections; attractiveness, learnability, operability, and understandability. Brown et al.[43] had other sub-classes under "usability' including error prevention, completeness, memorability, information needs, flexibility/customizability, learnability, performance speed, and competency. In the study of Yasini et al.,[42] "usability" included ease of use, readability, information needs, operability, flexibility, user satisfaction, completeness, and user contentment by look and feeling perceived after using the app. Anderson et al.[34] viewed usability as a sub-class of ease of use, while Yasini et al.[42] and Loy et al.[41] placed ease of use under usability. Reynoldson et al.[52] classified ease of use and usability as 2 separate classes that each one has its own sub-classes. Stoyanov et al.[33] placed usability as a part of functionality.

In the study of Cruz Zapata et al.,[40] usability included style, behavior, and structure subclasses. In the study of Loy,[41] it included ease of use, user support, and ability to adapt to different user needs.
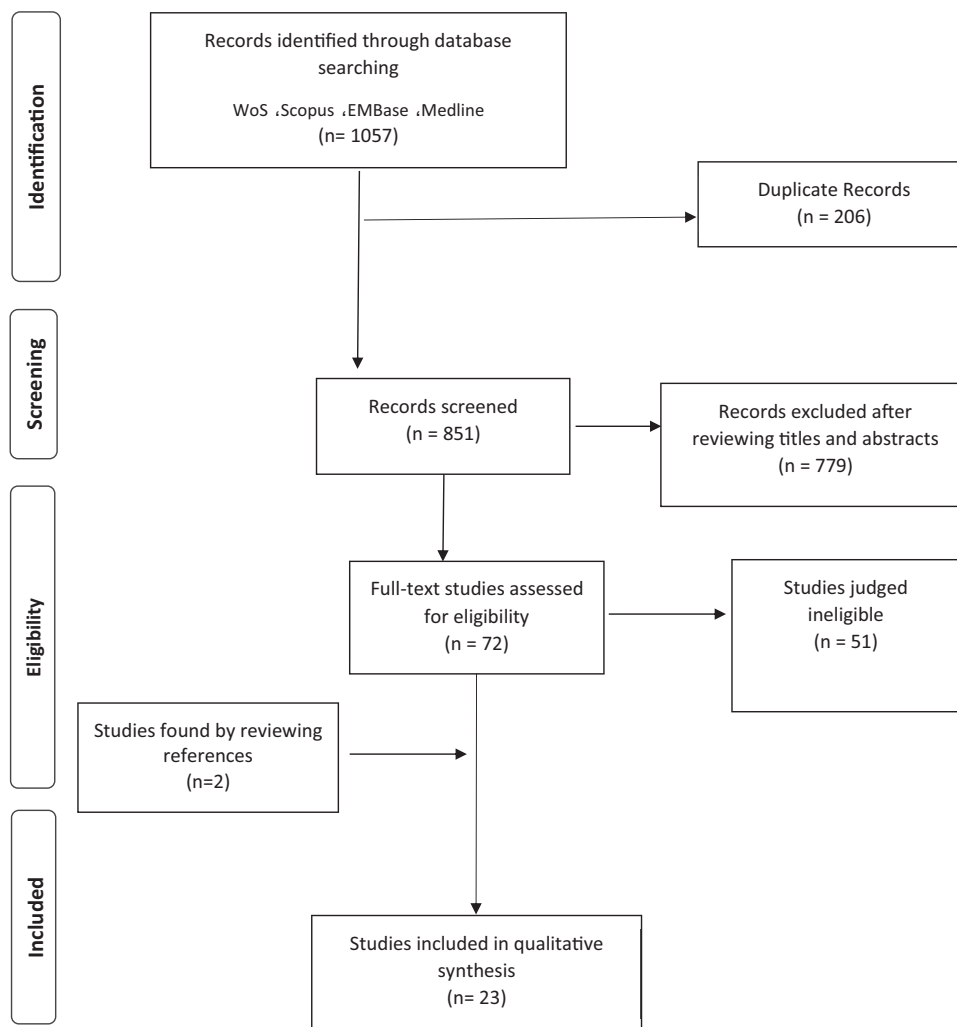
Figure 1. The flowchart that schematizes the approach to identify, screen and include the relevant studies in this review.

