Quality assessment of mHealth apps: a scoping review

Godwin Denk Giebel[1†], Christian Speckemeier[1*†],
Nils Frederik Schrader[1], Carina Abels[1], Felix Plescher[1],
Vivienne Hillerich[1], Desiree Wiedemann[1], Kirstin Börchers[2],
Jürgen Wasem[1], Nikola Blase[1‡] and Silke Neusser[1‡]

[1]Institute for Healthcare Management and Research, University of Duisburg-Essen, Essen, Germany,
[2]QM BÖRCHERS CONSULTING+, Herne, Germany

Introduction: The number of mHealth apps has increased rapidly during recent years. Literature suggests a number of problems and barriers to the adoption of mHealth apps, including issues such as validity, usability, as well as data privacy and security. Continuous quality assessment and assurance systems might help to overcome these barriers. Aim of this scoping review was to collate literature on quality assessment tools and quality assurance systems for mHealth apps, compile the components of the tools, and derive overarching quality dimensions, which are potentially relevant for the continuous quality assessment of mHealth apps.

Methods: Literature searches were performed in Medline, EMBASE and PsycInfo. Articles in English or German language were included if they contained information on development, application, or validation of generic concepts of quality assessment or quality assurance of mHealth apps. Screening and extraction were carried out by two researchers independently. Identified quality criteria and aspects were extracted and clustered into quality dimensions.

Results: A total of 70 publications met inclusion criteria. Included publications contain information on five quality assurance systems and further 24 quality assessment tools for mHealth apps. Of these 29 systems/tools, 8 were developed for the assessment of mHealth apps for specific diseases, 16 for assessing mHealth apps for all fields of health and another five are not restricted to health apps. Identified quality criteria and aspects were extracted and grouped into a total of 14 quality dimensions, namely "information and transparency", "validity and (added) value", "(medical) safety", "interoperability and compatibility", "actuality", "engagement", "data privacy and data security", "usability and design", "technology", "organizational aspects", "social aspects", "legal aspects", "equity and equality", and "cost(-effectiveness)".

Discussion: This scoping review provides a broad overview of existing quality assessment and assurance systems. Many of the tools included cover only a few dimensions and aspects and therefore do not allow for a comprehensive quality assessment or quality assurance. Our findings can contribute to the development of continuous quality assessment and assurance systems for mHealth apps.

Systematic Review Registration: https://www.researchprotocols.org/2022/7/e36974/, International Registered Report Identifier, IRRID (DERR1-10.2196/36974).

1 Introduction

The number of mobile phone users in 2023 was estimated at 7.3 billion worldwide, representing over 90% of the world's population (1, 2). The intensive use of mobile devices has affected many industries, including the proliferation of mobile healthcare (mHealth) apps (3). While a universally accepted definition is lacking (4), the term mHealth is broadly defined as using "portable devices with the capability to create, store, retrieve, and transmit data in real time between end users for the purpose of improving patient safety and quality of care" (5). As an integral part of eHealth, mHealth apps aim to improve access to evidence-based information and engage patients directly in treatments by enabling providers (e.g., doctors, healthcare facilities) to connect with patients (6, 7). As such, mHealth apps have the potential to improve healthcare through accessible, effective and cost-effective interventions (8). In times of demographic change and healthcare workforce shortages, high-quality apps might contribute to sustainable healthcare (9). Especially with the rise of chronic diseases, mHealth apps can be an opportunity for prevention and improved treatment, as these diseases require constant self-care and monitoring (10). Despite this potential, literature suggests a scarcity of high-quality mHealth apps (11). In line with this, a scoping review identified several problems and barriers to the utilization of mHealth apps, including issues related to validity, usability, as well as data privacy and security, among others (12). Particularly with the widespread use of mHealth apps, it is important to avoid quality issues such as misinformation, which can limit effectiveness or potentially harm the user. As in other areas of health care, high standards are needed for evidence-based and high-quality mHealth apps (12). Appropriate quality assessment and assurance is therefore needed both during the development and ongoing use of mHealth apps.

According to a World Health Organization's definition, quality of care is "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes" (13). While in general, health care quality is a multidimensional construct (14), quality dimensions in mHealth differ from those in other existing healthcare services (15). With its fast-track procedure, Germany was the first country in the world to create a system that makes selected, tested mHealth apps [called "Digital Health Applications" (DiGA)] an integral part of healthcare (16). The Federal Institute for Drugs and Medical Devices (BfArM) has set certain requirements the app must meet in order to be included in the so called "DiGA directory". These apps have to demonstrate scientifically proven evidence of a benefit, either in the form of medical benefits or patient-relevant structure and process improvements for the patient (16). Furthermore, they must meet requirements for product safety and functionality, privacy and information security, interoperability, robustness, consumer protection, usability, provider support, medical content quality and patient safety. Once listed in the directory, patients can request these mHealth apps from their health insurance company, or the apps can be prescribed directly (16).

Currently there is a need for further adjustments to the fast-track procedure on the part of providers, health insurers and manufacturers (9). For example, the National Association of Statutory Health Insurance Funds (GKV-Spitzenverband) requires, among other things, that quality specifications must be met for user-friendly and target group-oriented design, data protection and data security (17).

Quality assessment tools: In addition to country-specific approaches, a number of simple assessment tools have been developed, such as the Mobile App Rating Scale (MARS) (18), ENLIGHT (19) or the System Usability Scale (SUS) (20). These approaches (in the following called *quality assessment tools*) typically assess the quality of apps with a number of items and provide the user with a score. For example, as one of the most widely used evaluation tools, MARS was developed on the basis of a literature review of existing criteria for evaluating the quality of apps and subsequent categorization by a panel of experts. The resulting multidimensional rating scale covers the areas of engagement, functionality, aesthetics, information and subjective quality of apps. Resulting scores are intended to be used by researchers, guide app developers, or to inform health professionals and policymakers (18, 19).

Quality assurance systems: In addition, approaches have been developed (hereinafter referred to as *quality assurance systems*) which go beyond traditional scoring instruments, e.g., by providing a framework for assessing the mHealth apps along their product lifecycle. For example, Sadegh et al. (21) propose an mHealth evaluation framework through three different stages of the app's lifecycle. Similarly, Mathews et al. (22) detail a framework assessing technical, clinical, usability, and cost aspects pre- and post-market entry. To date, there is no overview in the literature that differentiates between quality assessment tools and quality assurance systems.

Therefore, and in view of the situation in Germany described above, this work pursues two objectives: (1) to collate literature on quality assessment tools and quality assurance systems for mHealth apps, compile the components of the tools, and group them into overarching quality dimensions, which are potentially relevant for the continuous quality assessment of mHealth apps; (2) to identify and characterize quality assurance systems with a view to continuous quality assurance.

Relevant information can be extracted from publications in which the tools are developed or validated. Studies in which the tools are used for the evaluation of apps are also potentially relevant, as they provide evidence that the respective tools have been applied for the assessment of an mHealth app by researchers. The method of scoping review was found feasible, as no single precise question regarding feasibility, appropriateness, meaningfulness or effectiveness had to be answered (23). While specific questions of effectiveness are traditionally answered by collating quantitative literature in a systematic review, scoping reviews are used to map literature and address a broader research question (e.g., identify gaps in research, clarify concepts, or report on types of evidence that inform clinical practice) (24).

This work is part of a larger research project (QuaSiApps—Ongoing Quality Assurance of Health Apps Used in Statutory Health Insurance Care), which is funded by the Innovation Fund of the Federal Joint Committee and aims to create a concept for continuous quality assurance of mHealth apps.

2 Methods

A scoping review was conducted to answer the following questions: Which quality assessment tools and quality assurance systems have been developed and/or used in the field of mHealth apps? Which items do they consist of? Which quality dimensions can be derived from the quality assessment tools and quality assurance systems? To answer these questions, we followed the key phases outlined by Levac et al. (25), including identifying relevant studies, study selection, charting the data, and collating, summarizing, and reporting the results. Reporting followed the PRISMA extension for scoping reviews (26). A review protocol was written and published prior to screening (27). The protocol contains detailed information on the databases searched, the search terms used, and the inclusion and exclusion criteria applied during the screening process.

2.1 Literature search

The electronic indexed databases Medline, EMBASE and PsycInfo were searched for primary literature on the topic. Studies containing description of a literature review were included, if the review served to develop the items of the assessment tool presented. However, the focus had to be on the development and description of a specific tool. Search strategies were developed through discussion (GG, NS, CS) and with aid of the working group leader (SN). The strategies were pilot tested and refined. The search strategies comprise of keywords and synonyms for assessment tools and mHealth. All bibliographic searches were adapted to the databases' requirements. Full search strategies and number of hits per keyword can be found in the review protocol. Searches were executed on July 26th, 2021. Reference lists of included articles were screened for further eligible literature. Further information such as the search string can be found in the corresponding research protocol (27).

