Review

Methodological Quality of Manuscripts Reporting on the Usability of Mobile Applications for Pain Assessment and Management: A Systematic Review

Ana F. Almeida [1], Nelson P. Rocha [1,2] and **Anabela G. Silva [3,4],***

[1] Institute of Electronics and Informatics Engineering of Aveiro (IEETA), Campus Universitário de Santiago, 3810-193 Aveiro, Portugal; fi.almeida7@ua.pt (A.F.A.); npr@ua.pt (N.P.R.)

[2] Department of Medical Sciences, Universidade de Aveiro - Edifício 30, Agras do Crasto - Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

[3] School of Health Sciences, Universidade de Aveiro - Edifício 30, Agras do Crasto - Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

[4] Center for Health Technology and Services Research, Universidade de Aveiro (CINTESIS.UA), Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

* Correspondence: asilva@ua.pt

check for
updates

Abstract: Background: There has been increasing use of mobile mHealth applications, including pain assessment and pain self-management apps. The usability of mHealth applications has vital importance as it affects the quality of apps. Thus, usability assessment with methodological rigor is essential to minimize errors and undesirable consequences, as well as to increase user acceptance. Objective: this study aimed to synthesize and evaluate existing studies on the assessment of the usability of pain-related apps using a newly developed scale. Methods: an electronic search was conducted in several databases, combining relevant keywords. Then titles and abstracts were screened against inclusion and exclusion criteria. The eligible studies were retrieved and independently screened for inclusion by two authors. Disagreements were resolved by discussion until consensus was reached. Results: a total of 31 articles were eligible for inclusion. Quality assessment revealed that most manuscripts did not assess usability using valid instruments or triangulation of methods of usability assessment. Most manuscripts also failed to assess the three domains of usability (effectiveness, efficiency and satisfaction). Conclusions: future studies should consider existing guidelines on usability assessment design, development and assessment of pain-related apps.

Keywords: pain; app; mobile app; mobile application; application; usability

---

1. Introduction

Pain is a global health problem, affecting all populations, regardless of age, sex, income, race/ethnicity, or geography [1]. It represents one of the main motives for seeking healthcare and a huge clinical, social and economic problem [1], conditioning life activities and being responsible for morbidity, absence from work, and temporary or permanent disability. Inadequate pain assessment and poor management have an impact on patients' psychological status, often resulting in anxiety and depression, and contribute to the longstanding maintenance of pain [2]. Therefore, choosing an appropriate instrument to assess pain is very important [2], and constitutes the first step for effective pain management [3].

Over the years, there has been a global push towards using information technologies to address health needs [4], particularly as the mobile Health (mHealth) application (app) market grows. A study of the status and trends of the mHealth market reported that in 2016 there were about 259.000 mHealth

apps available on major app stores [5]. Furthermore, there is an increasing number of smartphone apps devoted to pain assessment and pain self-management strategies available in app stores [6–8]. The first three pain-related apps that appear in the Play Store when searched using the word "pain"(search performed on the 15th of January 2020 on a smartphone) were Manage my Pain, Pain Diary and Pain Companion and the description of these apps refers that they have been downloaded 100,000, 50,000, 10,000 or more times, respectively, suggesting that a considerable number of persons is using them. Recent reviews have evaluated the quality of pain-related apps [9,10], focusing on its feasibility, acceptability or content and functioning. However, a major aspect to consider when assessing the quality of an app is usability [11], but usability has received less attention.

The International Organization for Standardization (ISO) 9241-11 [12] defines usability as the "extent to which a system, product or service can be used by specific users to achieve specific goals with effectiveness, efficiency, and satisfaction in a specific context of use". Effectiveness refers to the accuracy and completeness with which users achieve specified goals, and can be characterized using measures of accuracy and completeness such as task success. Efficiency refers to the resources used concerning the results achieved, and the time needed to complete tasks can be used as an indicator. Satisfaction is the extent to which the user's physical, cognitive and emotional responses that result from the use of a system, product or service meet the user's needs and expectations and can be assessed through interviews, focus groups or scales and questionnaires [12]. The assessment of usability can be formative if its main aim is to detect and solve problems or summative if its main aim is to meet the metrics associated with the system, product, or service task and goals [13]. A formative assessment is usually more common at earlier phases of a system or product development, and the type of instruments and procedures used depends on whether the assessment is formative or summative and on the development phase of the system, product, or service [14]. Independently of the phase of development, it is usually good practice to use a combination of approaches for usability assessment, involving, for example, the combination of quantitative or qualitative approaches, health professionals and patients or assessments in the laboratory and real context [15].