Similarly, "design" was identified in some studies[30,36,52] as a separate main criterion, but in others it was placed under 4 separate criteria including functionality,[33,34,38] usability,[48] consistency,[30] and engagement.[33,38] Design is a multi-dimensional criterion and may be considered from various viewpoints.

In some of the studies, criteria were used interchangeably or overlapped. For example, security, privacy, and safety overlapped in some sets of criteria, and there were different interpretations of these 3 concepts.[30,32,36,41,49]

Mobile health apps have various functionalities. Two of the assessment tools reviewed (Yasini[42] and Loy[41]) provided dynamic assessment criteria based on the use cases and features of specific mHealth apps. In these methodologies, the relevant criteria are selected for each app according to its use cases. For example, the criterion "accuracy of the calculations" will be only used for apps that provide at least one calculation. This can lead to more accurate and efficient assessment. As an example, the criteria to assess an app that geolocates the nearest pharmacy in real time would be completely different from the criteria to assess an app created to manage a chronic disease. Dynamic assessment of apps according to use cases is not in contradiction with providing a single and comprehensive set of criteria in a data base. Therefore, the first step would

be detecting the use cases offered by the app. This could be carried out by using a classification questionnaire. Once the use cases of an app are discovered, the appropriate criteria could be selected to assess the app. If the data base is well designed with relevant decision trees, the assessment criteria could be selected even automatically according to the answers given to the classifying questionnaire. It is clear that some criteria are applied to all apps and do not need to be selected for special functionalities.

We have classified the criteria extracted from the 23 studies reviewed to provide a consolidated and inclusive set, based on literature published through December 2016. There will never be a complete and perfect mHealth apps assessment criteria, because these criteria must apply to apps that are changing in development continuously. We need decisive, accurate and reliable criteria to assess the compliance of these apps to the existing regulations and best practices. We do not have to add a jungle of criteria to the existing jungle of apps. Today, many public and private institutions (The French National Authority for Health,[61] NHS in England,[62] the European commission,[63] etc.) are working to publish guidelines concerning mobile health applications. We hypothesized that all experts of all institutions could collaborate to create a community that publishes common exhaustive guidelines in this field. An open source project