2.2 Inclusion and exclusion criteria

Studies were eligible for inclusion if they fulfilled the following criteria: (1) included either development, or description, or further information on disease-independent concepts of quality assessment or quality assurance of mHealth apps, (2) were in English or German language, and (3) were published between January 1st, 2016 and July 26th, 2021. This means that studies were included if quality assessment tools and quality assurance systems were applied (application studies), developed (development studies) or validated (validation studies). In order to incorporate approaches

currently in use, the quality assessment tools and quality assurance systems used in application and validation studies were identified and included, even if they were published before January 1st, 2016. For application studies, the investigated mHealth apps had to be used by patients in outpatient treatment and needed to have more functions than improvement of adherence, text-messaging, reminder or screening for primary prevention or (video) consultation or disease education or reading out and controlling of devices.

Applied exclusion criteria were: (1) articles that did not include information on quality assessment tools or quality assurance systems, (2) the investigated quality assessment or quality assurance system was not disease-independent, (3) the assessed mHealth app had not more functions than the following: improvement of adherence, text-messaging, reminder or screening for primary prevention or (video) consultation or disease education or reading out and controlling of devices, (4) The mHealth app evaluated was not primarily for patient use, (5) the assessed mHealth app is not used in outpatient treatment, (6) articles that included only research protocols, conference abstracts, letters to the editor, or expression of opinions. Apart from the publication date, the inclusion and exclusion criteria were all set manually and not using the filter function of the databases. Further information on the inclusion and exclusion criteria, including the search timeframe, can be found in the review protocol (27).

2.3 Selection of relevant studies

Identified results were loaded into the EndNote reference management program (Clarivate Analytics, Philadelphia, US; version X9). Duplicates were removed automatically and manually during the screening process. All unique references were screened in terms of their potential relevance based on title and abstract. Documents considered potentially relevant were reviewed in full-text and retained if the study met inclusion criteria. Two researchers (GG, NS) performed all screening steps independently. Any disagreements were resolved by consulting a senior researcher (SN).

2.4 Extraction and analysis of data

Included studies were extracted in tables by two persons independently (GG, CS). Relevant data of included articles was marked and extracted using MAXQDA 2022 (Verbi Software GmbH, Berlin). In a first step, articles were categorized into application studies, validation studies, and development studies and were then extracted into pre-specified tables. The extraction table for application studies comprised author(s), year, country, the used quality assessment tool(s), investigated disease(s) or the field(s) of application, the number of investigated apps, the study type and the source of the tool used in the application study. Data extraction from validation studies included author(s), year, country, the validate quality assessment tool(s), investigated disease(s)/field(s) of application and the origin of the validated

tool. The extraction table for development studies included author (s), year, country, quality assessment tool, disease(s)/field(s) of application, the quality dimensions described and named by the author of the respective studies and the attribution to the quality dimensions developed in this scoping review.

Identified items were extracted and grouped into clusters in Microsoft Excel by one researcher (GG) and quality-checked by two researchers (FP, CA). Results were compared and discussed in case of disagreement. If necessary, a senior researcher was involved (SN). In case a criterion or aspect did not match into an existing cluster, a new cluster was created. Based on the information from the literature analyzed, the clusters were labeled. The labeled clusters were described and constituted the quality dimensions. The results were summarized, systemized and presented in tables.

3 Results

The selection process is shown in Figure 1. A total of 2,871 articles were identified in the three databases. Of these, 2,235 articles remained after duplicate removal and were screened according to title and abstract. One hundred and twenty-four articles were included in full-text screening and subsequently, 59 studies met inclusion criteria. See Supplementary Appendix A for a table of studies excluded in full-text screening. A further 11 articles were identified via citation searching. This refers to the

studies in which the tools mentioned in the application studies were developed. In total, 70 articles were included.

In 15 of the included articles, a quality assurance system or a quality assessment tool was developed (10, 21, 28–40). Five of the included articles were validation studies (8, 41–44). In addition to development and validation studies, a number of studies (n = 39) were identified in which quality assessment tools were applied. Of these, 19 studies employed the Mobile App Rating Scale (MARS) which was developed by Stoyanov et al. (18) or a modified version of it (33, 38, 45–61), and one study (62) used the user version of the MARS (uMARS) proposed by Stoyanov et al. (63). Another ten studies (39, 64–72) employed the System Usability Scale (SUS) developed by Brooke et al. (20). Two studies (73–75) used the (modified) Silberg scale. Further seven studies were identified in which additional quality assessment tools and quality assurance systems were applied (76–82). In total, 14 quality assurance systems and quality assessment tools were found in application or validation studies (18–20, 22, 63, 73, 83–90). Of note, the articles by Liu et al. (33), Tan et al. (38) and Wood et al. (39) report on both development and application and were therefore classified in both categories. An overview of all included articles is given in Supplementary Appendix B. In the 74 included articles, a total of 29 distinct approaches to quality assurance or quality assessment were identified. Five of these have been identified as quality assurance systems (21, 22, 30, 35, 90), while the remaining 24 tools are considered quality assessment tools. Figure 2 gives an overview of the different types of studies included.

FIGURE 1
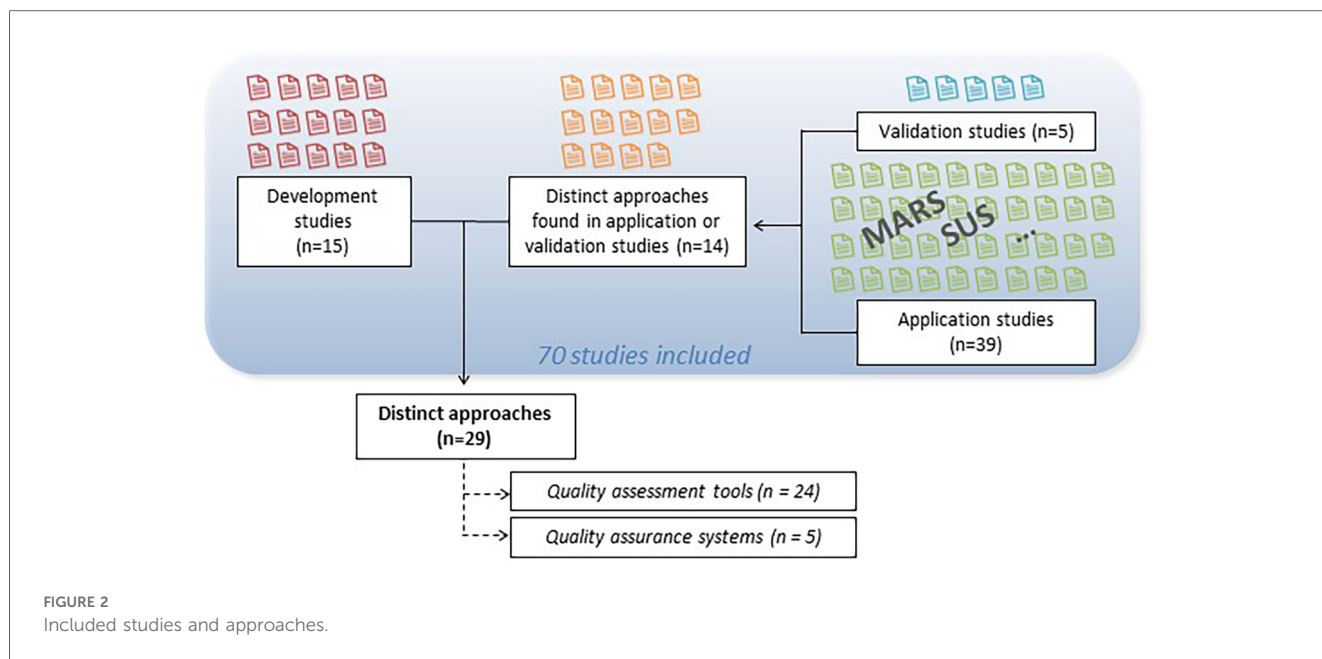Flow diagram depicting the selection of sources of evidence.
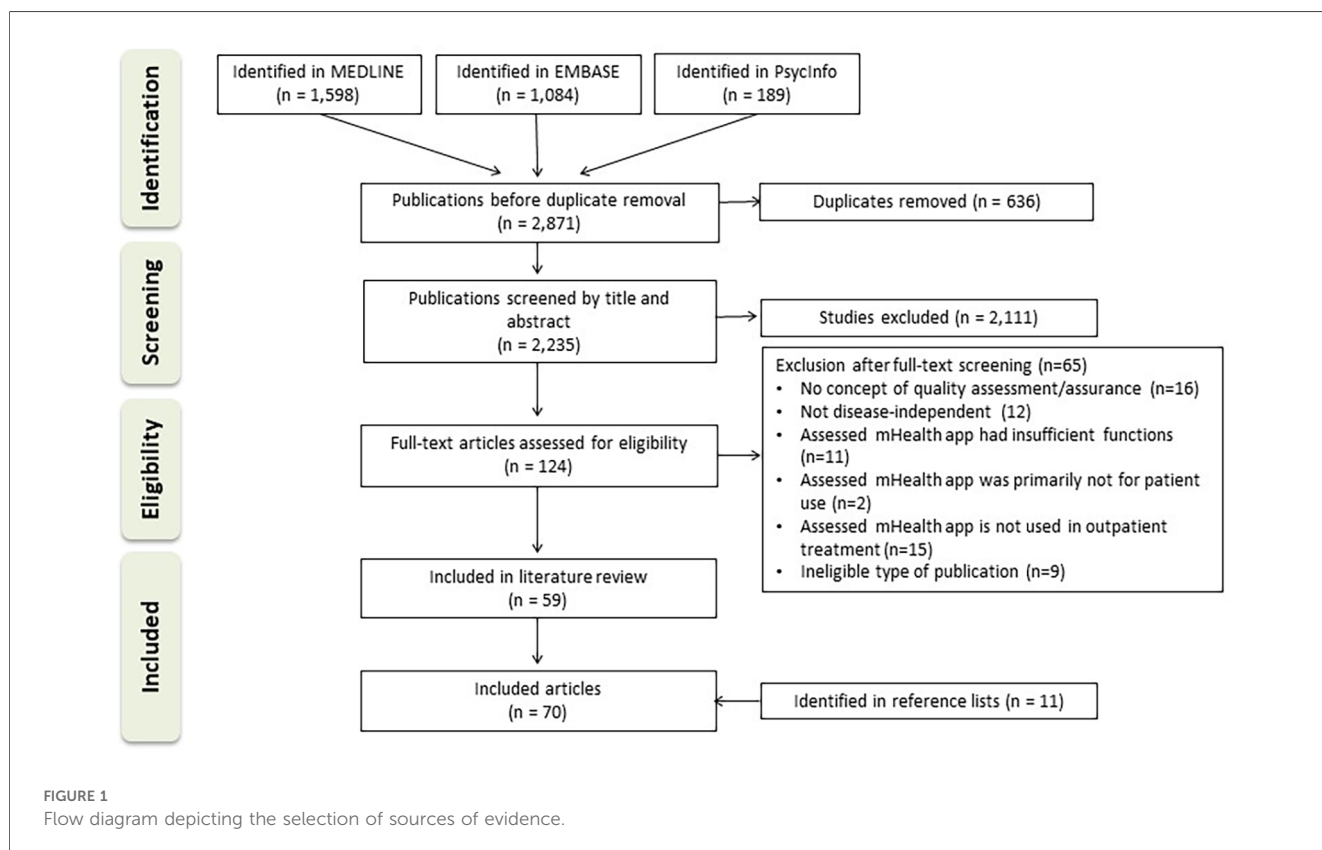
FIGURE 2
Included studies and approaches.