Usability assessment with methodological rigor is essential to minimize the likelihood of errors and undesirable consequences, to increase user acceptance, and to develop highly usable systems and products more likely to be successful in the market [16]. The adoption of usable mHealth systems or products will lead to safer and higher quality care, and a higher return of investment for institutions [17]. However, the assessment of usability is a complex process that requires, as referred, a combination of methods for a deeper and comprehensive evaluation of a product/service [18]. The methodological robustness and quality of these methods can have an impact on the results of the usability assessment and, therefore, should be carefully considered when interpreting the results of usability studies. Reviews on mobile applications for pain assessment or management have focused on the quality and content of the mobile applications [19,20] rather than on the quality of the process of assessment of usability. A robust existing scale, the Mobile App Rating Scale, also aims to assess the quality of the mobile apps [21]. A recent systematic review attempted to assess the quality of usability studies using a four-point scale [22] but failed to report on the process of development of the scale, or its reliability or validity, fundamental characteristics of any assessment or measurement instrument. Recently, a 15-items scale was developed to assess the methodological quality of usability studies for eHealth applications or general applications [23]. This scale was developed using a robust process, including a three-round Delphi method with 25 usability experts to generate a first version of the scale, which was then tested for comprehensibility by another three experts to generate a final version of the scale. This final version was shown to be feasible and both valid and reliable [23]. The present review aims to synthesize and evaluate existing studies on the assessment of the usability of pain-related apps using this newly developed scale [23].

2. Methods

2.1. Study Design

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).

2.1.1. Search Strategy and Selection of The Studies

An electronic search was conducted in Academic Search Complete, Scopus, PubMed, ScienceDirect, Web of Science, Scielo and Institute of Electrical and Electronics Engineers (IEEE) Xplore Digital Library on the 10th of May 2019, from database inception to the day the search was conducted. The following combination of words was used for all databases: pain AND (app OR "mobile app" OR "mobile application" OR application) AND usability. To be included in this systematic review studies had to: include the assessment of the usability of a pain-related mobile application as a study aim in any study sample and setting, be a full article published in a peer-reviewed journal/conference proceeding and be written in English, Portuguese, Spanish or French. The pain-related mobile application could target pain assessment, pain intervention or both. For this review, a mobile app was defined as "a software/set or program that runs on a mobile device and performs certain tasks for the user" [24]. Review articles were excluded.

Two researchers (AGS and AFA) independently reviewed the retrieved references against inclusion criteria. Disagreements were resolved by discussion until consensus was reached.

2.1.2. Data Extraction

All retrieved references were imported into the reference software Mendeley (Elsevier, North Holland) and checked for duplicates. Then titles and abstracts were screened against inclusion and exclusion criteria. Posteriorly, full texts of potentially eligible studies were retrieved and independently screened for inclusion by two authors of this review (AGS and AFA). The agreement was measured using a Cohen's K. Values below 0.20 indicate no concordance, between 0.21 and 0.39 minimal concordance, between 0.40 and 0.59 weak concordance, between 0.60 and 0.79 moderate concordance, between 0.80 and 0.90 strong concordance and above 0.90 almost perfect concordance [25].

Data from each included manuscript were extracted by one of the authors (AFA) using a standardized form including 1) manuscript authors, 2) name of the app, 3) app aim (i) pain assessment, i.e., one or two-way communication applications mainly intended to monitor pain characteristics or ii) pain management, i.e., applications designed to provide support/deliver pain-related interventions) [26], 4) individuals involved in the usability and characteristics, 5) use case, 6) domain of usability assessed (efficiency, effectiveness and/or satisfaction), 7) procedures for usability assessment, and 8) usability outcomes. The extracted data was checked by the second author (AGS) and the disagreements between authors at any point in the process were resolved through discussion until consensus was achieved.

Regarding the usability domains, efficiency refers to the resources used concerning the results achieved; effectiveness refers to the accuracy and completeness with which users achieve specified goals, and satisfaction is the extent to which the user's physical, cognitive and emotional responses that result from the use of a system, product or service meet the user's needs and expectations [12]. A study was considered to have assessed efficiency if the time needed to complete tasks was reported; effectiveness when measures of accuracy and completeness regarding specified goals were reported (e.g., task success) and satisfaction was assessed through interviews, focus group and scales/questionnaires (e.g., the System Usability Scale (SUS)) [27].