Table 1. Details of included studies

Reference Author (s)	Year	Subject category	Criteria examples	Short Description about provided tool
Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M.[33]	2015	General	Engagement Functionality Aesthetics Information subjective quality	In this study, a simple and reliable tool for classifying and assessing the quality of mHealth apps entitled Mobile Application Rating Scale (MARS), was developed. The process for development of the instrument involved a systematic review of existing guidelines/instruments for evaluation of websites and an expert panel. This tool assesses app quality on 4 dimensions. All items are rated on a 5-point scale from "1.Inadequate" to "5.Excellent."
Powell AC, Torous J, Chan S, et al.[35]	2016	Depression and Smoking Cessation	The criteria are not classified from general to specific. Some of the criteria that had greatest interrater reliability are: - Interactiveness/ Feedback - Password protection - Uploaded by health care agency - Number of consumer ratings - Explicit privacy policy - Encryption - Basis of research - Product advisory support	In this study a panel of 6 reviewers reviewed 20 apps by 22 measures. Rating scale for mental health apps presented in the Anxiety and Depression Association of America (ADAA) website including 5 measures on a five-point scale and the scale presented in the PsyberGuide website with 7 measures were 2 main sources of these measures.
Zapata BC, Fernández-alemán JL, Idri A, Toval A.[39]	2015	General	Usability criteria including: - Attractiveness - Learnability - Operability - Understandability	This paper has analyzed 22 studies that perform usability evaluations of mHealth applications. Usability model applied in this study was ISO/IEC 9126-1.
Schnall R, Rojas M, Bakken S, et al.[48]	2016	HIV prevention	Usability criteria including: - Visibility of system status - Match between system and the real world - User control and freedom - Consistency and standards - Help users recognize, diagnose, and recover from errors - Error prevention - Recognition rather than recall - Flexibility and efficiency of use - Aesthetic and minimalist design - Help and documentation	In a section of this study, the authors used a usability evaluation checklist for evaluating apps. This checklist developed by Bright et al.[53] based on Nielsen's 10 usability heuristics for user interface design.
Jin M, Kim J.[30]	2015	General	Accuracy Understandability Objectivity Consistency Suitability of design Accuracy of wording Security	In this study, after developing a version of an evaluation tool for mHealth apps from a review of previous studies, the provisional tool was modified and edited after verification by 5 experts with regard to its content validity. The answers were then collected to verify the validity and reliability of the tool.
Jeon E, Park HA, Min YH, Kim HY.[44]	2014	General	Quality of the health information including: - Authorship - Attribution - Disclosure - Currency	In this study, the quality of the health information provided by each assessed app was evaluated using a Silberg score modified by Griffiths and Christensen.
Brown W, Yen P-Y, Rojas M, Schnall R.[43]	2013	General	Usability Evaluation including: - Error prevention - Completeness - Memorability - Information needs - Flexibility /Customizability	This study tried to clarify the usefulness of the Health-IT Usability Evaluation Model for assessment the usability of mHealth apps. This model was developed to integrate multiple theories as a comprehensive framework for usability evaluation[54].

(continued)

Table 1. continued

Reference Author (s)	Year	Subject category	Criteria examples	Short Description about provided tool
			- Learnability - Performance speed - Competency - Other outcomes	
Reynoldson C, Stones C, Allsop M, et al.[52]	2014	Pain Self-Management	Product description Clinical content Interface design Ease of use was	In this study to develop criteria of each component, different resources were used. The section "ease of use" was developed using heuristics evaluation methods for mobile computing, designed by Bertini E, et al.[55] For interface design assessment an adapted version of heuristics for user interface design[56] and for clinical content assessment Clayton's program to assess and manage pain[57], SOCRATES mnemonic was used.
Martinez-Perez B, de la Torre-Diez I, Lopez-Coronado M.[49]	2015	Self-management of heart diseases (one app)	Android guidelines Style Behavior Structure Quality of Experience Content quality Security Ease of use Availability Performance Appearance Learning	This study tries to use 2 different tools for assessing the quality of a mHealth app named Heartkeeper. The first tool assesses the agreement with the Android guidelines that presented by Google and the second tool measures the users' Quality of Experience (QoE).
Van Singer M, Chatton A, Khazaal Y.[37]	2015	Panic disorder	Interactivity, Abbott's scale The health on the net code (HON) Brief DIS9CERN Silberg scale for accountability self-help model Content Quality	In this study panic disorder apps were assessed by using several tools that some of them adapted from quality evaluation studies of websites, and tools described in previous studies on the quality of mobile apps. Abbott's scale, health on the net code (HON), Brief DISCERN, Silberg scale for accountability were the main tools that were mentioned in this study.
Xiao Q, Wang Y, Sun L, Lu S, & Wu Y.[47]	2016	Cardiovascular diseases apps in China	Authorship Attribution Disclosure Currency Ease to Use Usefulness Privacy	In this study, selected apps were assessed by a quality assessment scale with 7 dimensions and 20 items derived from the Silberg scale and Technology Acceptance Model.
Taki S, Campbell K J, Russell CG, Elliott R, Laws R, & Denney-Wilson E[36].	2015	Infant Feeding Apps	Currency Author Design Navigation Content Accessibility Security Interactivity and connectivity Software Issues	In a section of this study a quality assessment tool was developed based on items from the HRWEF (Health-Related Website Evaluation Form) tool used for websites and tools used in previous studies.
Yasini M, Beranger J, Desmarais P, Perez L, & Marchand G.[42]	2016	General	Medical aspects and content validity Legal consistency Ethical issues IT security Usability	In this study, criteria were presented based on a literature review and expert experiences in 5 sections. For each section, a working group including at least 5 experts was created. These criteria were also set according to the various use cases provided by the app.

