An umbrella review of usability assessment methods for mobile apps for physical rehabilitation

Sylvia Hach, Gemma Alder, Verna Stavric, Denise Taylor, Nada Signal

Submitted to: JMIR mHealth and uHealth
on: May 29, 2023

Table of Contents

An umbrella review of usability assessment methods for mobile apps for physical rehabilitation

Sylvia Hach[1*] BA, MA, BSc, PhD; Gemma Alder[1*] BSc, MSc, PhD; Verna Stavric[1*] BSc, MHSc, PhD; Denise Taylor[1*] PhD, MSc, Grad Dip Physiotherapy, Certificate Health Economics; Nada Signal[1*] PhD, MHSc, BHSc

[1]Auckland University of Technology Health and Rehabilitation Research Institute Auckland NZ
[*]these authors contributed equally

Corresponding Author:
Sylvia Hach BA, MA, BSc, PhD
Auckland University of Technology
Health and Rehabilitation Research Institute
Faculty of Health and Environmental Sciences
Auckland
NZ

Abstract

Background: Usability has been touted as one determiner of success of mobile Health (mHealth) interventions. While multiple systematic reviews of usability assessment approaches for different mHealth solutions for physical rehabilitation are available, there is a lack of synthesis of this portion of the literature.

Objective: To summarise systematic reviews examining usability assessment tools in mHealth interventions including physical rehabilitation.

Methods: An umbrella review was conducted according to a published registered protocol (CRD42022338785). A topic-based search was conducted of Pubmed, Cochrane, IEEEXplore, Epistemonikos and Web of Science from January 2015 to April 2023 for systematic reviews investigating usability assessment tools in mHealth interventions including physical exercise rehabilitation. Eligibility screening, data extraction and assessment of the methodological quality (AMSTAR2) was completed and data tabulated for synthesis.

Results: A total of 12 systematic reviews were included of which three (25%) did not refer to any theoretical usability framework. The sample referenced a total of 32 usability assessment tools and 66 custom-made, as well as hybrid, measures. Information on psychometric properties was included for nine instruments (28%) with satisfactory internal consistency and structural validity. A lack of reliability, responsiveness and cross-cultural validity data was found. The methodological quality of the systematic reviews was limited with eight studies (67%) displaying two or more critical weaknesses.

Conclusions: There is significant diversity in the usability assessment of mHealth for rehabilitation and a link to theoretical models is often lacking. There is widespread use of custom-made measures and pre-existing measures often do not display sufficient psychometric strength. As a result, existing mHealth usability evaluations are difficult to compare. It is proposed that multi-method usability assessment is employed and that, in the selection of usability measures, there is a focus on explicit reference to their theoretical underpinning and acceptable psychometric properties. Clinical Trial: PROSPERO CRD42022338785; https://www.crd.york.ac.uk/prospero/#recordDetails

(JMIR Preprints 29/05/2023:49449)
DOI: https://doi.org/10.2196/preprints.49449

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

An umbrella review of usability assessment methods for mobile apps for physical rehabilitation

Running title: Usability assessment methods of physical rehabilitation mobile apps

Sylvia Hach[1], Gemma Alder[1], Verna Stavric[1], Denise Taylor[1], & Nada Signal[1]

[1]Health and Rehabilitation Research Institute, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand

Author contact details:

Sylvia Hach

Health and Rehabilitation Research Institute, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand
Email: sylvia.hach@gmail.com

Abstract

Background: Usability has been touted as one determiner of success of mobile Health (mHealth) interventions. Multiple systematic reviews of usability assessment approaches for different mHealth solutions for physical rehabilitation are available. However, there is a lack of synthesis in this portion of the literature which results in clinicians and developers devoting a significant amount of time and effort in analysing and summarising a large body of systematic reviews.

Objective: To summarise systematic reviews examining usability assessment instruments, or measurements tools, in mHealth interventions including physical rehabilitation.

Methods: An umbrella review was conducted according to a published registered protocol. A topic-based search was conducted of Pubmed, Cochrane, IEEEXplore, Epistemonikos, Web of Science and CINAHL Complete from January 2015 to April 2023 for systematic reviews investigating usability assessment instruments in mHealth interventions including physical exercise rehabilitation. Eligibility screening included date, language, participant, and article type. Data extraction and assessment of the methodological quality (AMSTAR2) was completed and tabulated for synthesis.

Results: A total of 12 systematic reviews were included of which three (25%) did not refer to any theoretical usability framework and the remaining most commonly referencing the ISO framework. The sample referenced a total of 32 usability assessment instruments and 66 custom-made, as well as hybrid, instruments. Information on psychometric properties was included for nine instruments (28%) with satisfactory internal consistency and structural validity. A lack of reliability, responsiveness and cross-cultural validity data was found. The methodological quality of the systematic reviews was limited with eight studies (67%) displaying two or more critical weaknesses.

Conclusions: There is significant diversity in the usability assessment of mHealth for rehabilitation and a link to theoretical models is often lacking. There is widespread use of custom-made instruments and pre-existing instruments often do not display sufficient psychometric strength. As a result, existing mHealth usability evaluations are difficult to compare. It is proposed that multi-method usability assessment is employed and that, in the selection of usability assessment instruments, there is a focus on explicit reference to their theoretical underpinning and acceptable psychometric properties. This could be facilitated by a closer collaboration between researchers, developers, and clinicians throughout the phases of mHealth tool development.

Protocol registration: PROSPERO CRD42022338785; https://www.crd.york.ac.uk/prospero/#recordDetails

Keywords: Usability, quality evaluation, mobile health, physical exercise, rehabilitation, overview, umbrella review, psychometrics.

Funding details: none

Word count: 3780 (excluding abstract, tables, references & Supplementary Files) (abstract: 354)

Introduction

The development of mobile health (mHealth) [1,2] solutions has seen exponential growth in recent times, driven particularly by the global pandemic [3,4]. Mobile health has been heralded as a tool to provide access to quality rehabilitation input for patients outside of the time they are able to spend with clinicians [5] and for patients in geographically remote areas [6]. Furthermore, similar to the observed trend of increased health information seeking on the internet [7], the democratisation of access to rehabilitation could be achieved by individuals actively seeking standalone mHealth solutions.

However, there is also increasing awareness that mHealth solutions available to clinicians and their patients often lack quality evaluations [8,9]. Many mHealth solutions only have short-term (<30 days) data from small sample sizes to support their effectiveness [10]. Moreover, only limited standardised outcome measures are typically utilised [11,12].

Usability is one key aspect commonly included in the evaluation of mHealth solutions [5,9,11]. Usability has been touted as a determiner of the success of mHealth interventions [13]. Usability is often delineated from two related concepts. First, the concept of utility which captures a system's ability to meet user needs [14]. Second, user experience is commonly understood as a broader concept of the experience of utilising a mHealth solution and may include measures of user beliefs [15]. However, usability may or may not be part of how user experience is captured, and many different definitions of usability appear in the literature [15–17].

The diversity in definitions of usability is mirrored by the diversity in usability models or frameworks. The five most commonly cited models of usability are that of ISO9241-11 [18] and its revision [19], ISO/IEC25010 [20], Nielsen's usability model [21], and, in the context of health in particular, the People At the Centre of Mobile Application Development (PACMAD) model [14,22]. These models identify factors such as efficiency, or the resources expended to achieve a task, effectiveness, the level of accuracy and completeness of a task achieved using a mobile solution, and satisfaction or positive user interaction whilst operating the mobile solution as components of usability. The key difference between the PACMAD and the aforementioned frameworks is that these and other factors such as errors, are seen as arising from three different sources - the user themselves, the task, and the context. This could be argued to be of particular importance for mHealth, where users may experience limitations such as perceptual or cognitive (aging) barriers [23]. These additionally impact on task demands and therefore represent an important consideration in the design of mHealth tools.

Usability assessment has been included in several good practice guidelines for the development of mHealth solutions [24–28] as well as in many evaluation frameworks [29,30], and can be regarded as a crucial step for evaluation at different stages of the typical mHealth development cycles. To date, however, no accepted standard for the assessment of usability of mHealth solutions exists. This means that researchers and developers of mHealth are faced with difficult decisions when designing mHealth evaluation procedures that strike the balance between responsiveness, reliability and validity and are unable to compare existing solutions for the purpose of innovating. Further, clinicians are unable to be guided in their prescription of mHealth solutions and there are significant barriers for consumers to engage with existing solutions.