2.2. Methodological Quality Assessment

The methodological quality of included studies was independently assessed by two reviewers (AGS and AFA) using the Critical Assessment of Usability Studies Scale [23]. This is a recently developed scale, which is both valid and reliable (Intraclass Correlation Coefficient—ICC = 0.81) and

Int. J. Environ. Res. Public Health **2020**, *17*, 785

4 of 12

scores vary between 0 and 100% [23]. This scale is composed of fifteen questions on the procedures used to assess usability: 1) Did the study use valid measurement instruments of usability (i.e., there is evidence that the instruments used assess usability)? 2) Did the study use reliable measurement instruments of usability (i.e., there is evidence that the instruments used have similar results in repeated measures in similar circumstances)? 3) Was there coherence between the procedures used to assess usability (e.g., instruments, context) and study aims? 4) Did the study use procedures of assessment for usability that were adequate to the development stage of the product/service? 5) Did the study use procedures of assessment for usability adequate to study participants' characteristics? 6) Did the study employ triangulation of methods for the assessment of usability? 7) Was the type of analysis adequate to the study's aims and variables measurement scale? 8) Was usability assessed using both potential users and experts? 9) Were participants who assessed the product/service usability representative of the experts' population and/or of the potential user's population? 10) Was the investigator that conducted usability assessments adequately trained? 11) Was the investigator that conducted usability assessments external to the process of product/service development? 12) Was the usability assessment conducted in the real context or close to the real context where product/service is going to be used? 13) Was the number of participants used to assess usability adequate (whether potential users or experts)? 14) Were the tasks that serve as the base for the usability assessment representative of the functionalities of the product/service? 15) Was the usability assessment based on continuous and prolonged use of the product/service over time? Items 12 and 15 may be considered as not applicable depending on the phase of product development.

A pilot test of three manuscripts was undertaken and results discussed to clarify potential differences regarding the understanding of the scale items before moving to the assessment of the remaining manuscripts. Disagreements were resolved by discussion until reaching a consensus. The agreement was measured using an ICC (Model 2,1) calculated using SPSS version 24, and an ICC of at least 0.7 was considered acceptable [28].

3. Results

3.1. Study Selection

Searches resulted in 1091 references. After removing duplicates (n = 185), 906 references were screened based on title and abstract and 69 full articles were retrieved. Of these, 31 articles were eligible for inclusion (Figure 1). The main reasons for exclusion were: studies without reference to mobile apps (n = 13) or no pain-related app (n = 5), studies that didn't assess usability (n = 15), studies that were reviews or protocols (n = 5). Cohen's K between the two researchers involved in study selection was 0.77, which indicates moderate concordance.

3.2. Mobile Apps

The 31 manuscripts included covered a total of 32 pain apps (one manuscript assessed three different apps [29]; one manuscript assessed two different apps [19]; two manuscripts assessed the Pain Droid [30,31], and another two assessed the Pain Squad+ [32,33]). Twenty-five of these apps were categorized as pain assessment, and included pain scales and pain diaries to record the users' pain episodes and pain characteristics. Of these 25 pain assessment apps, 23 were intended for patient use and two for health professionals use (INES-DIO and iPhone pain app). The remaining seven apps were categorized as pain management and included self-management strategies (e.g., meditation, guided relaxation), but all of them included pain assessment features too.
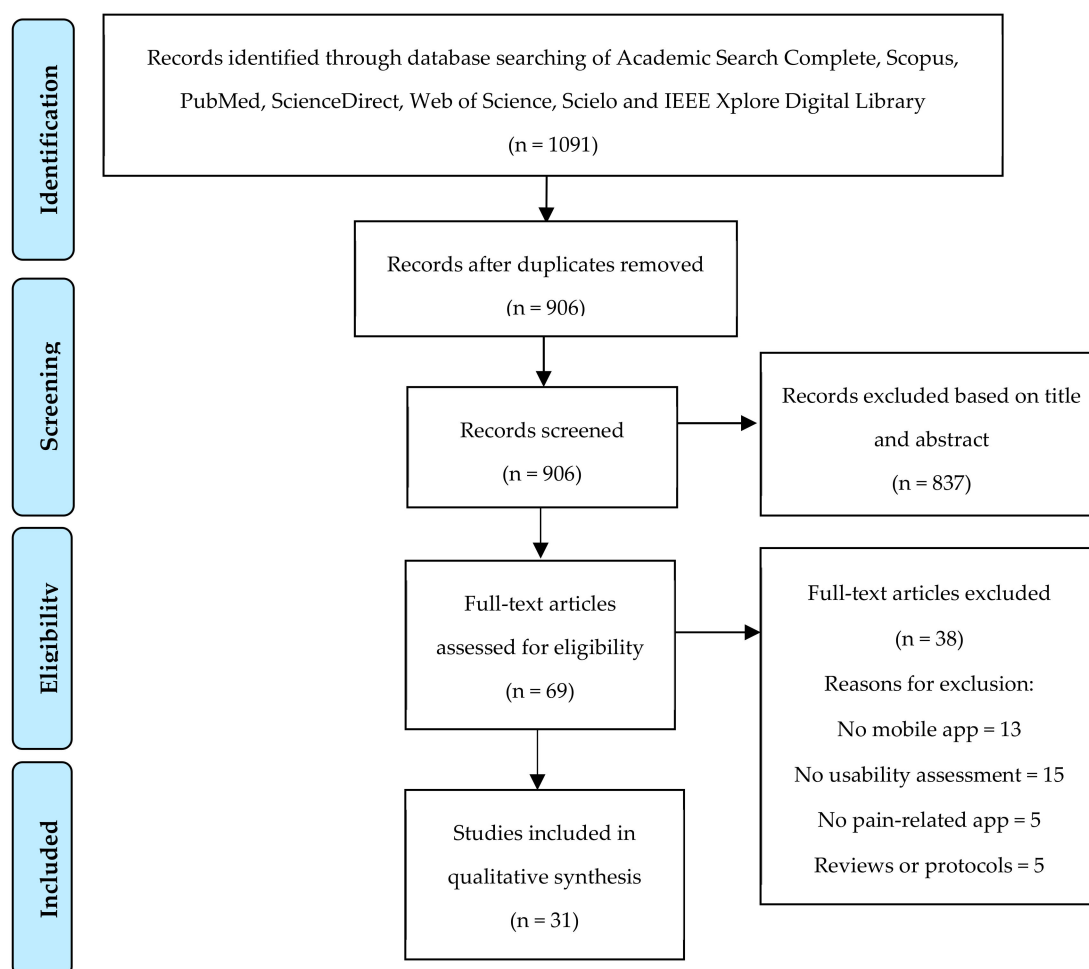
Figure 1. Study selection flowchart (CONSORT).