(continued)

Table 1. continued

Reference Author (s)	Year	Subject category	Criteria examples	Short Description about provided tool
Stoyanov SR, Hides L, Kavanagh DJ, & Wilson H.[38]	2016	General	Engagement Functionality Aesthetics Information subjective quality	In this study, a simple and reliable tool (uMARS) that can be used by end-users to assess the quality of mobile Health apps was created. Actually, uMARS is a user version of the original MARS[33].
McNiel P, & McArthur EC.[45]	2016	General	Currency Relevance Authority Accuracy Purpose	The authors adapted the CRAAP (Currency, Relevance, Authority, Accuracy, Purpose) Test to present it as a guideline for students to assess the credibility of mobile apps, and identify evidenced-based mHealth apps.
Loy JS, Ali EE, & Yap KYL.[41]	2016	Medication-Related Problems	Appropriateness Reliability Usability Privacy	A tool consisted of 4 sections for quality assessment of mobile apps that target Medication-Related Problems (MRPs) was developed. Features considered to be important for apps targeting MRPs included: monitoring, interaction checker, dose calculator, medication information and medication record. The quality scores become different based on the presence of each considered feature.
Anderson K, Burford O, & Emmerton L.[34]	2016	Chronic Disease	Engagement Functionality Ease of use Information management	In a section of this study, a checklist for assessment of apps dealing with chronic diseases was synthesized using peer-reviewed studies and checklists. Also face and construct validity were assessed by the authors.
Iribarren SJ, Schnall R, Stone PW, & Carballo-Dieguez A.[50]	2016	Tuberculosis Prevention and Treatment	Inform Instruct Record Display Guide Remind/alert Communicate	In this study authors developed a tool for assessment and scoring the functionalities of apps dealing with tuberculosis prevention and treatment. The tool includes 7 functionality criteria and 4 subcategories based on the Institute for Healthcare Informatics report[58].
Huckvale k, Car M, Morrison C, and Car j.[46]	2012	asthma self-management	Attribution Purpose clearly stated Privacy policy Information referencing Contact details Funding model clear Editorial/advertising policy Justification of claims	In a section of this study, a set of quality criteria that derived from HON code were used for asthma self-management app quality assessment. These criteria include 8 best-practice principles involving attribution, transparency of information and traceability.
Chen J, Cade JE, & Allman-Farinelli M.[51]	2015	Weight Loss	Accountability Scientific coverage Content accuracy Technology-enhanced features Usability	In this study, a tool developed based on a modified version of the instrument developed by Gan and Allman-Farinelli[59]. The usability of apps was assessed by the validated SUS (System Usability Scale) 10-item.
Scott K, Richards D, & Adhikari R.[32]	2015	General	Security risks Safety measures	In this study, after a literature review a tool was developed containing a set of 9 features, categorized as security risks (first 3) and safety measures (last 6). This tool also scoring the Security risk and safety.
Cruz Zapata B, Hernández Ninirola A, Idri A, Fernandez-Alemán JL, Toval A.[40]	2014	Mobile PHRs	Style Behavior Structure	In this study, a questionnaire was developed based on the known usability standards and recommendations, related literature, and the official design guidelines released for Android and iOS. This questionnaire was then validated through an online survey.