3.1 Characteristics of included quality assessment tools

An overview of the identified 24 quality assessment tools is presented in Table 1. Of these, 8 were developed for the

assessment of mHealth apps for specific diseases or disease areas and 11 for assessing mHealth apps for all fields of health. Another five are not restricted to health apps, but have been included as they have been used for the assessment of health apps in application studies. Included articles dealing with quality

TABLE 1 Characteristics of the included quality assessment tools.

Author (year)	Country	Tool name	Field of application
Baumel et al. (2017) (19)	US	ENLIGHT	All fields of health
Berry et al. (2018) (28)	UK	Mobile Agnew Relationship Measure (mARM) Questionnaire	Mental health
Brooke et al. (1996) (20)	UK	System Usability Scale (SUS)	Not restricted to health
Brown et al. (2013) (29)	US	Health-ITUEM	All fields of health
Doak et al. (1996) (83)	US	Suitability Assessment of Materials (SAM)	All fields of health
Glattacker et al. (2020) (31)	Germany	Usability questionnaire	Allergic Rhinitis
Huang et al. (2020) (32)	Singapore	App-HONcode	Medication Management in Diabetes
Huckvale et al. (2015) (84)	UK	Untitled	All fields of health
Jusob et al. (2022) (10)	UK	Untitled	Chronic diseases
Lewis et al. (1995) (85)	US	After-Scenario Questionnaire (ASQ), Post-Study System Usability Questionnaire (PSSUQ), Computer System Usability Questionnaire (CSUQ)	Not restricted to health
Liu et al. (2021) (33)	China	Untitled	Traditional Chinese Medicine and Modern Medicine
Llorens-Vernet and Miro (2020) (40)	Spain	Mobile App Development and Assessment Guide (MAG)	All fields of health
Minge and Riedel (2013) (34)	Germany	meCUE	Not restricted to health
O'Rourke et al. (2020) (36)	Austria	App Quality Assessment Tool for Health-Related Apps (AQUA)	All fields of health
Pifarre et al. (2017) (37)	Spain	Untitled	Tobacco-quitting
Reichheld (2004) (86)	US	Net Promoter Score (NPS)	Not restricted to health
Ryu and Smith-Jackson (2006) (87)	US	Mobile Phone Usability Questionnaire (MPUQ)	Not restricted to health
Schnall et al. (2018) (88)	US	Health Information Technology Usability Evaluation Scale (Health-ITUES)	All fields of health
Shoemaker et al. (2014) (89)	US	Patient Education Materials Assessment Tool (PEMAT)	All fields of health
Silberg et al. (1997) (73)	Sweden	Silberg Scale	All fields of health
Stoyanov et al. (2015) (18)	Australia	Mobile App Rating Scale (MARS)	All fields of health
Stoyanov et al. (2016) (63)	Australia	User Version of the Mobile Application Rating Scale (uMARS)	All fields of health
Tan et al. (2020) (38)	Australia	Untitled	Allergic Rhinitis and/or asthma
Wood et al. (2018) (37)	Australia	Untitled	Cystic fibrosis

assessment tools predominantly stem from the USA ($n = 8$), Australia ($n = 4$), and the UK ($n = 4$). The identified quality assessment tools were developed in different ways, e.g., by adapting existing measures, based on findings from literature and guideline review, by conducting focus groups or by mixed-methods approaches. The quality assessment tools were developed for utilization developers, academics, healthcare providers, government officials and users.

None of the 24 articles includes a definition of a concept for quality. Of note, Brooke et al. (20) define the concept of usability in the context of the SUS. The tools are diverse with regard to their extent. Some tools consider single aspects, such as engagement (28, 86). For example, the net promoter score (86), which was used by de Batlle et al. (64), consists of only one question. In contrast, other tools cover a wider range of aspects (19, 40).

3.2 Characteristics of included quality assurance systems

The identified approaches were assigned to quality assurance systems if they assessed the apps over time. An overview of the included quality assurance systems is presented in Table 2.

The five included quality assurance systems stem from the US ($n = 2$), Australia ($n = 1$), Iran ($n = 1$), and France ($n = 1$). Similar to the quality assessment tools identified, none of the five articles includes a definition of a concept for quality. The quality assurance systems were developed to be used by developers, health practitioners, government officials, and users. Camacho et al. (30) tailored an existing implementation framework and developed a process to assist stakeholders, clinicians, and users with the implementation of mobile health technology. The Technology Evaluation and Assessment Criteria for Health apps (TEACH-apps) consists of the four parts (1) preconditions, (2) preimplementation, (3) implementation, and (4) maintenance and evolution. The authors recommend to repeat the process at least biannually, in order to adapt for changing consumer preferences over time (30). Mathews et al. (22) propose a digital health scorecard consisting of four domains (technical, clinical, usability, cost), which aims to serve as framework guiding the evolution and successful delivery of validated mHealth apps over the product's lifecycle. Moshi et al. (35) have developed criteria for evaluation of mHealth apps within health technology assessment (HTA) frameworks. The multidimensional module also contains items allowing for post-market surveillance. Sadegh et al. (21) have conducted an mHealth evaluation framework throughout the lifecycle in three stages, namely (1) service requirement analysis, (2) service development, and (3) service delivery. Finally, Yasini et al. (90) have developed a multidimensional scale for quality assessment of mHealth apps.