Numerous systematic reviews have explored usability assessment approaches for various mHealth solutions in the context of physical rehabilitation. However, there is a lack of synthesis in this area of the literature. This may contribute to clinicians and developers needing to devote a significant amount of time and effort in analysing and summarising a large body of systematic reviews. An umbrella review can act as "a means for a rapid review of the evidence to address a broad and high-quality evidence base" [31]. Specifically, an umbrella review allows for a broader scope compared to individual systematic reviews which may focus on individual treatment options or individual conditions [32–34]. Hence, the aim of this umbrella review was to provide a 'user-friendly' summary of the use of usability assessment instruments, or measurement tools, for researchers, clinicians, and consumers of mHealth irrespective of the specific area of application (e.g., diabetes, tuberculosis, sleep). Specifically, the objective was to summarise systematic reviews that investigated usability assessment instruments in mHealth interventions including those related to physical exercise rehabilitation. It is envisaged that such a summary will firstly aide researchers, developers, and clinicians to gain an overview of usability assessment instruments without needing to explore primary literature. Secondly, the presented summary may aide the development of mHealth usability assessment standards.

Methods

The umbrella review protocol was developed based on the Cochrane Handbook of Systematic Reviews of Interventions [33] and other relevant methodology sources [e.g., 34] and was registered with PROSPERO (CRD42022338785). StArt software [35] was utilised for the first and second level screening of results datasets and extracting relevant information.

Inclusion criteria

Based on the objectives of the study, the following inclusion criteria were formulated; (i) Articles published between 01/01/2015 and 27/04/2023. The date range reflected the launch of Apple ResearchKit in 2015 which accelerated mHealth development and research [36]; (ii) containing data on human participants; (iii) with the 'unit of searching' [33] being "systematic reviews" [37,38] in order to reduce the effect of cumulative bias that may arise when including non-systematic reviews; (iv) examining usability assessment instruments of mobile applications for health professionals and for healthcare consumers and (v) published in the English language to enable all contributing authors to perform screening, extraction and synthesis of the search results. No post-hoc modifications were made to the inclusion criteria. Systematic reviews of usability assessment instruments of other (mobile) solutions such as wearables, sensors, Virtual Reality, blockchain, Internet of Things (IoT), simulated data or solutions for healthcare professionals only were excluded.

Search methods and search terms

The following databases were searched with a combination of the search terms mobile application*, mobile app, usab*, usab* criteria, usab* evaluat*, systematic review, mhealth, mobile health, physical exercise: Pubmed, Cochrane, IEEEXplore, Epistemonikos, Web of Science and CINAHL Complete combined using Boolean operator OR and AND and customised for each database in accordance with their filtering specifications. The results sets were imported into StArt [35].

Data collection and analysis

A preliminary search of existing systematic reviews was conducted prior to finalising the search terms in order to scope the extent and type of existing evidence [33] .The subsequent final search terms produced a results set that was more refined in focus and feasible in terms of the size of the expected results set. Following the removal of duplicates, two-level screening was performed: Title and abstract screen was performed by the primary author [SH] and a randomly selected subset of articles (n=118, ~10%) was screened by a second author [VS], kappa .87. Second-level, full text screening was performed by the primary author [SH] using StArt for data extraction from the final results set. A data extraction form including basic reference details, as well as information such as population of interest and interventions studied, was discussed and agreed upon by three authors [SH, GA, NS] prior to data extraction (see review protocol PROSPERO CRD42022338785 for more detail).

Quality assessment was completed using AMSTAR2 by the primary author [SH] and a second author [VS] separately, kappa .823. Any disagreement was discussed and resolved via consensus. In line with recommendations by Shea et al. [39], a discussion to determine AMSTAR2 critical domains for the present overview occurred among two authors [SH, NS]. Criteria 2, 4 and 7 were retained on the premise of constituting critical criteria as defined by the original publication [39]. The original critical criteria 9, 11, 13 and 15 were classified as non-critical for the purpose of the present overview due to pertaining to meta-analytic steps which none of the included systematic reviews performed. Instead, the following criteria were classified as critical: Criterion 5 due to the variety of study designs and target user groups/clinical contexts included within the systematic reviews; Criterion 16 due to the context of mHealth usability where the borders between academic enquiry and commercialisation are more blurred and funding could constitute a significant source of bias/conflict of interest. A summary rating was produced according to recommendations by Shea et al. [39].

Finally, in order to gauge potential skewing of the data caused by significant overlap of primary studies contained within the systematic reviews included in the present overview [40], overlap assessment was achieved via citation matrix [41,42] for the systematic reviews including the System Usability Scale (SUS) as an exemplar. The SUS was chosen given it is one of the most-well known instruments [43]and preliminary searches of the literature demonstrated its frequency of use and reference.

Data availability

Lead authors of protocols of systematic reviews were also contacted. If pre-print manuscripts were available, these were included in the present overview. Systematic reviews not able to be obtained through the authors of the protocol were listed separately for the purpose of the overview [33]. All data accessed and created as part of the present overview are included as part of the article and its Supplementary Files.

Results

The initial database search returned 1479 results which were reduced to 1375 after removal of duplicates (see Figure 1). Title and abstract screening resulted in 27 articles included for full text screening. Fifteen of the full text articles retrieved (see Supplementary Table S1) were ineligible because they did not review usability assessment measures, include sufficient detail on usability assessment instruments (e.g., including binary information only), did not include a literature review or examined non-health mobile service categories (see Figure 1).

A total of 12 systematic reviews examining usability assessment instruments were included. Data was extracted (see Supplementary Table S2) as per the registered protocol. Across the systematic reviews included, there was coverage of primary studies from the start of records to 2020. Three of the systematic reviews included examined usability assessment instruments within a specific target user group (e.g., diabetic users [44]; users living with a mental health concern [45,46]). The remaining nine systematic reviews [47–55] focused on usability assessment instruments employed across different target user populations. Usability models/frameworks referenced included ISO [20] (referenced in [44,48,49,55]), Nielsen [21] (referenced in [45]) and the framework by the Canadian Institutes of Health Research (CIHR) and the Mental Health Commission Canada (MHCC) [56] (referenced in [47]). Three (25%) of the systematic reviews [46,50,51] included in the present overview did not refer to any theoretical framework (see Supplementary Table S2).

The systematic reviews included identified a total of 32 usability assessment instruments (see Table 1) and a further 66 custom-made usability assessment instruments as well as hybrid custom-made instruments (see Supplementary Table S3). The most commonly referenced usability assessment instrument was the System Usability Scale (SUS) [57], followed by the IBM Computer Usability Satisfaction Questionnaire [58], and the Usefulness Satisfaction and Ease of Use (USE) Questionnaire [59].

Data regarding the psychometric properties of nine instruments (28%) [57,58,60–65] was included in the systematic reviews as detailed in Table 1. Internal consistency was generally good across these instruments, content validity was provided through expert panel or focus groups [57,58,60,61,63–65] and exploratory and/or confirmatory factor analyses were employed in evidence of structural validity [57,60,61,64,65]. Details on convergent validity were included for three instruments [60,61,65] (see Table 1). Importantly, there was no evidence of reliability, responsiveness or cross-cultural validity assessment for the usability assessment instruments referenced most often (i.e., SUS, IBM Computer Usability Satisfaction Questionnaire and USE Questionnaire).

Further, eight of the systematic reviews [13,44–46,48–50,54](67%) referred to usability assessment methods other than assessment scales. These included focus groups, heuristic evaluation, think-aloud protocols as well as other methods (see Supplementary Table S4).