Table 1. continued

| Reference Author (s)	Year	Subject category	Criteria examples	Short Description about provided tool
Martinez-Perez, B, de la Torre-Diez I, Candelas-Plasencia S, & Lopez-Coronado M.[31]	2013	General	Quality of Experience including: - Content quality - Security - Ease of use - Availability - Performance - Appearance - Learning	In this study, firstly in order to obtain a general classification of mHealth apps a review of these apps has been done. Secondly, a Quality of Experience (QoE) measuring tool was developed in the form of a survey by contribution of psychologists. Then this tool was evaluated by using a sample of applications selected with the aid of the classification obtained.
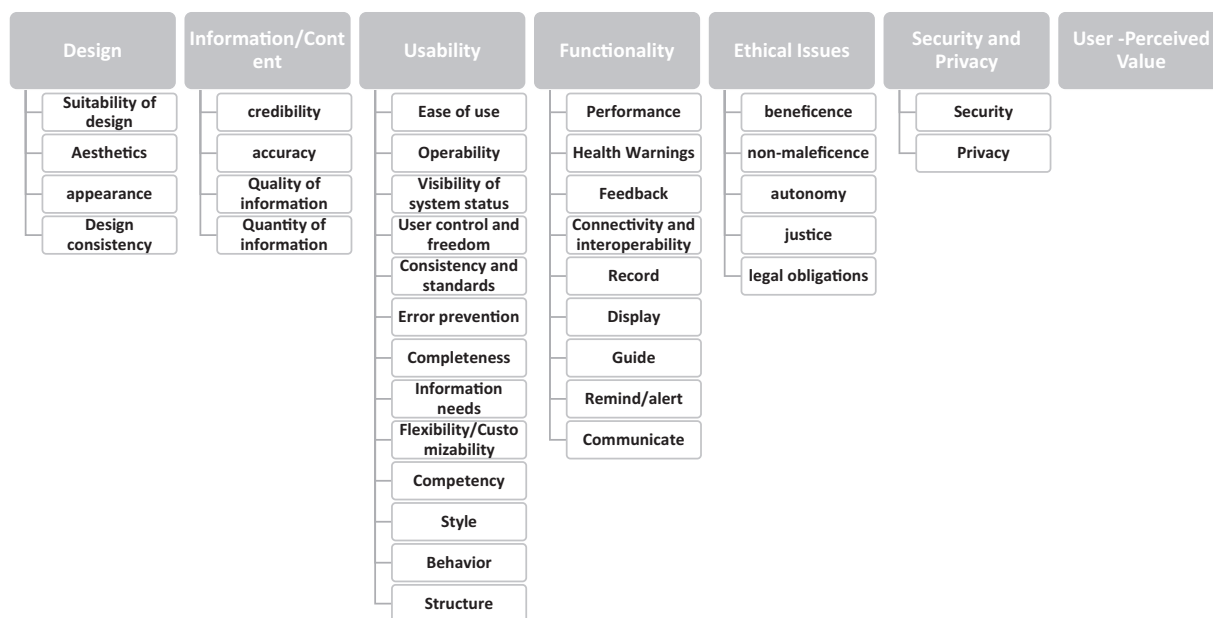
Figure 2. Outline of developed classification of mhealth apps evaluation criteria.

in this field allows us to ensure adaptability and transparency. To meet this aim, the results of this study as a framework of criteria and sub-criteria can be applied as an approved layout for further investigation. Furthermore, developing a new assessment tool for mHealth apps based on the classification presented in this study could be one of the perspectives of the utility of this research. A considerable number of papers was published on this topic after the cut-off for this review, including some reporting on the application of some existing reviewed assessment criteria (for example, uMARS[38]). The reports of experiences with existing criteria are likely to be a valuable source of input for any efforts to achieve one or more sets of standard criteria. This would be the subject of a further research that could be designed applying the criteria used in this review.

Limitation: We excluded non-English articles, and we did not take into account existing guidelines and standards about mHealth apps assessment that were not indexed in our search resources (EMBase, Medline, Web of sciences and Scopus). For example, some European states have published related guidelines in this field.[61] Another limitation was the general lack of good definitions among assessment criteria that could lead to some misinterpretations of the expert panel, for the construction of our consolidated set of criteria.