3.3 Derived quality dimensions

In total, 584 items were extracted from the identified quality assessment tools and quality assurance systems and were categorized into clusters, respectively the quality dimensions. The number of quality dimensions derived from each of the 29 articles ranged from one to 13 dimensions, with an average of 4.9 dimensions. These were grouped to a total of 14 distinct quality dimensions. Figure 3 gives an overview of the identified quality dimensions and quality aspects.

Items pertaining to the quality dimension "validity and (added) value" were contained in 21 of the included quality assessment tools and quality assurance systems. Items addressing the clear, complete and accurate presentation of relevant and useful content based on evidence-based information were included here. In addition, items concerned with the provision of information about the (scientific) sources used, the involvement of experts in the development and evaluation process and the patient-specific benefits were considered relevant for this quality dimension. Twenty-one articles contained items which were grouped to "usability and design". Usability provides information on how difficult / complex it is to operate and use the app. Usability can be indicated by ease of use. Both direct and long-term use should be taken into account. The design includes the presentation and associated clarity. The application itself, but also the results provided, should be clear and concise. Integrated functions should always be fit for purpose. The usability should be tested by usage tests before publication.

Eighteen of the 32 articles included information which was grouped to the quality dimension "engagement". Engagement describes the user's involvement and can be indicated by the extent of use or the intention to use the app long term. It can be strengthened by calls to action, the setting of goals and human attributes such as friendliness, trust, and acceptance. Users can be motivated by interactions, personalization, interesting content and resulting fun during use. In contrast to the intention to use, the subjective benefit is not part of this dimension but belongs to Validity & (Added) Value.

Fifteen of the included articles describing quality assessment tools and quality assurance systems contained dimensions which were sorted into "information and transparency". The dimension "data privacy and data security" was contained in 11 articles. Further dimensions are "technology" ($n = 9$), "equity and equality" ($n = 8$), "interoperability and compatibility" ($n = 7$), "(medical) safety" ($n = 7$), "actuality" ($n = 7$), "legal aspects" ($n = 5$), "(cost-)effectiveness" ($n = 5$), "social aspects" ($n = 4$), and "organizational aspects" ($n = 4$). An overview of the quality dimensions derived from the included studies can be found in Supplementary Appendix C. The full descriptions of these dimensions, which

TABLE 2 Characteristics of included quality assurance systems.

Author (year)	Country	Tool name	Field of application
Camacho et al. (2020) (30)	US	Technology Evaluation and Assessment Criteria for Health Apps (TEACH-Apps)	All fields of health
Mathews et al. (2019) (22)	US	Digital Health Scorecard	All fields of health
Moshi et al. (2020) (35)	Australia	Health technology assessment module	All fields of health
Sadegh et al. (2018) (21)	Iran	Untitled	All fields of health
Yasini et al. (2016) (90)	France	Multidimensional assessment program	All fields of health

FIGURE 3
Quality dimensions with corresponding criteria.
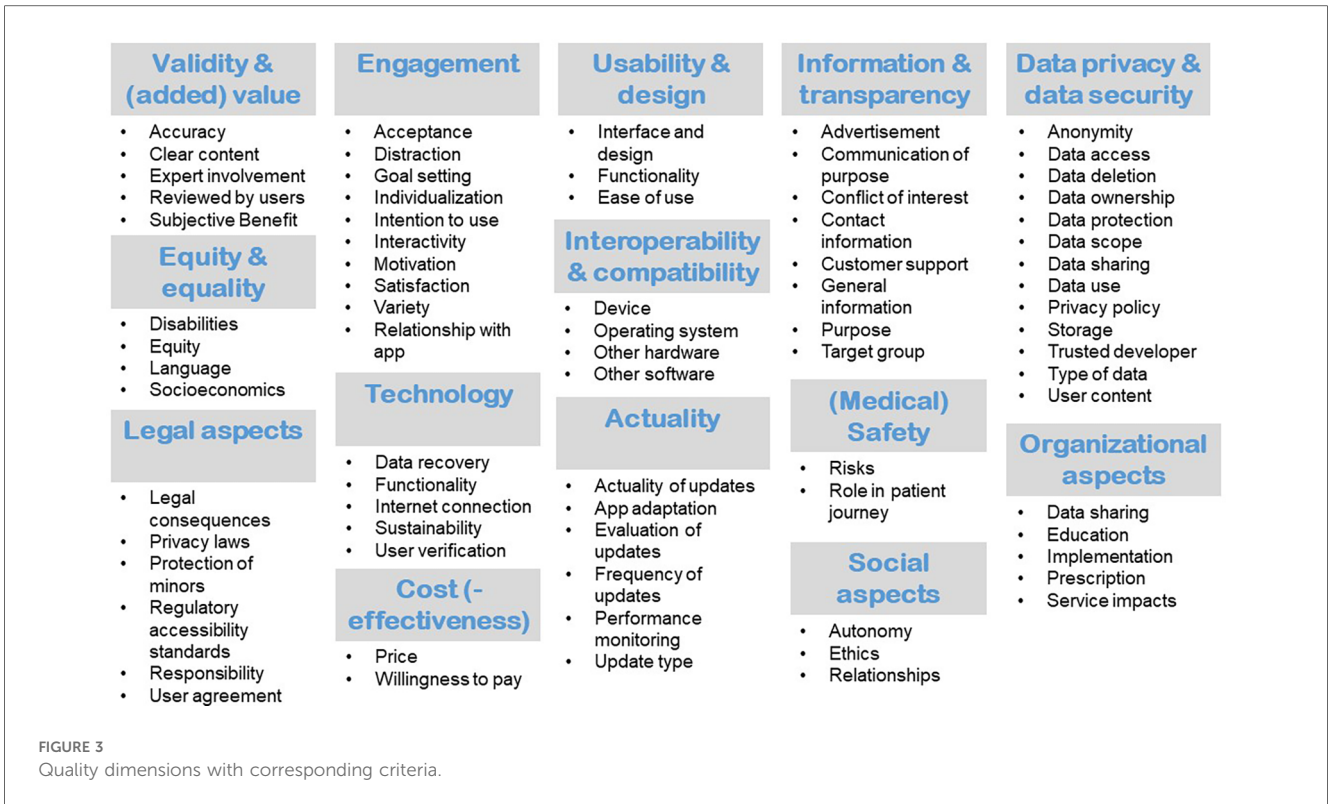
were developed based on the extracted criteria from the quality assessment tools and quality assurance systems, can be found in Supplementary Appendix D. The frequency of the individual quality dimensions in the 33 quality assessment tools and quality assurance systems included is illustrated in Figure 4.

4 Discussion

The aim of this scoping review was to identify relevant quality dimensions by searching and analyzing quality assessment tools and quality assurance systems. Thereby, the aim was not to obtain a complete survey of all available quality assessment tools. Such a list was meanwhile provided by Hajesmaeel-Gohari et al. (91). A total of 70 articles were included in the review, of which

FIGURE 4
Word cloud including the quality dimensions.

29 articles contained distinct approaches of quality assessment tools and quality assurance systems.

Of the identified approaches, some include one or two aspects, while others allow a more comprehensive assessment. For example, the Net Promoter Score (86), which de Batlle et al. (64) used alongside the SUS to evaluate an mHealth-enabled integrated care model, consists of just one item. The NPS is based on the question „How likely is it that you would recommend [name of company/product/website/services] to a friend or colleague?". In their study, de Batlle et al. (64) used the NPS to measure acceptability and thus, the score was assorted to the quality dimension "engagement" in our scoping review.

A large number of approaches from application studies were identified. In these studies, the MARS or a modified version of the MARS ($n = 19$) and the SUS ($n = 10$) were used most frequently. The MARS is a 23-item questionnaire with questions on engagement, functionality, aesthetics, the quality of the information contained and general questions for the subjective assessment of the app (18). The MARS was developed to enable a multidimensional assessment of app quality by researchers, developers, and health-professionals. The items contained in the MARS were assigned to six quality dimensions in this review, namely "information and transparency", "validity and (added) value", "engagement", "usability and design", "technology", and "equity and equality". The SUS, which is also frequently used, was developed by Brooke et al. in 1996 (20) with the intention of providing a "quick and dirty" tool for measuring the usability in industrial systems evaluation. It consists of ten elements, which were grouped into the dimensions "engagement" and "usability and design" in this review.
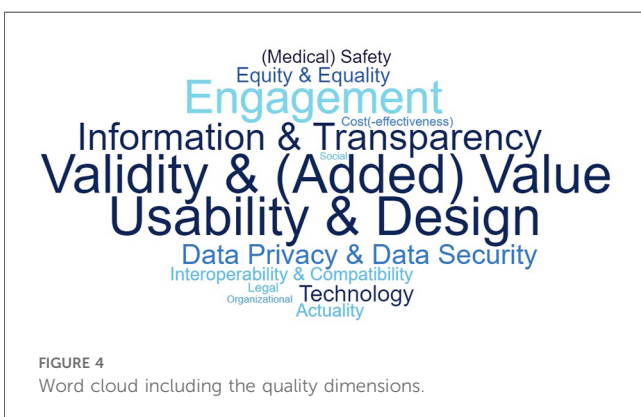
Many quality assessment tools contain questions that are easy to answer, but which reflect opinions rather than facts. For example, the items "I feel critical or disappointed in the app." (28) or "Visual appeal: How good does the app look?" (18) can be answered in subjective ways by different people. The tools therefore consist of parameters that are used to approximate the quality of the app. In this context, it is questionable how the quality of an app can be fully measured. This may also include the consideration of problems with the app.