3.3. Usability Assessment

All manuscripts except three [34–36] assessed the domain "satisfaction". Only eight manuscripts described methods compatible with the assessment of all the three domains of usability (efficiency, effectiveness, and satisfaction), combining objective and subjective indicators, such as interviews, task completion rates and measuring the time participants needed to complete each task [19,29,32,37,38]. However, the results are not clearly and adequately reported in some manuscripts. For example, two studies of De La Vega et al. report, in their methods, recording of errors and use of SUS, however, they do not provide data on the number of errors neither present the final score of the SUS in the results section [39,40]. Similarly, other authors report the use of SUS, but do not provide its final score [31,41,42].

The procedures most commonly used for usability assessment were interviews and open-ended questions, used in 18 manuscripts, as well as verification of completion rates, also used in 18 manuscripts (Table 1). Other approaches to assess usability included using validated questionnaires or scales, observation, task completion times, think aloud and error rate.

Regarding the involvement of end-users in usability assessment, 23 manuscripts assessed apps that were intended for individuals with specific pathologies and used a sample of individuals with characteristics similar to the target group of end-users for usability assessment [19,29,31–34,36–38,40–53]. Participants in these studies included, among others, individuals with fibromyalgia and other rheumatological diseases [37,38,40,43,44], and individuals with headaches [29,32,33,45–48], cancer-related pain [36,49–51], or chronic pain in general [19,41,52,53], individuals with Down syndrome [34] and individuals who are wheelchair users [31,42]. Of these

23 studies, five also targeted specific age groups: four used children or adolescents [32,33,49,51] and one used older individuals [29]. From the remaining manuscripts (*n* = 8), two used children and adolescents [54,55] for usability assessment, four used individuals from the general population [35,39,56,57] and another two used health professionals [58,59]. The overview of the manuscripts analyzed is presented in the Table S1 from the Supplementary file.

Table 1. Methods of usability assessment reported in the manuscripts included in this systematic review.

References [1]	Method of Usability Assessment	*n* (%) (out of 31)
[32,38–40,44,48,53,55]	Think aloud approach	8 (25.81%)
[26,29,30,34,36,40,42,45,47–55]	Focus group, surveys or interviews	18 (58.06%)
[19,29,31,41–43,45,48,55–57]	Valid instrument/questionnaire (e.g., System Usability Scale, SUS)	11 (35.48%)
[32,34,37–40,46,48,53,54]	Observations / recording sessions / field notes	10 (32.26%)
[39,40,55]	Task error rate	3 (9.68%)
[29,32–37,39,44,47,49–51,54–58]	Completion rates	18 (58.06%)
[19,29,32,34,38,54–56]	Recording of task completion times	8 (25.81%)

[1] The total sum exceeds the number of manuscripts assessed since some used a combination of methods.