CONCLUSION

In conclusion, in this study 7 main categories of criteria and 37 subclasses were presented for health-related app assessment. There is a huge heterogeneity in assessment criteria for mHealth apps in different studies. This can be either due to the various assessment approaches used by researchers or different definitions for each criterion. Although in some cases providing a unique and scientific definition of a criterion or defining its place in an appropriate hierarchy of criteria may be very difficult, it seems necessary to reach a consensus by experts of this field about related concepts. It is also necessary to provide precise and mutually exclusive definitions for assessment criteria. Other findings indicate that, depending on their use cases, different kinds of mHealth apps may need different assessment criteria. Addressing these points may lead to improvement of existing tools and development of better and more standard mHealth app assessment tools.

FUNDING

CONTRIBUTORS

All authors contributed to the study conception and design. The search procedure (screening the papers and full text assessment) was carried out by RN and SRNK with arbitration and confirmation by MY. Drafting of the manuscript was carried out by RN, SRNK and MG. All authors contributed the analytic strategy to achieve the final classification of assessment criteria. MY critically revised the manuscript and provided insights on the review discussion. GM and MY approved the version to be published.

COMPETING INTERESTS

None.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

REFERENCES

1. Barton AJ. The regulation of mobile health applications. *BMC Med* 2012; 10 (1): 46.
2. Morse SS, Murugiah MK, Soh YC, Wong TW, Ming LC. Mobile health applications for pediatric care: review and comparison. *Ther Innov Regul Sci* 2018; 52 (3): 383–91.
3. Årsand E, Frøisland DH, Skrøvseth SO, *et al*. Mobile health applications to assist patients with diabetes: lessons learned and design implications. *J Diabetes Sci Technol* 2012; 6 (5): 1197–206.
4. Tripp N, Hainey K, Liu A, *et al*. An emerging model of maternity care: smartphone, midwife, doctor? *Women Birth* 2014; 27 (1): 64–7.
5. Baldwin JL, Singh H, Sittig DF, Giardina TD. Patient portals and health apps: pitfalls, promises, and what one might learn from the other. *Healthc* 2017; 5 (3) 81–5.
6. Boyce B. Nutrition apps: opportunities to guide patients and grow your career. *J Acad Nutr Diet* 2014; 114 (1): 13–4.
7. Goyal S, Cafazzo JA. Mobile phone health apps for diabetes management: current evidence and future developments. *QJM* 2013; 106 (12): 1067–9.
8. Singh K, Drouin K, Newmark LP, *et al*. Developing a framework for evaluating the patient engagement, quality, and safety of mobile health applications. *Issue Brief (Commonw Fund)* 2016; 5 (1): 11.
9. Hoyt RE, Yoshihashi AK. *Health Informatics: Practical Guide for Healthcare and Information Technology Professionals*. Lulu; 2014.
10. García-Gómez JM, de la Torre-Díez I, Vicente J, Robles M, López-Coronado M, Rodrigues JJ. Analysis of mobile health applications for a broad spectrum of consumers: a user experience approach. *Health Informatics J* 2014; 20 (1): 74–84.
11. Silva BMC, Rodrigues JJPC, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. *J Biomed Inform* 2015; 56: 265–72.
12. Lee Y, Moon M. Utilization and content evaluation of mobile applications for pregnancy, birth, and child care. *Healthc Inform Res* 2016; 22 (2): 73–80.
13. Agu E, Pedersen P, Strong D, *et al*. The smartphone as a medical device: assessing enablers, benefits and challenges. In: 10th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2013. IEEE; 2013: 76–80.
14. Lee Ventola C. Mobile devices and apps for health care professionals: uses and benefits. *P T* 2014; 39 (5): 356–64.