Many of the approaches identified focus on usability. Presumably because usability is quite easy to measure and provides app developers with important insights. For example, patient safety is rarely addressed in the identified studies. Thus, the frequency of items in different questionnaires does not necessarily indicate their relevance to the healthcare system. This could be due to the target group of the approach and the complexity of the survey. However, in the next step, it is necessary to determine which aspects are relevant to the quality of apps from the perspective of the healthcare system.

Besides the quality assessment tools and given the objective of QuaSiApps (to develop a continuous quality assurance system), a particular focus of this scoping review was to identify approaches which consider mHealth apps over time. Interestingly, only five quality assurance systems could be identified. Compared to quality assessment tools, these quality assurance systems were somewhat more extensive overall and their items contributed to between five and 13 quality dimensions. For example, the items of the health technology assessment module presented by Moshi et al. (35) contributed to a total of 13 of the 14 quality dimensions in our review.

The descriptions of the 14 dimensions were derived based on the extracted criteria from the quality assessment tools and quality assurance systems. The main aim of this work was to identify relevant quality dimensions in the context of mHealth apps in order to conduct focus groups with patients and expert interviews with other stakeholders. This was done to gain further insights into each dimension and to investigate their relevance. Based on this, a set of criteria for evaluating the provision of mHealth apps will be developed.

In addition to the use within the research project, our findings can also be used as an orientation in the development of an mHealth app or related assessment instruments such as checklists. The dimensions should not be seen as a simple rating tool. Developers and researchers should critically reflect on each quality dimension.

In the following, the application of the quality dimensions "Information and Transparency" as well as "Validity & (Added) Value" will be briefly presented using the example of an mHealth app for diabetes management. In the context of "Information and Transparency", it should be ensured that the information is presented transparently and that the relevant target group is clearly defined. For example, information should be provided on the responsible manufacturer, the costs involved and how to deal with problems during the use, or what forms of support are generally available. More indication-specific, the quality dimension "Validity & (Added) Value" should ask whether the content and functions are evidence-based and in-line with published guidelines. Recorded vital signs such as blood glucose must be clear,

complete, accurate, relevant and useful. The information provided should be supported by (scientific) sources. Endocrinologists and other stakeholders should be involved in the development and evaluation process. The final application should be subject to clinical trials to demonstrate the patient benefit.

In a next step in the QuaSiApps project, these quality assurance systems will be analyzed and checked for their transferability to the German context. Interestingly, some quality assurance systems in particular show a certain degree of flexibility, thereby taking into account the dynamic developments in the mHealth sector. For example, Yasini et al. (90) developed a multidimensional scale that is completed in a web-based, self-administered questionnaire. The resulting report is both app-specific and applicable to all types of mHealth apps.

The appropriateness of the identified 14 dimensions has to be examined from a bottom-up patients' perspective as well as from a top-down healthcare system perspective to develop a quality assurance system feasible for the German health care system. As described, there are common dimensions for the quality assessment of mHealth apps that are included in many of the approaches analyzed, such as usability, data privacy and validity. Concerning the additional dimensions that we found, the question arises as to how they relate to these classic dimensions. The International Organization for Standardization (ISO) defines quality as the "degree to which a set of inherent characteristics […] of an object […] fulfils requirements" (92). It is to be discussed whether quality dimensions such as "cost (-effectiveness)" represent inherent quality characteristics of an object and thus, their suitability needs to be discussed.

As mentioned above, the approaches included in this review differ from traditional quality assurance concepts. A variety of framework concepts for quality assurance in the healthcare sector exist. They are similar to each other, but have different focuses depending on their objectives (e.g., whether they were designed for quality improvement in the healthcare system, to compare the quality of healthcare internationally, as a template for the accreditation of healthcare services, etc.). In Germany, the Institute for Quality Assurance and Transparency in Healthcare (IQTIG) acts as the central scientific institute for quality assurance in the healthcare sector. The framework concept, whose requirements are based on the principles of patient-centeredness, contains the quality dimensions "effectiveness", "safety", "responsiveness", "timeliness", "appropriateness", and "coordination and continuity" (93).

The BfArM's fast-track procedure includes requirements in its checklists relating to "product safety and functionality", "privacy and information security", "interoperability", "robustness", "consumer protection", "usability", "provider support", "medical content quality" and "patient safety". Thereby, the fast-track procedure ensures pre-selection by including criteria which are also included in many of the quality assessment tools and quality assurance systems identified in this review. The next step in the QuaSiApps project will be to analyze the transferability of results obtained in this scoping review into a concept for continuous quality assurance of DiGAs, also against the criteria already used in the fast-track procedure. QuaSiApps includes literature reviews, focus groups with users and patients, and interviews

with health care stakeholders. Based on the results, proposals for procedural purposes and quality dimensions will be formulated. These will be agreed and refined in expert workshops. The project aims to develop a set of quality aspects and corresponding quality characteristics, quality requirements, quality indicators and measurement tools.

While we are not aware of a literature review specifically on quality assurance systems, at the time of our search several literature reviews on the quality assessment of mHealth apps had already been published (91, 94–99). The most recent review of these included literature published up to December 2022 (99). The authors identified a set of 216 evaluation criteria and 6 relevant dimensions ("context", "stakeholder involvement", "development process", "evaluation", "implementation", and "features and requirements"). Although the systemization of the dimensions differs from ours, the content is comparable. This could be indicative for the relevance of the findings.

For example, Azad-Khaneghah et al. (94) conducted a systematic review to identify rating scales used to evaluate usability and quality of mHealth apps. They note that the identified scales ask about different criteria and it is therefore unclear whether the scales actually measure the same construct. Similar to our review, a theoretical basis for the construct of app quality could only be identified to a very limited extent, which is also reflected by the lack of definition of the term "quality" in the included literature. Similarly, McKay et al. (96), who conducted a systematic review of evaluation approaches for apps in the area of health behavior change, criticize the incompleteness of the evaluation criteria, resulting in the authors being unable to propose a uniform best-practice approach to the evaluation of mHealth apps.

Transferring one of the identified systems without adapting to the German healthcare system would not be appropriate. Therefore, further steps are necessary to develop a quality assurance system operating on the system level. The 14 dimensions identified need to be further explored to determine whether they address the potential risks to the quality of health care, and they need to be reflected by stakeholders in the German health care system. Our review has a number of limitations. We searched three databases and also included literature from the field of psychology by searching PsycInfo, but it cannot be ruled out that a more extensive search might have led to additional results. In addition, the exclusive use of bibliometric databases and the omission of secondary literature are limitations in this context. With regard to the search strategy, there is currently disagreement on terminology (27). Therefore, different strategies were tested beforehand and the results were compared to ensure an optimal search strategy. A further potential limitation arises from the inclusion of literature published between January 1st, 2016 and July 26th, 2021. Since we also included older quality assessment tools and quality assurance systems via application studies published during this period, relevant instruments developed before 2016, such as the SUS (20), the NPS (86), or the Silberg Scale (73) were also covered. Our review only included articles up to mid-2021. However, a current review pointed out older assessment tools

such as the SUS (1996) (20), the MARS (2015) (18), the PSSUQ (1995) (85), and the uMARS (2016) (63) are still among the most commonly used in mHealth assessment (99). Therefore, we are confident that our search strategy has enabled us to include a large proportion of quality assessment tools and quality assurance systems currently in use.