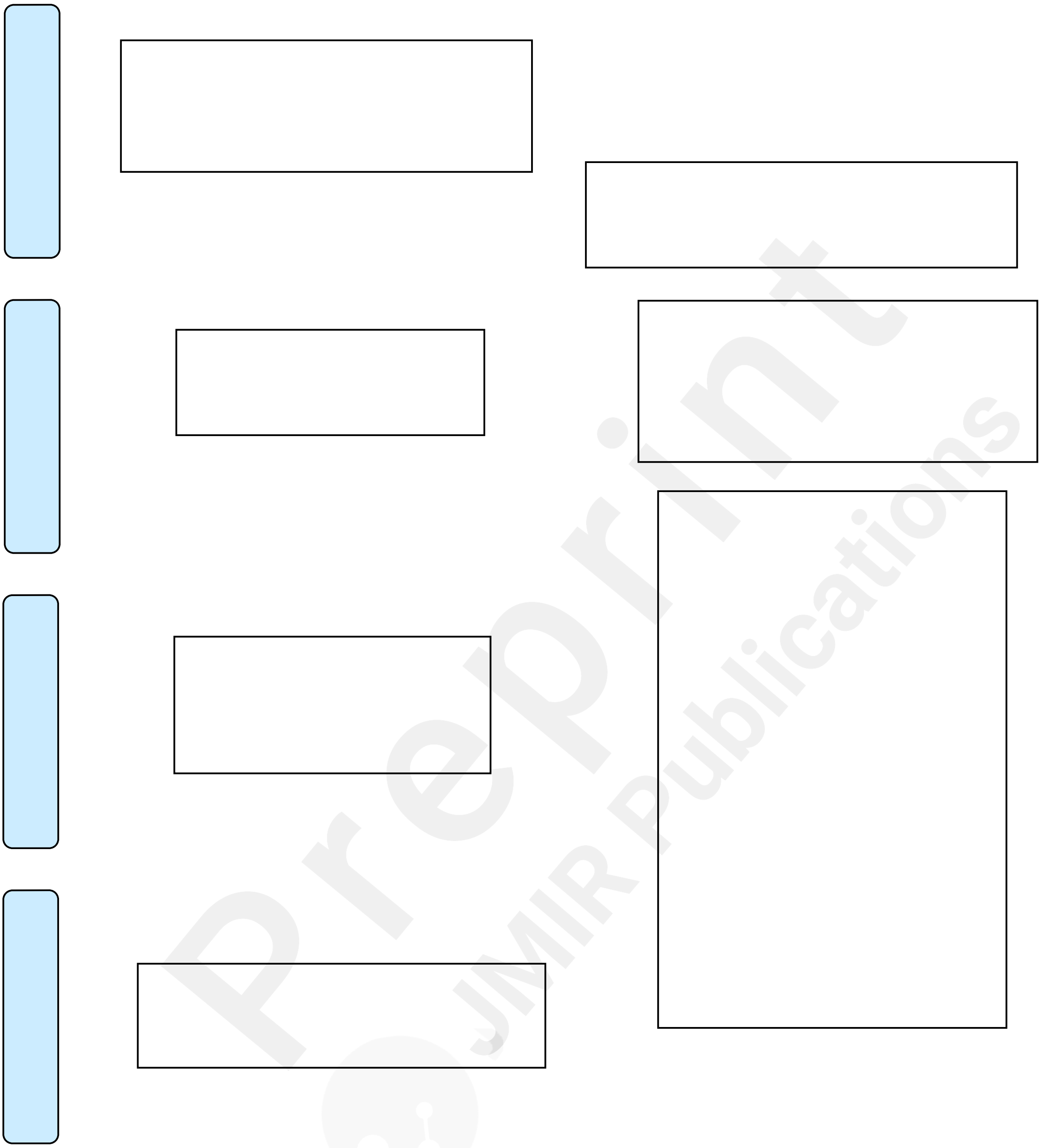
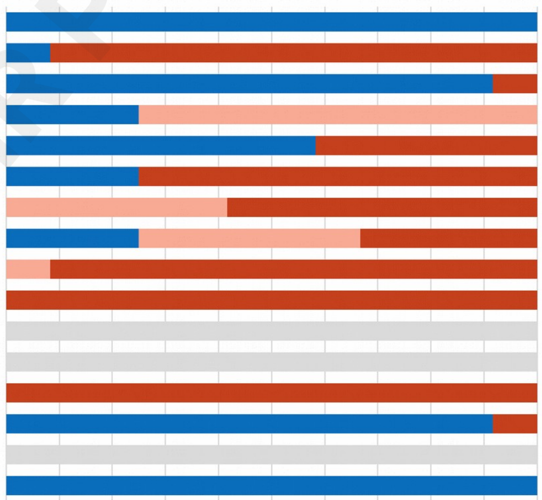


Figure 1: PRISMA flowchart

Quality assessment of the systematic reviews using AMSTAR2, revealed that eight articles (67%) [44–46,48–50,54,55] exhibited at least two critical weaknesses (see Figure 2), three systematic reviews (25%) [47,51,52] were affected by one critical weakness and one review (8%) [53] had only non-critical weaknesses. The most frequently unfulfilled assessment criteria included: the sources of funding enquiry for the included studies (AMSTAR Criterion 10); accounting for risk of bias when interpreting results (AMSTAR criterion 13); use of a satisfactory technique for assessing risk of bias (AMSTAR criterion 9); and inclusion of a review protocol (AMSTAR criterion 2) (see Supplementary Table S5).

Finally, visualisation of citation overlap for systematic reviews including primary studies utilising the Systems Usability Scale (SUS) showed minimal overlap with four (10%) of 41 primary studies included in two of the systematic reviews (see Supplementary Table S6). With the exception of the citation of the original publication of the SUS instrument [i.e., 57], all other references included in the overview were unique to one of the systematic reviews included.

AMSTAR2 domains For reasons of clarity, the original question has been reduced to some extent.	**“Yes”, “Partially Yes”, “No”, and “No Meta-analysis conducted” answers (%)** 0 20 40 60 80 100
1. Components of PICO? 2. Review protocol?* 3. Explanation of study design? 4. Comprehensive literature search strategy?* 5. Study selection in duplicate?* 6. Data extraction in duplicate? 7. List of excluded studies and justify the exclusions?* 8. Study characteristics? 9. Satisfactory technique for assessing risk of bias? 10. Sources of funding for included studies? 11. Appropriate methods for meta-analysis? 12. Assess potential impact of risk of bias on the results? 13. Account for risk of bias when interpreting/discussing? 14. Satisfactory explanation for/discussion of heterogeneity? 15. Publication bias assessed and discussed? 16. Potential sources of conflict of interest?*	

Note: * denotes critical criterion as determined for the present overview

Figure 2: Overview of methodological quality of reviews according to AMSTAR2

Discussion

The exponential growth of research evidence related to the effectiveness of mobile solutions for rehabilitation [66–69], and the proliferation of technological solutions that afford new modes of treatment delivery [70,71] underscore the critical need for high quality mHealth usability evaluation. Usability attributes such as efficiency, learnability and memorability [21] are particularly important to consider for mHealth users who may face challenges due to neurological compromise [72], age-related issues [23], or limited technology experience [13]. This umbrella review aimed to summarise usability assessment instruments for mHealth researchers, clinicians, and consumers to guide the development, assessment, and selection of high quality mHealth tools.

The review identified, firstly, significant diversity and common use of custom-made instruments when usability assessment instruments were employed to evaluate mHealth tools for rehabilitation. Secondly, there was a notable lack of theoretical grounding for selection of the assessment of usability. Thirdly, a scarcity of psychometric data for widely utilised instruments for mHealth usability assessment was evident in the systematic reviews included.

Heterogeneity of instruments, including non-standardised instruments

Regarding the first critical point, a wide range of different instruments for the assessment of usability was evident across the systematic reviews included. This range included adaptations of pre-existing usability assessment instruments for the context of mobile applications [73,74] as well as assessment instruments, such as the Mobile App Rating Scale (MARS) [62], specifically designed for usability assessment of mHealth tools. In addition, both completely custom-made instruments and hybrids [cf. 75] of pre-existing instruments with custom elements was prevalent in the mHealth usability literature.

While the use of hybrid assessment instruments and adaptations of pre-existing assessment instruments may increase flexibility and thereby possibly improve the experience for respondents, the fact that most studies are limited in sample size prevents validation of hybrid and adapted instruments [51]. Alternative approaches to increasing flexibility and improving respondent experience while ensuring psychometric integrity are needed instead. A good example of this may be seen in the creation of a hybrid version of the SUS with the inclusion of pictorial elements which increased respondent motivation [75]. Importantly, acceptable validity, consistency and sensitivity was also evidenced allowing future users of the hybrid measure to place greater trust in the quality of the data.

Theoretical underpinning

Secondly, and similar to what has been found to be the case for individual-level studies assessing the usability of specific mHealth tools [47], this review revealed that some systematic reviews examining the broader literature related to usability assessment lacked connection to theoretical models of usability. This observation resonates with previous criticisms of the quality of reviews of health-related mobile applications [8] as well as research exploring technology adoption in fields beyond mHealth [76]. The latter exposed a reliance on a wide array of theoretical models of technology adoption in the literature and in some cases several within one review. To address this, it has been suggested that generic models for different service categories (e.g., information, transaction) be developed [76]. A theoretically grounded, generic guide for mHealth usability assessment could similarly promote broader adoption and enhance comparison of usability across studies and use cases.

Psychometric properties and psychometric testing

Thirdly, systematic reviews included in our overview also reported significant limitations regarding the psychometric properties of pre-existing instruments. For example, the MARS tool, which has been put forward as an instrument for standardised use in mHealth usability assessment [51], lacks structural validity. Moreover, other constructs such as internal consistency and criterion validity have been documented as significant areas of future work for measuring the implementation of interventions [53], with usability assessment playing a significant role.

While consistent with previous research, the present umbrella review did not specifically search for psychometric evaluations of usability assessment instruments; instead, it relied on summaries of psychometric evaluations presented as part of the included systematic reviews. As a result, it is likely that psychometric evaluation of other instruments is available. For example, psychometric evaluation of the popular USE questionnaire [59] is available and, consistent with our observation, has been shown to be affected by a lack of reliability and validity [77]. Furthermore, outside of the academic literature, there is a still greater portion of mHealth solutions on the market that likely will not have undergone empirical evaluation of usability.

While some of the acceptable psychometric information was referenced for the System Usability Scale (SUS) [e.g., 78], both the IBM Usability Satisfaction Questionnaire and the USE questionnaire appear to lack reliability assessment. Reliability, or the freedom of measurement error [79], may be regarded as crucial with regard to any metrics which are gathered following, rather than during, a user's interaction with an application. The inability to separate true change in users' estimate of the usability of mHealth tools from random variation, or measurement error, originating from recall bias

[9,80,cf. 81], for example, means that mHealth tool iterations [82] are unable to be evaluated appropriately.