15. Moon BC, Chang H. Technology acceptance and adoption of innovative smartphone uses among hospital employees. *Healthc Inform Res* 2014; 20 (4): 304–12.
16. Ozdalga E, Ozdalga A, Ahuja N. The smartphone in medicine: a review of current and potential use among physicians and students. *J Med Internet Res* 2012; 14 (5): e128.
17. Whitlow ML, Drake E, Tullmann D, Hoke G, Barth D. Bringing technology to the bedside: using smartphones to improve interprofessional communication. *Comput Inform Nurs* 2014; 32 (7): 305–11.
18. Wu RC, Tzanetos K, Morra D, Quan S, Lo V, Wong BM. Educational impact of using smartphones for clinical communication on general medicine: more global, less local. *J Hosp Med* 2013; 8 (7): 365–72.
19. Yasini M, Marchand G. Toward a use case based classification of mobile health applications. *Stud Health Technol Inform* 2015; 210: 175–9.
20. Research2guidance. mHealth economics 2017—current status and future trends in mobile health. 2017. https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ Accessed December 24, 2017.
21. Domnich A, Arata L, Amicizia D, *et al*. 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* 2016; 16 (1)83.
22. Slade M, Oades L, Jarden A. *Wellbeing, Recovery and Mental Health*. Cambridge: Cambridge University Press; 2017.
23. Albrecht UV, von Jan U. mHealth apps and their risks—taking stock. *Stud Health Technol Inform* 2016; 226: 225–8.
24. Huckvale K, Morrison C, Ouyang J, Ghaghda A, Car J. The evolution of mobile apps for asthma: an updated systematic assessment of content and tools. *BMC Med* 2015; 13 (1): 58.
25. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Behav Med Pract Policy Res* 2014; 4 (4): 363–71.
26. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. *Telemed J E Health* 2015; 21 (2): 97–104.
27. Zhang C, Zhang X, Halstead-Nussloch R. Assessment metrics, challenges and strategies for mobile health apps. *Issues Inform Syst* 2014; 15 (2): 59.
28. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. *J Med Internet Res* 2017; 19 (4): e125.
29. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; 151 (4): 264–9.
30. Jin M, Kim J. Development and evaluation of an evaluation tool for healthcare smartphone applications. *Telemed J E Health* 2015; 21 (10): 831–7.
31. Martínez-Pérez B, de la Torre-Díez I, Candelas-Plasencia S, López-Coronado M. Development and evaluation of tools for measuring the quality of experience (QoE) in mHealth applications. *J Med Syst* 2013; 37 (5): 9976.
32. Scott K, Richards D, Adhikari R. A review and comparative analysis of security risks and safety measures of mobile health apps. *Aust J Inform Syst* 2015; 9.
33. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth Uhealth* 2015; 3 (1): e27.
34. Anderson K, Burford O, Emmerton L. App chronic disease checklist: protocol to evaluate mobile apps for chronic disease self-management. *JMIR Res Protoc* 2016; 5 (4): e204.
35. Powell AC, Torous J, Chan S, *et al*. Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. *JMIR Mhealth Uhealth* 2016; 4 (1): e15.
36. Taki S, Campbell KJ, Russell CG, Elliott R, Laws R, Denney-Wilson E. Infant feeding websites and apps: a systematic assessment of quality and content. *Interact J Med Res* 2015; 4 (3): e18.
37. Van Singer M, Chatton A, Khazaal Y. Quality of smartphone apps related to panic disorder. *Front Psychiatry* 2015; 6: 96.

38. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). *JMIR MHealth Uhealth* 2016; 4 (2): e72.

39. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. *J Med Syst* 2015; 39 (2): 1–19.

40. Cruz Zapata B, Hernández Niñirola A, Idri A, Fernández-Alemán JL, Toval A. Mobile PHRs compliance with android and iOS usability guidelines. *J Med Syst* 2014; 38 (8): 81.

41. Loy JS, Ali EE, Yap KYL. Quality assessment of medical apps that target medication-related problems. *J Manag Care Spec Pharm* 2016; 22 (10): 1124–40.