In addition, we only included articles in German and English language. It is notable that the majority of the quality assessment tools and quality assurance systems included are from English-speaking countries, which may indicate that some tools published in other languages may not have been identified. Nevertheless, we limited our search to articles in bibliographic databases, most of which are published in English. Further, descriptions for quality dimensions were formulated based on the information contained in the quality assessment tools and quality assurance systems. Thus, these descriptions are based on the subjective perception of the researchers and are not based on existing definitions. One reason for choosing this approach was the lack of international agreement on the underlying concepts. As described by Nouri et al. (97), there are major differences in the classification and definition of the individual criteria, so that usability, for example, has very different subcategories depending on the scale or is even seen as a subcategory of functionality, as in Stoyanov et al. (18). The ISO defines usability as the "extent to which a system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (100). While this definition identifies the fundamentals of usability and makes clear that effectiveness, efficiency and satisfaction are key criteria (101), ISO 9241-11:2018 is not intended to describe usability evaluation methods (100). Finally, the study protocol (27) announced the assessment of the suitability of the criteria and derived dimensions for the continuous quality assurance of mHealth apps in Germany as part of this review. In the light of our results, this seems unattainable without taking further steps (e.g., focus groups with patients). Results will be published elsewhere.

Concluding, this review serves as a building block of a continuous quality assurance system for mHealth apps in Germany. Based on our findings, we agree with Nouri et al. (97) that it is challenging to define suitable evaluation criteria for the wide range of functionalities and application areas of apps. In addition, apps are constantly evolving, which means that quality assessment tools and quality assurance systems will also need to constantly adapt.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

Author contributions

GG: Investigation, Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal Analysis, Data curation.

CS: Data curation, Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing, Conceptualization, Funding acquisition, Validation. NS: Writing – review & editing, Investigation. CA: Conceptualization, Funding acquisition, Methodology, Validation, Writing – review & editing. FP: Writing – review & editing, Validation, Formal Analysis. VH: Writing – review & editing, Investigation. DW: Investigation, Writing – review & editing. KB: Writing – review & editing, Supervision, Methodology, Conceptualization. JW: Writing – review & editing, Supervision, Methodology, Conceptualization. NB: Writing – review & editing, Supervision, Methodology. SN: Conceptualization, Writing – review & editing, Supervision, Methodology, Funding acquisition.

Conflict of interest

KB is managing director of the company QM BÖRCHERS CONSULTING+.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frhs.2024.1372871/full#supplementary-material

References

1. Statista. Population—Statistics & Facts (2023). Available online at: https://www.statista.com/topics/776/population/#topicOverview (accessed January 8, 2024).

2. Statista. Forecast Number of Mobile Users Worldwide from 2020 to 2025. (2023). Available online at: https://www.statista.com/statistics/218984/number-of-global-mobile-users-since-2010/ (accessed January 7, 2024).

3. Galetsi P, Katsalaki K, Kumar S. Exploring benefits and ethical challenges in the rise of mHealth (mobile healthcare) technology for the common good: an analysis of mobile applications for health specialists. *Technovation*. (2023) 121:102598. doi: 10.1016/j.technovation.2022.102598

4. Park YT. Emerging new era of mobile health technologies. *Healthc Inform Res*. (2016) 22(4):253–4. doi: 10.4258/hir.2016.22.4.253

5. Akter S, Ray P. Mhealth—an ultimate platform to serve the unserved. *Yearb Med Inform*. (2010) 19(1):94–100. doi: 10.1055/s-0038-1638697

6. Watanabe K, Okusa S, Sato M, Miura H, Morimoto M, Tsutsumi A. Mhealth intervention to promote physical activity among employees using a deep learning model for passive monitoring of depression and anxiety: single-arm feasibility trial. *JMIR Form Res*. (2023) 7:e51334. doi: 10.2196/51334

7. Pal S, Biswas B, Gupta R, Kumar A, Gupta S. Exploring the factors that affect user experience in mobile-health applications: a text-mining and machine-learning approach. *J Bus Res*. (2023) 156:113484. doi: 10.1016/j.jbusres.2022.113484

8. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the mobile application rating scale (MARS). *PLoS One*. (2020) 15(11):e0241480. doi: 10.1371/journal.pone.0241480

9. Schliess F, Affini Dicenzo T, Gaus N, Bourez JM, Stegbauer C, Szecsenyi J, et al. The German fast track toward reimbursement of digital health applications (DiGA): opportunities and challenges for manufacturers, healthcare providers, and people with diabetes. *J Diabetes Sci Technol*. (2022) 18(2):470–6. doi: 10.1177/19322968221121660

10. Jusob FR, George C, Mapp G. A new privacy framework for the management of chronic diseases via mHealth in a post-COVID-19 world. *Z Gesundh Wiss*. (2022) 30(1):37–47. doi: 10.1007/s10389-021-01608-9

11. Agnew JMR, Hanratty CE, McVeigh JG, Nugent C, Kerr DP. An investigation into the use of mHealth in musculoskeletal physiotherapy: scoping review. *JMIR Rehabil Assist Technol*. (2022) 9(1):e33609. doi: 10.2196/33609

12. Giebel GD, Speckemeier C, Abels C, Plescher F, Börchers K, Wasem J, et al. Problems and barriers related to the use of digital health applications: a scoping review. *J Med Internet Res*. (2023) 25:e43808. doi: 10.2196/43808

13. WHO. Quality of care. (2023) Available online at: https://www.who.int/health-topics/quality-of-care#tab=tab_1 (accessed January 6, 2024).

14. Schang L, Blotenberg I, Boywitt D. What makes a good quality indicator set? A systematic review of criteria. *Int J Qual Health Care*. (2021) 33(3):1–10. doi: 10.1093/intqhc/mzab107

15. Motamarri S, Akter S, Ray P, Tseng C-L. Distinguishing mHealth from other health care systems in developing countries: a study on service quality perspective. *Commun Assoc Inform Syst*. (2014) 34:669–92. doi: 10.17705/1CAIS.03434

16. BfArM. Das Fast-Track-Verfahren für digitale Gesundheitsanwendungen (DiGA) nach § 139e SGB V. (2023) Available online at: https://www.bfarm.de/SharedDocs/Downloads/DE/Medizinprodukte/diga_leitfaden.pdf?__blob=publicationFile (accessed January 1, 2024).

17. GKV-SV. Fokus: Digitale Gesundheitsanwendungen (DiGA) (2023) Available online at: https://www.gkv-spitzenverband.de/gkv_spitzenverband/presse/fokus/fokus_diga.jsp (accessed January 1, 2024).

18. 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. doi: 10.2196/mhealth.3422

19. Baumel A, Faber K, Mathur N, Kane JM, Muench F. Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. *J Med Internet Res*. (2017) 19(3):e82. doi: 10.2196/jmir.7270

20. Brooke J. SUS—a quick and dirty usability scale. In: *Usability Evaluation in Industry*. London: CRC Press (1996). p. 189–94.

21. Sadegh SS, Khakshour Saadat P, Sepehri MM, Assadi V. A framework for m-health service development and success evaluation. *Int J Med Inform*. (2018) 112:123–30. doi: 10.1016/j.ijmedinf.2018.01.003

22. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. *NPJ Digit Med*. (2019) 2:38. doi: 10.1038/s41746-019-0111-3

23. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between

a systematic or scoping review approach. *BMC Med Res Methodol.* (2018) 18(1):143. doi: 10.1186/s12874-018-0611-x

24. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc.* (2015) 13 (3):141–6. doi: 10.1097/XEB.0000000000000050

25. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci.* (2010) 5:69. doi: 10.1186/1748-5908-5-69

26. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med.* (2018) 169(7):467–73. doi: 10.7326/M18-0850

27. Giebel GD, Schrader NF, Speckemeier C, Abels C, Börchers K, Wasem J, et al. Quality assessment of digital health applications: protocol for a scoping review. *JMIR Res Protoc.* (2022) 11(7):1–6. doi: 10.2196/36974

28. Berry K, Salter A, Morris R, James S, Bucci S. Assessing therapeutic alliance in the context of mHealth interventions for mental health problems: development of the mobile agnew relationship measure (mARM) questionnaire. *J Med Internet Res.* (2018) 20(4):e90. doi: 10.2196/jmir.8252

29. Brown W 3rd, Yen PY, 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):1080–7. doi: 10.1016/j.jbi.2013.08.001

30. Camacho E, Hoffman L, Lagan S, Rodriguez-Villa E, Rauseo-Ricupero N, Wisniewski H, et al. Technology evaluation and assessment criteria for health apps (TEACH-apps): pilot study. *J Med Internet Res.* (2020) 22(8):e18346. doi: 10.2196/18346

31. Glattacker M, Boeker M, Anger R, Reichenbach F, Tassoni A, Bredenkamp R, et al. Evaluation of a mobile phone app for patients with pollen-related allergic rhinitis: prospective longitudinal field study. *JMIR Mhealth Uhealth.* (2020) 8(4):e15514. doi: 10.2196/15514