Moreover, the widespread use of custom-made and hybrid assessment instruments leads to the loss of the original instrument's integrity and compromises its already-documented psychometric strengths [83]. Consequently, establishing the validity of results from individual usability investigations becomes challenging and comparison across studies is difficult. Hence, there is an urgent need to assess the accuracy and appropriateness [84] of individual usability assessment instruments to capitalise on the promise of mHealth tools in rehabilitation [5,6].

Another important psychometric aspect of usability assessment instruments that the systematic reviews included in the present overview highlight as missing from the published literature is responsiveness. mHealth development usually involves iterative design and testing cycles [e.g., 30,82] with associated formative and summative usability evaluation [45]. Across the life of mHealth development, iterative cycles are likely to span different stages of development and be undertaken in different clinical contexts [54,85]. Integrating usability assessment into this process requires instruments that are generic enough to capture user responses to a wide variety of mHealth strategies but also fine-grained enough to possess sufficient responsiveness [79].

Finally with regard to the argument of lacking psychometric assessment, none of the pre-existing mHealth usability assessment instruments referenced as part of the literature included in this overview appear to have been informed by a breadth of cultural perspective or undergone cross-cultural validity testing. Given the global potential of mHealth to address inequities in access to and outcomes from rehabilitation [5,6], it is particularly important to establish cross-cultural validity of the usability assessment instruments employed in mHealth development. In addition, with the pervasiveness of technology, there is a certain element of unpredictability of the context in which mHealth tools will be trialled and used 'in the wild' [86,87]. For that reason, an alternative argument could be made for innovative, culturally responsive methodology for mHealth tool design including usability testing [88]. A key difference in such attempts is user participation at multiple stages of development and responsiveness to expanding the stages of development as guided by stakeholders. This process likely includes constant negotiation and may be resource heavy but is arguably needed if the aim is to create mHealth solutions impacting indigenous outcomes, for example [89,90].

Considering the identified issues, including lack of theoretical grounding, common use of custom-made assessment instruments and the scarcity of psychometric data for widely utilised mHealth usability assessment instruments, multi-method usability assessment appears paramount. This is consistent with recommendations made by a number of research groups [44,55,85,91] and reinforces the argument often advanced in favour of Ecological Momentary Assessment approaches, which are recognised for their advantage over retrospective assessment [80]. It is therefore proposed that standards be developed which specify the timepoints in the mHealth life cycle at which usability assessment is completed, with an emphasis on what methods to use. Moreover, these standards should mandate that individual assessment instruments are grounded in a theoretical framework and possess a minimum threshold for psychometric properties [e.g., 53,92].

Recommendations

The establishment of a universal usability scoring system or algorithm would further facilitate the integration of these assessments into an overall framework [93]. It has been observed that, at present, less than half of existing evaluation frameworks include such a scoring system, but that such systems could support funding decisions [29] and advance the vision of prescribable mHealth applications [10]. While technological advancement often outpaces academic enquiry necessitating new approaches to mHealth evaluation frameworks [94], usability factors are enduring [16] and investing resources into establishing standards will therefore be valuable.

Limitations

In the context of an area of practice where the lines between commercial and academic work are blurred and usability assessment constitutes a common practice in the global commercial environment [95], the present overview is limited in only including English language systematic reviews published within the academic literature indexed in the databases included. Furthermore, the quality of the included systematic reviews was found to be limited and the fit of the AMSTAR2 tool with methodological papers is not perfect. However, AMSTAR2 could be argued to be more detailed compared to instruments developed for umbrella reviews specifically [31] and, in line with the AMSTAR2 recommendations [39], the authors modified the list of critical criteria to reflect the specific aim of the overview. Finally, with regard to the review's methodology, two limitations are of note. First, while the search syntax for the present overview included the keyword 'physical exercise', for pragmatic reasons, no validation step was included to confirm that all mHealth tools examined as part of the primary studies included within the systematic reviews included a physical exercise component. Regardless, the observations presented here are valid for mHealth tools for rehabilitation overall and provide valuable guidance to developers, researchers and clinicians. Second, for practical reasons, data selection could only be performed by the primary author [SH] with a subset of articles being screened by a second author [VS]. However, agreement on study selection was high (>80%), supporting the quality of the review.

Conclusions

There is considerable variety in approaches to, and instruments for the assessment of usability in mHealth for rehabilitation, many of which lack theoretical foundation. Clinicians are therefore advised to critically evaluate mHealth literature and solutions, paying particular attention to the population(s) in which usability testing was performed and the specific usability assessment instruments employed. Future research efforts should be focused on producing high quality systematic reviews and psychometric evaluations of usability assessment instruments. A collaborative effort between researchers, designers, and developers is essential to establish mHealth tool development standards. These standards should stress the incorporation of usability assessment instruments underpinned by a robust theoretical base. The present overview represents a valuable reference tool in this endeavour. Inclusion of multi-method usability assessment within the wider mHealth development cycle could also be part of these standards, which will ensure that we can capitalise on the widely heralded promise of mHealth to promote access to and outcomes from rehabilitation.

Acknowledgements

The wider team of researcher and clinicians at the AUT Research Innovation Centre for workshop and input; Exsurgo for valuable conversations on usability from a commercial perspective

Conflict of Interest

The authors report there are no competing interests to declare.

References

1. Istepanian RSH, Alanzi T. Mobile health (m-health): Evidence- based progress or scientific retrogression. In: Feng DD, editor. Biomedical Information Technology 2nd ed Elsevier Science & Technology; 2022. p. 717–734.
2. Istepanian RSH, Laxminarayan S, Pattichis CS. M-Health: Emerging mobile health systems. In: Mecheli-Tzanakou E, editor. Topics in Biomedical Engineering: International book series Springer; 2006. p. 1773–1787. doi: 10.4018/978-1-60566-050-9.ch136
3. Cao J, Lim Y, Sengoku S, Guo X, Kodama K. Exploring the shift in international trends in mobile health research from 2000 to 2020: Bibliometric analysis. JMIR Mhealth Uhealth 2021;9(9):1–17. PMID:34494968
4. mHealth App Economics 2017- Current status and future trends in mobile health. 2017. Available from: www.research2guidance.com
5. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, Ruggiero KJ. mHealth: A mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427–436. PMID:23918764
6. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: A systematic review. Telemedicine and e-Health 2014;20(1):75–82. PMID:24205809
7. Chu JT, Wang MP, Shen C, Viswanath K, Lam TH, Chan SSC. How, When and Why People Seek Health Information Online: Qualitative Study in Hong Kong. Interact J Med Res 2017;6(2):e24. doi: 10.2196/ijmr.7000
8. 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 Elsevier; 2016;51(6):1051–1059. PMID:27659122
9. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. Arch Phys Med Rehabil United States; 2019 Jan;100(1):115–127. PMID:30171827
10. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. NPJ Digit Med Springer US; 2018;1(1):1–12. PMID:31304297
11. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. Int J Med Inform Elsevier; 2019;126(March):95–104. PMID:31029270
12. Fiedler J, Eckert T, Wunsch K, Woll A. Key facets to build up eHealth and mHealth interventions to enhance physical activity, sedentary behavior and nutrition in healthy subjects - an umbrella review. BMC Public Health 2020 Oct;20(1):1605. PMID:33097013
13. 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. PMID:25600193
14. Harrison R, Flood D, Duce D. Usability of mobile applications: Literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1. doi: 10.1186/2194-0827-1-1
15. Weichbroth P. Usability of mobile applications: A systematic literature study. IEEE Access IEEE; 2020;8:55563–55577. doi: 10.1109/ACCESS.2020.2981892
16. Albert B, Tullis T. Introduction. Measuring the user experience: Collecting, analyzing, and presenting usability metrics Elsevier Science & Technology; 2013. p. 1–14. doi: http://dx.doi.org/10.1016/B978-0-12-415781-1.00001-7
17. Nielsen J. Defining usability. User experience re-mastered: Your guide to getting the right design 2019. p. 1–22.
18. Part 11: Usability: Definitions and concepts. Ergonomics of human-system interaction. 2022. Available from: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en
19. Bevan N, Carter J, Harker S. ISO 9241-11 revised: What have we learnt about usability since 1998?

Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2015;9169:143–151. doi: 10.1007/978-3-319-20901-2_13
20. ISO/IEC 25010. ISO 25000 software and data quality. 2022. Available from: https://iso25000.com/index.php/en/iso-25000-standards/iso-25010
21. Nielsen J. Usability engineering. 1st ed. Morgan Kaufmann; 1994. ISBN: 0125184069
22. Alturki R, Gay V. Usability testing of fitness mobile application: Methodology and quantitative results. CS & IT-CSCP 2017. p. 97–114. doi: 10.5121/csit.2017.71108
23. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). Int J Med Inform Elsevier; 2018;114(September 2017):66–75. PMID:29673606
24. National Health Service. Digital Assessment Questionnaire V2.1. NHS Digital. 2018.
25. MHRA. Guidance on applying human factors and usability engineering of medical devices including drug-device combination products in Great Britain. 2021.
26. ORCHA. 2022. Available from: https://orchahealth.com/about-us/unique-approach/
27. New Zealand Ministry of Health. Health Navigator app library. 2022. Available from: www.healthnavigator.org.nz/app-library/
28. New Zealand Ministry of Health. Health applications assessment guidance. Wellington, New Zealand; 2017.
29. Moshi MR, Tooher R, Merlin T. Suitability of current evaluation frameworks for use in the health technology assessment of mobile medical applications: A systematic review. Int J Technol Assess Health Care 2018;34(5):464–475. PMID:30201060
30. Bonten TN, Rauwerdink A, Wyatt JC, Kasteleyn MJ, Witkamp L, Riper H, van Gemert-Pijnen LJEWC, Cresswell K, Sheikh A, Schijven MP, Chavannes NH. Online guide for electronic health evaluation approaches: Systematic scoping review and concept mapping study. J Med Internet Res 2020;22(8):1–22. PMID:32784173
31. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. Int J Evid Based Healthc 2015;13(3):132–140. PMID:26360830
32. Pollock M, Fernandes RM, Pieper D, Tricco AC, Gates M, Gates A, Hartling L. Preferred Reporting Items for Overviews of Reviews (PRIOR): A protocol for development of a reporting guideline for overviews of reviews of healthcare interventions. Syst Rev Systematic Reviews; 2019;8(1):1–9. PMID:31870434
33. Pollock M, Fernandes RM, Becker LA, Pieper D, Hartling L. Chapter V: Overviews of reviews. Cochrane Handbook for Systematic Reviews of Interventions 2021; Available from: https://training.cochrane.org/handbook/current/chapter-v#_Ref524711112
34. Hunt H, Pollock A, Campbell P, Estcourt L, Brunton G. An introduction to overviews of reviews: Planning a relevant research question and objective for an overview. Syst Rev Systematic Reviews; 2018;7(1):1–9. PMID:29490699
35. Silva C, Zamboni A, Hernandes E, Di Thomazzo A, Belgamo A, Fabbri S. LaPES. State of the Art through Systematic Review. 2022.
36. Davis TL, Diclemente R, Prietula M. Taking mHealth forward: Examining the core characteristics. JMIR Mhealth Uhealth JMIR Publications Inc.; 2016;4(3). PMID:27511612
37. Chandler J, Cumpston M, Thomas J, Higgins JPT, Deeks JJ, Clarke MJ. Chapter I: Introduction. In: Higgins JPT, Green S, Chandler J, Cumpston M, Li T, Page M, Welch V, editors. Cochrane handbook for systematic reviews of interventions Version 6. 2022.
38. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. BMJ 2009;339. PMID:19622552
39. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V,

Kristjansson E, Henry DA. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ (Online) 2017;358(j4008). PMID:28935701
40. Gates M, Gates A, Guitard S, Pollock M, Hartling L. Guidance for overviews of reviews continues to accumulate, but important challenges remain: A scoping review. Syst Rev Systematic Reviews; 2020;9(1):1–19. PMID:33148319
41. Bougioukas KI, Vounzoulaki E, Mantsiou CD, Savvides ED, Karakosta C, Diakonidis T, Tsapas A, Haidich AB. Methods for depicting overlap in overviews of systematic reviews: An introduction to static tabular and graphical displays. J Clin Epidemiol Elsevier Inc; 2021;132:34–45. PMID:33309886
42. Hennessy EA, Johnson BT. Examining overlap of included studies in meta-reviews: Guidance for using the corrected covered area index. Res Synth Methods 2020;11(1):134–145. PMID:31823513
43. Lewis JR. The System Usability Scale: Past, Present, and Future. Int J Hum Comput Interact Taylor & Francis; 2018 Jul 3;34(7):577–590. doi: 10.1080/10447318.2018.1455307
44. Georgsson M. A review of usability methods used in the evaluation of mobile health applications for diabetes. Stud Health Technol Inform 2020;273:228–233. PMID:33087617
45. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability evaluations of mobile mental health technologies: Systematic review. J Med Internet Res 2020;22(1):1–19. PMID:31904579
46. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: A review of measurement, reporting, and validity. Psychiatric Services 2019;70(7):538–544. PMID:30914003
47. 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 Taylor & Francis; 2021;16(7):712–721. PMID:31910687
48. Vera F, Noël R, Taramasco C. Standards, processes and instruments for assessing usability of health mobile apps: A systematic literature review. Stud Health Technol Inform 2019;264:1797–1798. PMID:31438349
49. Saeed N, Manzoor M, Khosravi P. An exploration of usability issues in telecare monitoring systems and possible solutions: A systematic literature review. Disabil Rehabil Assist Technol 2020;15(3):271–281. PMID:30794009
50. Nouri R, Kalhori SRN, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: A systematic review. Journal of the American Medical Informatics Association 2018;25(8):1089–1098. PMID:29788283
51. Muro-Culebras A, Escriche-Escuder A, Martin-Martin J, Roldán-Jiménez C, De-Torres I, uiz-Muñoz M, Gonzalez-Sanchez M, Mayoral-Cleries F, Biró A, Tang W, Nikolova B, Salvatore A, Cuesta-Vargas AI. Tools for evaluating the content, efficacy, and usability of mobile health apps according to the consensus-based standards for the selection of health measurement instruments: Systematic review. JMIR Mhealth Uhealth 2021;9(12). PMID:34855618
52. Wakefield BJ, Turvey CL, Nazi KM, Holman JE, Hogan TP, Shimada SL, Kennedy DR. Psychometric properties of patient-facing ehealth evaluation measures: Systematic review and analysis. J Med Internet Res 2017;19(10):1–13. PMID:29021128
53. Kien C, Schultes MT, Szelag M, Schoberberger R, Gartlehner G. German language questionnaires for assessing implementation constructs and outcomes of psychosocial and health-related interventions: A systematic review. Implementation Science Implementation Science; 2018;13(1):1–16. PMID:30541590
54. Niknejad N, Ismail W, Bahari M, Nazari B. Understanding telerehabilitation technology to evaluate stakeholders' adoption of telerehabilitation services: A systematic literature review and directions for further research. Arch Phys Med Rehabil Elsevier Inc.; 2021;102(7):1390–1403. PMID:33484693
55. 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. PMID:25600193
56. Mental Health Commission of Canada (MHCC), Canadian Institutes of Health Research (CIHR).

Mental health apps: How to make an informed choice. 2016.
57. Brooke J. SUS: A “quick and dirty” usability scale. Usability Evaluation In Industry CRC Press; 1996;207–212. doi: 10.1201/9781498710411-35
58. Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. Int J Hum Comput Interact Taylor & Francis Group; 2009;7(1):57–78. doi: 10.1080/10447319509526110
59. Lund A. Measuring usability with the USE questionnaire. Usability Interface 2001;8(2):3–6.
60. Yen PY, Wantland D, Bakken S. Development of a customizable Health IT Usability Evaluation Scale. AMIA Annual Symposium Proceedings American Medical Informatics Association; 2010;2010:917. PMID:21347112
61. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q Management Information Systems Research Center; 1989;13(3):319–339. doi: 10.2307/249008
62. 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 JMIR Publications; 2016 Jun 1;4(2):e72. PMID:27287964
63. 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 JMIR Publications; 2015;3(1):e27. PMID:25760773
64. Bakken S, Grullon-Figueroa L, Izquierdo R, Lee NJ, Morin P, Palmas W, Teresi J, Weinstock RS, Shea S, Starren J. Development, validation, and use of English and Spanish versions of the telemedicine satisfaction and usefulness questionnaire. J Am Med Inform Assoc J Am Med Inform Assoc; 2006;13(6):660–667. PMID:16929036
65. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mhealth app usability questionnaire (MAUQ): Development and validation study. JMIR Mhealth Uhealth 2019;7(4):1–15. doi: 10.2196/11500
66. Suso-Martí L, La Touche R, Herranz-Gómez A, Angulo-Díaz-Parreño S, Paris-Alemany A, Cuenca-Martínez F. Effectiveness of telerehabilitation in physical therapist practice: An umbrella and mapping review with meta-meta-analysis. Phys Ther 2021 May;101(5). PMID:33611598
67. Marcolino MS, Oliveira JAQ, D’Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mHealth interventions: Systematic review of systematic reviews. JMIR Mhealth Uhealth 2018;6(1). doi: 10.2196/mhealth.8873
68. Mönninghoff A, Kramer JN, Hess AJ, Ismailova K, Teepe GW, Tudor Car L, Müller-Riemenschneider F, Kowatsch T. Long-term effectiveness of mHealth physical activity interventions: Systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Apr;23(4):e26699. PMID:33811021
69. Elavsky S, Knapova L, Klocek A, Smahel D. Mobile health interventions for physical activity, sedentary behavior, and sleep in adults aged 50 years and older: A systematic literature review. J Aging Phys Act United States; 2019 Aug;27(4):565–593. PMID:30507266
70. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R. mHealth technologies to influence physical activity and sedentary behaviors: Behavior change techniques, systematic review and meta-analysis of randomized controlled trials. Ann Behav Med England; 2017 Apr;51(2):226–239. PMID:27757789
71. Koumpouros Y, Georgoulas A. A systematic review of mHealth funded R&D activities in EU: Trends, technologies and obstacles. Inform Health Soc Care Taylor & Francis; 2020;45(2):168–187. PMID:31743060
72. Rabinowitz AR, Juengst SB. Introduction to topical issue on mHealth for brain injury rehabilitation. Journal of Head Trauma Rehabilitation 2022;37(3):131–133. doi: 10.1097/htr.0000000000000794
73. 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 JMIR Publications;