42. Yasini M, Beranger J, Desmarais P, Perez L, Marchand G. mHealth quality: a process to seal the qualified mobile health apps. *Stud Health Technol Inform* 2016; 228: 205–9.

43. Brown W, Yen P-Y, Rojas M, Schnall R. Assessment of the Health IT usability evaluation model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *J Biomed Inform* 2013; 46 (6): 1080.

44. Jeon E, Park HA, Min YH, Kim HY. Analysis of the information quality of Korean obesity-management smartphone applications. *Healthc Inform Res* 2014; 20 (1): 23–9.

45. McNiel P, McArthur EC. Evaluating health mobile apps: information literacy in undergraduate and graduate nursing courses. *J Nurs Educ* 2016; 55 (8): 480.

46. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. *BMC Med* 2012; 10 (1): 144.

47. Xiao Q, Wang Y, Sun L, Lu S, Wu Y. Current status and quality assessment of cardiovascular diseases related smartphone apps in China. *Stud Health Technol Inform* 2016; 225: 1030–1.

48. Schnall R, Rojas M, Bakken S, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J Biomed Inform* 2016; 60: 243–51.

49. Martinez-Perez B, de la Torre-Diez I, Lopez-Coronado M. Experiences and results of applying tools for assessing the quality of a mHealth app named Heartkeeper. *J Med Syst* 2015; 39 (11)

50. Iribarren SJ, Schnall R, Stone PW, Carballo-Dieguez A. Smartphone applications to support tuberculosis prevention and treatment: review and evaluation. *JMIR MHealth Uhealth* 2016; 4 (2): e25.

51. Chen J, Cade JE, Allman-Farinelli M. The most popular smartphone apps for weight loss: a quality assessment. *JMIR MHealth Uhealth* 2015; 3 (4): e104.

52. Reynoldson C, Stones C, Allsop M, et al. Assessing the quality and usability of smartphone apps for pain self-management. *Pain Med* 2014; 15 (6): 898–909.

53. Bright TJ, Bakken S, Johnson SB. Heuristic evaluation of eNote: an electronic notes system. *AMIA Annu Symp Proc* 2006; 2006: 864.

54. Yen P-Y. *Health Information Technology Usability Evaluation: Methods, Models, and Measures*. New York: Columbia University; 2010.

55. Bertini E, Catarci T, Dix A, Gabrielli S, Kimani S, Santucci G. Appropriating heuristic evaluation methods for mobile computing. *Int J Mob Hum Comput Interact* 2008; 1 (1): 20–41.

56. Molich R, Nielsen J. Improving a human-computer dialogue. *Commun ACM* 1990; 33 (3): 338–48.

57. Simon C. Pain control in palliative care. *InnovAiT* 2008; 1 (4): 257–66.

58. Aitken M. *Patient Apps for Improved Healthcare: from Novelty to Mainstream*. Parsippany, NJ: IMS Institute for Healthcare Informatics; 2013.

59. Gan KO, Allman-Farinelli M. A scientific audit of smartphone applications for the management of obesity. *Aust N Z J Public Health* 2011; 35 (3): 293–4.

60. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: a systematic review of prevalent and innovative methods. *Am J Prev Med* 2016; 51 (6): 1051–9.

61. French National Authority for Health. Good practice guidelines on health apps and smart devices (mobile health or mhealth). 2016. https://www.has-sante.fr/portail/jcms/c_2681915/en/good-practice-guidelines-on-health-apps-and-smart-devices-mobile-health-or-mhealth Accessed May 20, 2017.

62. NHS Innovations South East. App development: an NHS guide for developing mobile healthcare applications. 2014. http://www.innovationssoutheast.nhs.uk/files/4214/0075/4193/98533_NHS_INN_AppDevRoad.pdf Accessed December 19, 2016.

63. European Commission. First draft of guidelines: EU guidelines on assessment of the reliability of mobile health applications 2016. www.health-community.be/sites/default/files/u16/2nddraftguidelines.pdf Accessed December 19, 2016.