32. Huang Z, Lum E, Car J. Medication management apps for diabetes: systematic assessment of the transparency and reliability of health information dissemination. *JMIR Mhealth Uhealth.* (2020) 8(2):e15364. doi: 10.2196/15364

33. Liu X, Jin F, Hsu J, Li D, Chen W. Comparing smartphone apps for traditional Chinese medicine and modern medicine in China: systematic search and content analysis. *JMIR Mhealth Uhealth.* (2021) 9(3):e27406. doi: 10.2196/27406

34. Minge M, Riedel L. meCUE—Ein modularer Fragebogen zur Erfassung des Nutzungserlebens. (2013).

35. Moshi MR, Tooher R, Merlin T. Development of a health technology assessment module for evaluating mobile medical applications. *Int J Technol Assess Health Care.* (2020) 36(3):252–61. doi: 10.1017/S0266462320000288

36. O'Rourke T, Pryss R, Schlee W, Probst T. Development of a multidimensional app quality assessment tool for health-related apps (AQUA). *Digit Psychol.* (2020) 1 (2):13–23. doi: 10.24989/dp.v1i2.1816

37. Pifarre M, Carrera A, Vilaplana J, Cuadrado J, Solsona S, Abella F, et al. TControl: a mobile app to follow up tobacco-quitting patients. *Comput Methods Programs Biomed.* (2017) 142:81–9. doi: 10.1016/j.cmpb.2017.02.022

38. Tan R, Cvetkovski B, Kritikos V, O'Hehir RE, Lourenco O, Bousquet J, et al. Identifying an effective mobile health application for the self-management of allergic rhinitis and asthma in Australia. *J Asthma.* (2020) 57(10):1128–39. doi: 10. 1080/02770903.2019.1640728

39. Wood J, Jenkins S, Putrino D, Mulrennan S, Morey S, Cecins N, et al. High usability of a smartphone application for reporting symptoms in adults with cystic fibrosis. *J Telemed Telecare.* (2018) 24(8):547–52. doi: 10.1177/1357633X17723366

40. Llorens-Vernet P, Miro J. Standards for mobile health-related apps: systematic review and development of a guide. *JMIR Mhealth Uhealth.* (2020) 8(3):e13057. doi: 10.2196/13057

41. Broekhuis M, van Velsen L, Hermens H. Assessing usability of eHealth technology: a comparison of usability benchmarking instruments. *Int J Med Inform.* (2019) 128:24–31. doi: 10.1016/j.ijmedinf.2019.05.001

42. Dawson RM, Felder TM, Donevant SB, McDonnell KK, Card EB 3rd, King CC, et al. What makes a good health 'app'? Identifying the strengths and limitations of existing mobile application evaluation tools. *Nurs Inq.* (2020) 27(2):e12333. doi: 10. 1111/nin.12333

43. Llorens-Vernet P, Miro J. The mobile app development and assessment guide (MAG): Delphi-based validity study. *JMIR Mhealth Uhealth.* (2020) 8(7):e17760. doi: 10.2196/17760

44. Miro J, Llorens-Vernet P. Assessing the quality of mobile health-related apps: interrater reliability study of two guides. *JMIR Mhealth Uhealth.* (2021) 9(4):e26471. doi: 10.2196/26471

45. Alhuwail D, Albaj R, Ahmad F, Aldakheel K. The state of mental digi-therapeutics: a systematic assessment of depression and anxiety apps available for arabic speakers. *Int J Med Inform.* (2020) 135:104056. doi: 10.1016/j.ijmedinf.2019. 104056

46. Amor-Garcia MA, Collado-Borrell R, Escudero-Vilaplana V, Melgarejo-Ortuno A, Herranz-Alonso A, Arranz Arija JA, et al. Assessing apps for patients with genitourinary tumors using the mobile application rating scale (MARS): systematic search in app stores and content analysis. *JMIR Mhealth Uhealth.* (2020) 8(7):e17609. doi: 10.2196/17609

47. Choi YK, Demiris G, Lin SY, Iribarren SJ, Landis CA, Thompson HJ, et al. Smartphone applications to support sleep self-management: review and evaluation. *J Clin Sleep Med.* (2018) 14(10):1783–90. doi: 10.5664/jcsm.7396

48. Davalbhakta S, Advani S, Kumar S, Agarwal V, Bhoyar S, Fedirko E, et al. A systematic review of smartphone applications available for Corona virus disease 2019 (COVID19) and the assessment of their quality using the mobile application rating scale (MARS). *J Med Syst.* (2020) 44(9):164. doi: 10.1007/s10916-020-01633-3

49. Escoffery C, McGee R, Bidwell J, Sims C, Thropp EK, Frazier C, et al. A review of mobile apps for epilepsy self-management. *Epilepsy Behav.* (2018) 81:62–9. doi: 10. 1016/j.yebeh.2017.12.010

50. Fuller-Tyszkiewicz M, Richardson B, Little K, Teague S, Hartley-Clark L, Capic T, et al. Efficacy of a smartphone app intervention for reducing caregiver stress: randomized controlled trial. *JMIR Ment Health.* (2020) 7(7):e17541. doi: 10.2196/17541

51. Gong E, Zhang Z, Jin X, Liu Y, Zhong L, Wu Y, et al. Quality, functionality, and features of Chinese mobile apps for diabetes self-management: systematic search and evaluation of mobile apps. *JMIR Mhealth Uhealth.* (2020) 8(4):e14836. doi: 10.2196/14836

52. Mehdi M, Stach M, Riha C, Neff P, Dode A, Pryss R, et al. Smartphone and mobile health apps for tinnitus: systematic identification, analysis, and assessment. *JMIR Mhealth Uhealth.* (2020) 8(8):e21767. doi: 10.2196/21767

53. Muntaner-Mas A, Martinez-Nicolas A, Lavie CJ, Blair SN, Ross R, Arena R, et al. A systematic review of fitness apps and their potential clinical and sports utility for objective and remote assessment of cardiorespiratory fitness. *Sports Med.* (2019) 49 (4):587–600. doi: 10.1007/s40279-019-01084-y

54. Myers A, Chesebrough L, Hu R, Turchioe M, Pathak J, Creber R. Evaluating commercially available mobile apps for depression self-management. *AMIA Annu Symp Proc.* (2021) 25:906–14.

55. Nguyen M, Hossain N, Tangri R, Shah J, Agarwal P, Thompson-Hutchison F, et al. Systematic evaluation of Canadian diabetes smartphone applications for people with type 1, type 2 and gestational diabetes. *Can J Diabetes.* (2021) 45 (2):174–8.e1. doi: 10.1016/j.jcjd.2020.07.005

56. Pearsons A, Hanson CL, Gallagher R, O'Carroll RE, Khonsari S, Hanley J, et al. Atrial fibrillation self-management: a mobile telephone app scoping review and content analysis. *Eur J Cardiovasc Nurs.* (2021) 20(4):305–14. doi: 10.1093/eurjcn/zvaa014

57. Sereda M, Smith S, Newton K, Stockdale D. Mobile apps for management of tinnitus: users' survey, quality assessment, and content analysis. *JMIR Mhealth Uhealth.* (2019) 7(1):e10353. doi: 10.2196/10353

58. Shang J, Wei S, Jin J, Zhang P. Mental health apps in China: analysis and quality assessment. *JMIR Mhealth Uhealth.* (2019) 7(11):e13236. doi: 10.2196/13236

59. Terhorst Y, Rathner E-M, Baumeister H, Sander L. «Hilfe aus dem app-store?»: eine systematische Übersichtsarbeit und evaluation von apps zur anwendung bei depressionen ['help from the app store?': a systematic review of depression apps in German app stores]. *Verhaltenstherapie.* (2018) 28(2):101–12. doi: 10.1159/000481692

60. Terhorst Y, Messner EM, Schultchen D, Paganini S, Portenhauser A, Eder AS, et al. Systematic evaluation of content and quality of English and German pain apps in European app stores. *Internet Interv.* (2021) 24:100376. doi: 10.1016/j. invent.2021.100376

61. Wang X, Markert C, Sasangohar F. Investigating popular mental health mobile application downloads and activity during the COVID-19 pandemic. *Hum Factors.* (2023) 65(1):50–61. doi: 10.1177/0018720821998110

62. Choi SK, Yelton B, Ezeanya VK, Kannaley K, Friedman DB. Review of the content and quality of mobile applications about Alzheimer's disease and related dementias. *J Appl Gerontol.* (2020) 39(6):601–8. doi: 10.1177/0733464818790187

63. 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. doi: 10.2196/mhealth.5849