2015;4(3):e18. doi: 10.2196/IJMR.4323
74. Van Singer M, Chatton A, Khazaal Y. Quality of smartphone apps related to panic disorder. Front Psychiatry 2015;6(JUL):1–7. doi: 10.3389/fpsyt.2015.00096
75. Baumgartner J, Ruettgers N, Hasler A, Sonderegger A, Sauer J. Questionnaire experience and the hybrid System Usability Scale: Using a novel concept to evaluate a new instrument. International Journal of Human Computer Studies Elsevier Ltd; 2021;147:102575. doi: 10.1016/j.ijhcs.2020.102575
76. Ovčjak B, Heričko M, Polančič G. Factors impacting the acceptance of mobile data services - A systematic literature review. Comput Human Behav 2015;53:24–47. doi: 10.1016/j.chb.2015.06.013
77. Gao M, Kortum P, Oswald F. Psychometric evaluation of the USE (Usefulness, Satisfaction, and Ease of use) questionnaire for reliability and validity. Proceedings of the Human Factors and Ergonomics Society 2018;3:1414–1418. doi: 10.1177/1541931218621322
78. Peres SC, Pham T, Phillips R. Validation of the System Usability Scale (SUS): SUS in the wild. Proceedings of the Human Factors and Ergonomics Society 2013. p. 192–196. doi: 10.1177/1541931213571043
79. McDowell I. The theoretical and technical foundations of health measurement. Measuring Health. 2009. doi: 10.1093/acprof:oso/9780195165678.003.0002ISBN:9780195165678
80. Demers M, Winstein CJ. A perspective on the use of ecological momentary assessment and intervention to promote stroke recovery and rehabilitation. Top Stroke Rehabil Taylor & Francis; 2021;28(8):594–605. PMID:33272137
81. Boyd K, Bond R, Vertesi A, Dogan H, Magee J. How people judge the usability of a desktop Graphic User Interface at different time points: Is there evidence for memory decay, recall bias or temporal bias? Interact Comput 2019;31(2):221–230. doi: 10.1093/iwc/iwz019
82. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: Mixed-methods approach. JMIR Mhealth Uhealth 2019;7(4). doi: 10.2196/11656
83. Switzer GE, Wisniewski SR, Belle SH, Dew MA, Schultz R. Selecting, developing, and evaluating research instruments. Soc Psychiatry Psychiatr Epidemiol 1999;34(8):399–409. PMID:10501709
84. Hughes DJ. Psychometric validity: Establishing the accuracy and appropriateness of psychometric measures. The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development 2018. p. 751–779. doi: 10.1002/9781118489772.ch24
85. Zein S, Salleh N, Grundy J. A systematic mapping study of mobile application testing techniques. Journal of Systems and Software Elsevier Inc.; 2016;117:334–356. doi: 10.1016/j.jss.2016.03.065
86. Brown B, Reeves S, Sherwood S. Into the wild: Challenges and opportunities for field trial methods. Conference on Human Factors in Computing Systems - Proceedings 2011;1657–1666. doi: 10.1145/1978942.1979185
87. Kjeldskov J, Stage J. New techniques for usability evaluation of mobile systems. International Journal of Human Computer Studies 2004;60(5–6):599–620. doi: 10.1016/j.ijhcs.2003.11.001
88. Rolleston AK, Bowen J, Hinze A, Korohina E, Matamua R. Collaboration in research: weaving Kaupapa Māori and computer science. AlterNative 2021;17(4):469–479. doi: 10.1177/11771801211043164
89. Stowell E, Lyson MC, Saksono H, Wurth RC, Jimison H, Pavel M, Parker AG. Designing and evaluating mHealth interventions for vulnerable populations: A systematic review. Conference on Human Factors in Computing Systems - Proceedings Association for Computing Machinery; 2018. doi: 10.1145/3173574.3173589
90. McMillan B, Hickey E, Patel MG, Mitchell C. Quality assessment of a sample of mobile app-based health behavior change interventions using a tool based on the National Institute of Health and Care Excellence behavior change guidance. Patient Educ Couns Patient Educ Couns; 2016;99(3):429–435. PMID:26607787
91. Bernhaupt R, Mihalic K, Obrist M. Usability evaluation methods for mobile applications. Handbook of Research on User Interface Design and Evaluation for Mobile Technology 2011;745–758. doi:

10.4018/978-1-59904-871-0.ch044
92. Lewis CC, Mettert KD, Dorsey CN, Martinez RG, Weiner BJ, Nolen E, Stanick C, Halko H, Powell BJ. An updated protocol for a systematic review of implementation-related measures. Syst Rev Systematic Reviews; 2018;7(1):66. PMID:29695295
93. Oyebode O, Alqahtani F, Orji R. Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. IEEE Access 2020;8:111141–111158. doi: 10.1109/ACCESS.2020.3002176
94. Quintana Y, Torous J. A framework for evaluation of mobile apps for youth mental health. Homewood Research Institute Report. 2020.
95. Hosseiniravandi M, Kahlaee AH, Karim H, Ghamkhar L, Safdari R. Home-based telerehabilitation software systems for remote supervising: a systematic review. Int J Technol Assess Health Care England; 2020 Apr;36(2):113–125. PMID:32151291
96. Lewis JR. An After-Scenario Questionnaire for usability studies: Psychometric evaluation over three trials. ACM SIGCHI Bulletin ACM PUB27 New York, NY, USA; 1991;23(4):79. doi: 10.1145/126729.1056077
97. McNiel P, McArthur EC. Evaluating health mobile apps: Information literacy in undergraduate and graduate nursing courses. Journal of Nursing Education Slack Incorporated; 2016 Aug 1;55(8):480. PMID:27459438
98. Huis in't Veld RMHA, Kosterink SM, Barbe T, Lindegård A, Marecek T, Vollenbroek-Hutten MMR. Relation between patient satisfaction, compliance and the clinical benefit of a teletreatment application for chronic pain. J Telemed Telecare SAGE PublicationsSage UK: London, England; 2010;16(6):322–328. PMID:20798426
99. Baumel A, Faber K, Mathur N, Kane JM, Muench F, Research P. Enlight: A comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. J Med Internet Res 2017;19(3). doi: 10.2196/jmir.7270
100. Brown W, 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 Academic Press; 2013 Dec 1;46(6):1080–1087. PMID:23973872
101. Gediga G, Hamborg K-C, Düntsch I, È Nther Gediga G, Hamborg K, È Ntsch I DU. The IsoMetrics usability inventory: An operationalization of ISO 9241-10 supporting summative and formative evaluation of software systems. Behaviour and Information Technology Taylor & Francis Group; 2010;18(3):151–164. doi: 10.1080/014492999119057
102. Grau I, Kostov B, Gallego JA, Grajales F, Fernández-Luque L, Sisó-Almirall A. [Assessment method for mobile health applications in Spanish: The iSYScore index]. Semergen Semergen; 2016;42(8):575–583. PMID:26879598
103. Jin M, Kim J. Development and evaluation of an evaluation tool for healthcare smartphone applications. Telemed J E Health Telemed J E Health; 2015;21(10):831–837. PMID:26431261
104. Hart SG. NASA-task load index (NASA-TLX); 20 years later. Proceedings of the Human Factors and Ergonomics Society 2006;904–908. doi: 10.1177/154193120605000909
105. Price M, Sawyer T, Harris M, Skalka C. Usability evaluation of a mobile monitoring system to assess symptoms after a traumatic injury: A mixed-methods study. Journal of Medical Internet Research: Mental health JMIR Mental Health; 2016;3(1):e5023. doi: 10.2196/MENTAL.5023
106. Lewis JR. Psychometric evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. Proceedings of the Human Factors Society 1992. p. 1259–1263.
107. Loy JS, Ali EE, Yap KYL. Quality assessment of medical apps that target medication-related problems. J Manag Care Spec Pharm J Manag Care Spec Pharm; 2016;22(10):1124–1140. PMID:27668561
108. 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 Springer; 2013;37(5):1–8. PMID:24014255

109. Chin JP, Diehl VA, Norman KL. Development of an instrument measuring user satisfaction of the human-computer interface. Conference on Human Factors in Computing Systems - Proceedings Association for Computing Machinery; 1988 May 1;Part F130202:213–218. doi: 10.1145/57167.57203
110. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the internet: Caveant lector et viewor—Let the reader and viewer beware. JAMA American Medical Association; 1997;277(15):1244–1245. doi: 10.1001/JAMA.1997.03540390074039
111. Kirakowski J, Corbett M. SUMI: the Software Usability Measurement Inventory. British Journal of Educational Technology 1993;24(3):210–212. doi: 10.1111/J.1467-8535.1993.TB00076.X
112. Parmanto B, Pulantara IW, Schutte JL, Saptono A, McCue MP. An integrated telehealth system for remote administration of an adult autism assessment. Telemedicine and e-Health Mary Ann Liebert Inc.; 2013;19(2):88–94. PMID:23230821

AUS manuscript FINAL R1

Tables

Table 1: Overview of usability assessment scales identified by reviews included within the present overview

				Psychometric properties as identified by systematic reviews included in the present overview						
	Reference	Systematic Review(s) identifying scale	Count	Internal consistency (Cronbach α)	Reliability (Intraclass Corr-elation)	Content validity	Structural validity	Cross-cultural validity	Criterion, convergent, concurrent, discrim-inant validity	Responsive-ness
app adaptation Abbott's scale	[74]	Nouri et al. (2018)	1	NR	NR	NR	NR	NR	NR	NR
After Scenario Questionnaire	[96]	Inal et al. (2020)	1	NR	NR	NR	NR	NR	NR	NR
app adaptation Brief DISCERN	[74]	Nouri et al. (2018)	1	NR	NR	NR	NR	NR	NR	NR
app adaptation CRAAP checklist	[97]	Nouri et al. (2018)	1	NR	NR	NR	NR	NR	NR	NR
Ease of Use and Usefulness Scale (EUUS)	[98]	Kien et al. (2018)	1	NR	NR	NR	NR	NR	NR	NR
Enlight	[99]	Azad-Khaneghah et al. (2021)	1	NR	NR	NR	NR	NR	NR	NR
Health Information Technology Usability Evaluation Scale (Health-ITUES)	[60]	Azad-Khaneghah et al. (2021) Muro-Culebras et al. (2021)	2	.85-.92	No	Expert panel & factor analysis	Exploratory & Confirm-atory Factor Analysis	No	Correlation with the Poststudy System usability question-naire	Statistically significant difference was demon-strated with intervention group
	Reference	Systematic Review(s) identifying scale	Count	Internal consistency	Reliability (Intraclas	Content validity	Structural validity	Cross-cultural validity	Criterion, convergent ,	Responsive-ness

Records identified through database searching (n = 1479)

Duplicates removed (n = 104)

Records screened (n = 1375)

Records excluded during title and abstract screening (n = 1348)

Full text articles assessed for eligibility (n = 27)

Full text articles excluded, with reasons (n = 15)
- No review of usability assessment measures (n=12)
- Insufficient detail included on usability assessment measures (e.g., binary information only) (n=1)
- No literature review included (n=1)
- Examined non-health mobile service

Studies included in the overview (n = 12)

				(Cronbach α)	s Corr-elation)				concurrent, discrim-inant validity	
Health IT Usability Evaluation Model (Health-ITUEM)	[100]	Nouri et al. (2018) Vera et al. (2019)	2	NR	NR	NR	NR	NR	NR	NR
app adaptation Health-Related Website Evaluation Form (HRWEF)	[73]	Nouri et al. (2018)	1	NR	NR	NR	NR	NR	NR	NR
app adaptation Health On the Net (HON) code	[74]	Nouri et al. (2018)	1	NR	NR	NR	NR	NR	NR	NR
IBM Computer Usability Satisfaction Questionnaire	[58]	Azad-Khaneghah et al. (2021) Georgsson (2020) Ng et al. (2019) Wakefield et al. (2017) Zapata et al. (2015)	5	.89	No	Expert panel	No	NR	No	No
ISOMETRIC	[101]	Azad-Khaneghah et al. (2021)	1	NR	NR	NR	NR	NR	NR	NR
iSYScore index	[102]	Muro-Culebras et al. (2021)	1	No	No	Expert panel	No	NR	No	No
app adaptation Kim Model	[103]	Nouri et al. (2018)	1	NR	NR	NR	NR	NR	NR	NR

	Reference	Systematic Review(s) identifying scale	Count	Internal consistency (Cronbach α)	Reliability (Intraclass Corr-elation)	Content validity	Structural validity	Cross-cultural validity	Criterion, convergent, concurrent, discrim-inant validity	Responsive-ness
Measurement Scales for Perceived Usefulness and Perceived Ease of Use	[61]	Muro-Culebras et al. (2021)	1	.97 (usefulness) .91 (ease of use)	No	Focus group	Exploratory Factor Analysis	No	Convergent & discriminant validity	No
Mobile App Rating Scale (MARS)	[63]	Muro-Culebras et al. (2021) Nouri et al. (2018) Vera et al. (2019)	3	.90	.79	Expert panel	No	No	No	No
Mobile App Rating Scale (user version) (uMARS)	[62]	Muro-Culebras et al. (2021) Nouri et al. (2018)	2	.90	.66 (1-2 months) .70 (3 months)	Expert panel & focus groups	No	No	No	No
NASA Task Load Index (TLX)	[104]	Zapata et al., (2015)	1	NR	NR	NR	NR	NR	NR	NR
NICE guidelines tool	[90]	Azad-Khaneghah et al. (2021)	1	NR	NR	NR	NR	NR	NR	NR
Perceived Useful and Ease of Use Questionnaire (PUEU)	[105]	Azad-Khaneghah et al. (2021) Inal et al. (2020)	2	NR	NR	NR	NR	NR	NR	NR
Post-Study System Usability Scale (PSSUS)/PSSUQ	[106]	Inal et al. (2020) Niknejad et al. (2021) Vera et al. (2019)	3	NR	NR	NR	NR	NR	NR	NR

AUS manuscript FINAL R1

	Reference	Systematic Review(s) identifying scale	Count	Internal consistency (Cronbach α)	Reliability (Intraclass Corr-elation)	Content validity	Structural validity	Cross-cultural validity	Criterion, convergent, concurrent, discrim-inant validity	Responsive-ness
Quality Assessment tool for Evaluating Medical Apps (QAEM)	[107]	Azad-Khaneghah et al. (2021)	1	NR	NR	NR	NR	NR	NR	NR
Quality of Experience (QOE)	[108]	Azad-Khaneghah et al. (2021) Nouri et al. (2018)	2	NR	NR	NR	NR	NR	NR	NR
Questionnaire for User Interaction Satisfaction 7.0 (QUIS)	[109]	Georgsson (2020) Saeed et al. (2020)	2	NR	NR	NR	NR	NR	NR	NR
app adaptation Silberg score	[110]	Azad-Khaneghah et al. (2021) Nouri et al. (2018)	2	NR	NR	NR	NR	NR	NR	NR

AUS manuscript FINAL R1

	Reference	Systematic Review(s) identifying scale	Count	Internal consistency (Cronbach α)	Reliability (Intraclass Corr-elation)	Content validity	Structural validity	Cross-cultural validity	Criterion, convergent, concurrent, discrim-inant validity	Responsive-ness
Software Usability Measurement Inventory (SUMI)	[111]	Azad-Khaneghah et al. (2021)	1	NR	NR	NR	NR	NR	NR	NR
System Usability Scale (SUS)	[57]	Azad-Khaneghah et al. (2021) Georgsson (2020) Inal et al. (2020) Muro-Culebras et al. (2021) Ng et al. (2019) Niknejad et al. (2021) Nouri et al. (2018) Vera et al. (2019) Wakefield et al. (2017) Zapata et al., (2015)	10	.911	No	Focus group	Exploratory & Confirm-atory Factor Analysis	No	No	No
Telehealth Usability Questionnaire (TUQ)	[112]	Georgsson (2020) Inal et al. (2020) Niknejad et al. (2021)	3	NR	NR	NR	NR	NR	NR	NR
Telemedicine Satisfaction and Usefulness Questionnaire (TSUQ)	[64]	Wakefield et al. (2017)	1	.96 (Video visits) .92 (Use & impact)	No	Expert panel	Exploratory Factor Analysis	No	Significant discriminant validity (Hispanics vs non-Hispanics)	No