64. de Batlle J, Massip M, Vargiu E, Nadal N, Fuentes A, Ortega Bravo M, et al. Implementing mobile health-enabled integrated care for complex chronic patients: patients and professionals' acceptability study. *JMIR Mhealth Uhealth.* (2020) 8(11):e22136. doi: 10.2196/22136

65. Eiring O, Nytroen K, Kienlin S, Khodambashi S, Nylenna M. The development and feasibility of a personal health-optimization system for people with bipolar disorder. *BMC Med Inform Decis Mak.* (2017) 17(1):102. doi: 10.1186/s12911-017-0481-x

66. Grainger R, Townsley HR, Ferguson CA, Riley FE, Langlotz T, Taylor WJ. Patient and clinician views on an app for rheumatoid arthritis disease monitoring: function, implementation and implications. *Int J Rheum Dis.* (2020) 23(6):813–27. doi: 10.1111/1756-185X.13850

67. Henshall C, Davey Z. Development of an app for lung cancer survivors (iEXHALE) to increase exercise activity and improve symptoms of fatigue, breathlessness and depression. *Psychooncology.* (2020) 29(1):139–47. doi: 10.1002/pon.5252

68. Hoogeveen IJ, Peeks F, de Boer F, Lubout CMA, de Koning TJ, Te Boekhorst S, et al. A preliminary study of telemedicine for patients with hepatic glycogen storage disease and their healthcare providers: from bedside to home site monitoring. *J Inherit Metab Dis*. (2018) 41(6):929–36. doi: 10.1007/s10545-018-0167-2

69. Kizakevich PN, Eckhoff R, Brown J, Tueller SJ, Weimer B, Bell S, et al. PHIT for duty, a mobile application for stress reduction, sleep improvement, and alcohol moderation. *Mil Med*. (2018) 183(suppl_1):353–63. doi: 10.1093/milmed/usx157

70. Pelle T, van der Palen J, de Graaf F, van den Hoogen FHJ, Bevers K, van den Ende CHM. Use and usability of the Dr. Bart app and its relation with health care utilisation and clinical outcomes in people with knee and/or hip osteoarthritis. *BMC Health Serv Res*. (2021) 21(1):444. doi: 10.1186/s12913-021-06440-1

71. Seitz MW, Haux C, Smits KPJ, Kalmus O, Van Der Zande MM, Lutyj J, et al. Development and evaluation of a mobile patient application to enhance medical-dental integration for the treatment of periodontitis and diabetes. *Int J Med Inform*. (2021) 152:104495. doi: 10.1016/j.ijmedinf.2021.104495

72. Veazie S, Winchell K, Gilbert J, Paynter R, Ivlev I, Eden KB, et al. Rapid evidence review of mobile applications for self-management of diabetes. *J Gen Intern Med*. (2018) 33(7):1167–76. doi: 10.1007/s11606-018-4410-1

73. Silberg W, Lundberg G, Musacchio R. Assessing, controlling, and assuring the quality of medical information on the internet: caveant lector et viewor–let the reader and viewer beware. *JAMA*. (1997) 277(15):1244–5. doi: 10.1001/jama.1997.03540390074039

74. 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. doi: 10.3233/978-1-61499-658-3-1030

75. Zhang MW, Ho RC, Loh A, Wing T, Wynne O, Chan SWC, et al. Current status of postnatal depression smartphone applications available on application stores: an information quality analysis. *BMJ Open*. (2017) 7(11):e015655. doi: 10.1136/bmjopen-2016-015655

76. Athilingam P, Labrador MA, Remo EF, Mack L, San Juan AB, Elliott AF. Features and usability assessment of a patient-centered mobile application (HeartMapp) for self-management of heart failure. *Appl Nurs Res*. (2016) 32:156–63. doi: 10.1016/j.apnr.2016.07.001

77. Crosby LE, Ware RE, Goldstein A, Walton A, Joffe NE, Vogel C, et al. Development and evaluation of iManage: a self-management app co-designed by adolescents with sickle cell disease. *Pediatr Blood Cancer*. (2017) 64(1):139–45. doi: 10.1002/pbc.26177

78. Fougerouse P, Yasini M, Marchand G, Aalami O. A cross-sectional study of prominent US Mobile health applications: evaluating the current landscape. *AMIA Annu Symp Proc*. (2017) 2017:715–23.

79. Han MN, Grisales T, Sridhar A. Evaluation of a mobile application for pelvic floor exercises. *Telemed J E Health*. (2019) 25(2):160–4. doi: 10.1089/tmj.2017.0316

80. Huckvale K, Torous J, Larsen ME. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. *JAMA Netw Open*. (2019) 2(4):e192542. doi: 10.1001/jamanetworkopen.2019.2542

81. Turchioe M, Grossman LV, Baik D, Lee CS, Maurer MS, Goyal P, et al. Older adults can successfully monitor symptoms using an inclusively designed mobile application. *J Am Geriatr Soc*. (2020) 68(6):1313–8. doi: 10.1111/jgs.16403

82. Sedhom R, McShea MJ, Cohen AB, Webster JA, Mathews SC. Mobile app validation: a digital health scorecard approach. *NPJ Digit Med*. (2021) 4(1):111. doi: 10.1038/s41746-021-00476-7

83. Doak C, Doak L, Root J. Teaching patients with low literacy skills. *Am J Nurs*. (1996) 96(12):16M. doi: 10.1097/00000446-199612000-00022

84. Huckvale K, Prieto JT, Tilney M, Benghozi PJ, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Med*. (2015) 13:214. doi: 10.1186/s12916-015-0444-y

85. Lewis JR. IBM Computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int J Hum Comput Interact*. (1995) 7:57. doi: 10.1080/10447319509526110

86. Reichheld F. The one number you need to grow. *Harv Bus Rev*. (2004) 81:46–54.

87. Ryu Y, Smith-Jackson T. Reliability and validity of the mobile phone usability questionnaire (MPUQ). *J Usability Stud*. (2006) 2(1):39–53.

88. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (health-ITUES) for usability assessment of mobile health technology: validation study. *JMIR Mhealth Uhealth*. (2018) 6(1):e4. doi: 10.2196/mhealth.8851

89. Shoemaker SJ, Wolf MS, Brach C. Development of the patient education materials assessment tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. *Patient Educ Couns*. (2014) 96(3):395–403. doi: 10.1016/j.pec.2014.05.027

90. 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. doi: 10.3233/978-1-61499-678-1-205

91. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. *BMC Med Inform Decis Mak*. (2022) 22(1):22. doi: 10.1186/s12911-022-01764-2

92. ISO.org. ISO 9000:2015(en) Quality management systems — Fundamentals and vocabula. (2015). Available online at: https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:en (accessed January 1, 2024).

93. IQTIG. Methodische Grundlagen. Version 2.0. (2022). Available online at: https://iqtig.org/downloads/berichte-2/meg/IQTIG_Methodische-Grundlagen_Version-2.0_2022-04-27_barrierefrei.pdf (accessed January 1, 2024).

94. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. *Disabil Rehabil Assist Technol*. (2021) 16(7):712–21. doi: 10.1080/17483107.2019.1701103

95. Fiore P. How to evaluate mobile health applications: a scoping review. *Stud Health Technol Inform*. (2017) 234:109–14. doi: 10.3233/978-1-61499-742-9-109

96. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. *J Telemed Telecare*. (2018) 24(1):22–30. doi: 10.1177/1357633X16673538

97. Nouri R, Sharareh RNK, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. *J Am Med Inform Assoc*. (2018) 25(8):1089–98. doi: 10.1093/jamia/ocy050

98. Pham Q, Graham G, Carrion C, Morita PP, Seto E, Stinson JN, et al. A library of analytic indicators to evaluate effective engagement with consumer mHealth apps for chronic conditions: scoping review. *JMIR Mhealth Uhealth*. (2019) 7(1):e11941. doi: 10.2196/11941

99. Ribaut J, DeVito Dabbs A, Dobbels F, Teynor A, Mess EV, Hoffmann T, et al. Developing a comprehensive list of criteria to evaluate the characteristics and quality of eHealth smartphone apps: systematic review. *JMIR Mhealth Uhealth*. (2024) 12: e48625. doi: 10.2196/48625

100. ISO.org. ISO 9241-11:2018(en) Ergonomics of human-system interaction. Part 11: Usability: Definitions and concepts (2018). Available online at: https://www.iso.org/standard/63500.html (accessed January 8, 2024).

101. Bevan N, Carter J, Earthy J, Geis T, Harker S. New ISO standards for usability, usability reports and usability measures. *Lect Notes Comput Sci*. (2016) 9731:268–78. doi: 10.1007/978-3-319-39510-4_25