AUS manuscript FINAL R1

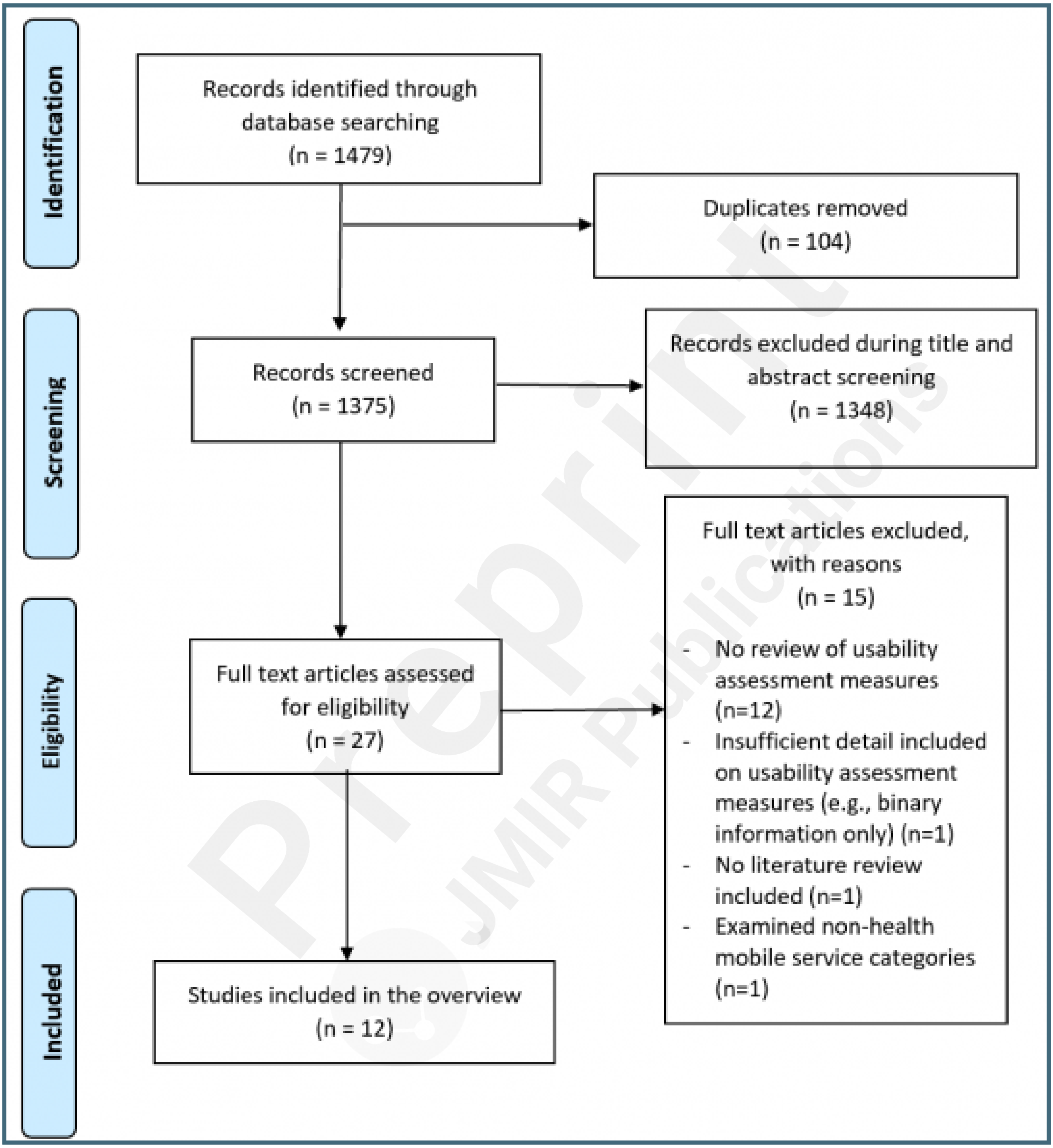
	Reference	Systematic Review(s) identifying scale	Count	Internal consistency (Cronbach α)	Reliability (Intraclass Corr-elation)	Content validity	Structural validity	Cross-cultural validity	Criterion, convergent, concurrent, discrim-inant validity	Responsive-ness
The mHealth App Usability Questionnaire for interactive mHealth apps (patient version) (MAUQ)	[65]	Muro-Culebras et al. (2021)	1	.895 .829 .900	No	Expert panel	Exploratory Factor Analysis	No	Correlation with PSSUQ & SUS	No
The mHealth App Usability Questionnaire for stand-alone mHealth apps (patient version) (MAUQ)	[65]	Muro-Culebras et al. (2021)	1	.847 .908 .717	No	Expert panel	Exploratory Factor Analysis	No	Correlation with PSSUQ & SUS	No
Usefulness, Satisfaction and Ease of Use Questionnaire (USE questionnaire)	[59]	Azad-Khaneghah et al. (2021) Inal et al. (2020)) Kien et al. (2018) Ng et al. (2019)	4	NR	NR	NR	NR	NR	NR	NR

NR: not reported as part of the systematic reviews included in the present overview.

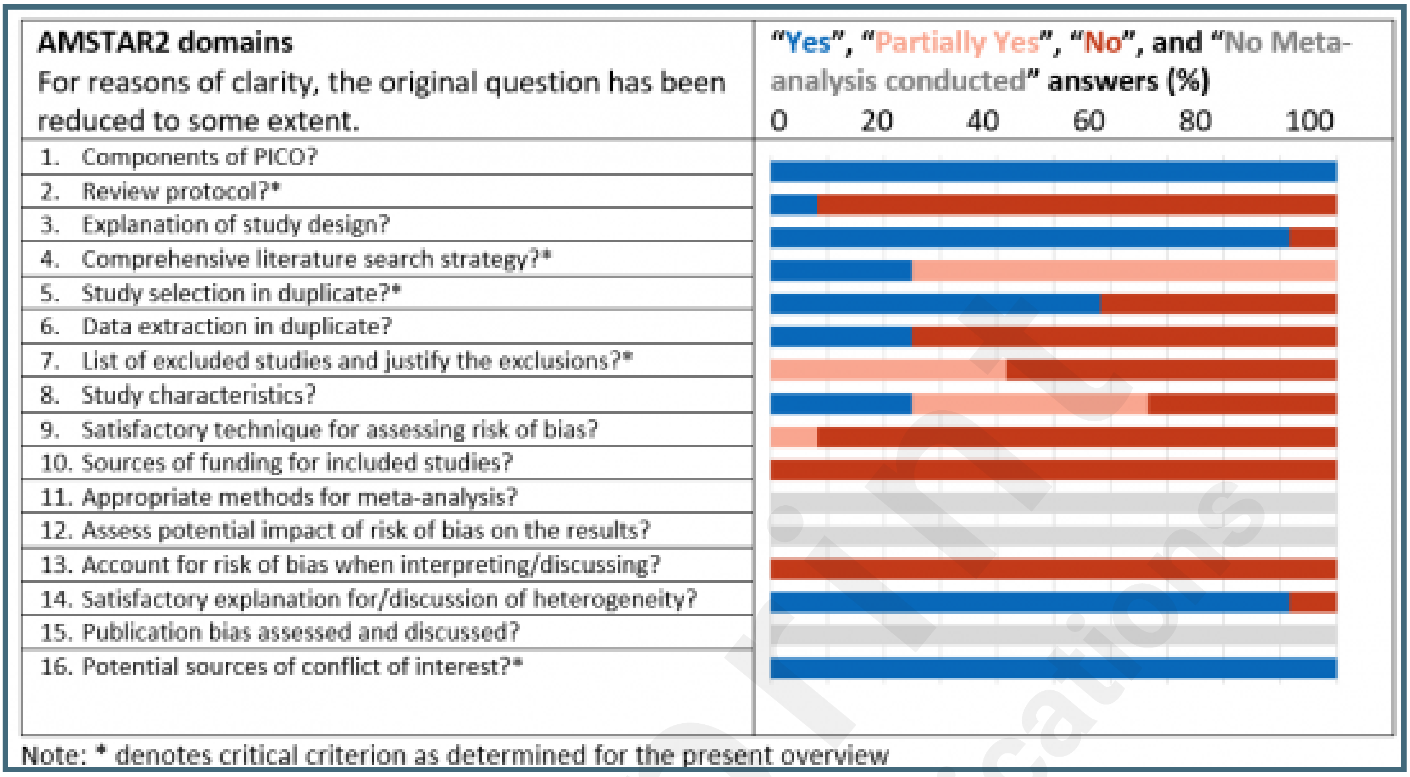
Supplementary Files

Figures

PRISMA flowchart.



Overview of methodological quality of reviews according to AMSTAR2.



Multimedia Appendixes

Supplementary files.
URL: http://asset.jmir.pub/assets/d7234ce68354544e16c86f3ea734afc4.docx

Related publication(s) - for reviewers eyes onlies

Letter to the editor.
URL: http://asset.jmir.pub/assets/0ef4562c4f4acc92c7b008b230a37c47.pdf