3.4. Methodological Quality Assessment

The ICC for reviewer's agreement was 0.71, IC95% [0.42–0.86]. The mean (±SD) score, in percentage, for the 31 manuscripts was 53.93% (SD = 13.01%), ranging between 20% and 73.33%. A more detailed analysis shows that all manuscripts, except one (out of 31) [51] assessed usability using appropriate procedures for the app development phase, but only two (out of 31) reported to have used an investigator adequately trained [35,40], and only one (out of 31) refers that the investigator responsible for usability assessment was external to the product development process [29]. Of the 31 manuscripts included, 11 used valid instruments to assess usability [19,29,31,41–43,45,48,55–57], but only nine triangulated methods of usability assessment [19,39,40,43,47,49,50,55,57]. The detailed results on the methodological quality assessment are presented in Table 2.

Table 2. Methodological quality assessment of studies included in the systematic review.

Authors	\multicolumn{15}{c}{Methodological Quality Items}	Score	%
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	—	——
Bedson 2019 [44]	0	0	1	1	1	0	0	1	1	0	0	1	0	1	1	8	53.33%
Suso-Ribera 2018 [41]	1	1	0	1	1	0	0	1	1	0	0	1	1	1	1	10	66.67%
Dantas 2016 [29]	1	1	1	1	1	0	1	0	0	0	1	0	0	1	1	9	60.00%
de la Vega 2014 [39]	0	0	1	1	0	1	0	1	1	0	0	0	1	1	0	7	46.67%
de la Vega 2018 [40]	0	0	1	1	0	1	0	0	1	1	0	0	1	1	0	7	46.67%
Diana 2012 [37]	0	0	1	1	1	0	0	0	1	0	0	N/A	1	1	N/A	5	38.46%
Docking 2018 [59]	0	0	1	1	1	0	1	1	1	0	0	0	1	0	0	7	46.67%
Fledderus 2015 [53]	0	0	1	1	1	0	1	1	1	0	0	N/A	1	0	N/A	7	53.85%
Fortier 2016 [51]	0	0	1	0	0	0	1	1	1	0	0	1	0	0	0	5	33.33%
Birnie 2018 [54]	0	0	1	1	1	0	1	0	1	0	0	1	1	1	0	7	46.67%
Boceta 2019 [58]	0	0	0	1	1	0	1	0	1	0	0	1	1	1	1	8	53.33%
Caon 2019 [56]	1	1	0	1	1	0	1	0	0	0	0	N/A	0	1	N/A	6	46.15%
Cardos 2017 [57]	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	6	40.00%
de Knegt 2016 [34]	0	0	1	1	1	0	1	0	1	0	0	0	1	1	0	7	46.67%
Kaltenhauser 2018 [35]	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	3	20.00%
Minen 2018 [48]	1	1	1	1	1	0	1	0	1	0	0	N/A	1	1	N/A	9	69.23%
Neubert 2018 [52]	0	0	1	1	1	0	1	1	1	0	0	1	1	1	0	9	60.00%

Int. J. Environ. Res. Public Health **2020**, *17*, 785

7 of 12

Table 2. *Cont.*

Authors						Methodological Quality Items										Score	%
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	—	—-
Reynoldson 2014 [19]	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	9	60.00%
Spyridonis 2012 [42]	1	1	1	1	1	0	1	0	1	0	0	0	0	1	0	8	53.33%
Spyridonis 2014 [31]	1	1	1	1	1	0	1	1	1	0	0	0	0	1	0	9	60.00%
Stefke 2018 [45]	1	1	1	1	1	0	1	0	0	0	0	1	0	1	1	9	60.00%
Stinson 2013 [49]	0	0	1	1	1	1	1	1	1	0	0	1	1	1	1	11	73.33%
Sun 2018 [55]	1	1	1	1	0	1	1	1	0	0	0	1	1	1	1	11	73.33%
Turner-Bowker 2011 [46]	0	0	1	1	1	0	0	0	1	0	0	0	1	0	0	5	33.33%
Vanderboom 2014 [43]	1	0	1	1	1	1	1	0	1	0	0	1	0	1	1	10	66.67%
Yen 2016 [38]	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	7	46.67%
Hochstenbach 2016 [50]	0	0	1	1	1	1	1	0	1	0	0	1	0	1	1	9	60.00%
Huguet 2015 [47]	0	0	1	1	1	1	1	0	1	0	0	1	1	1	1	10	66.67%
Jaatun 2013 [36]	0	0	0	1	1	0	0	1	1	0	0	N/A	1	1	N/A	6	46.15%
Jibb 2017 [32]	1	1	1	1	1	0	1	0	1	0	0	1	1	1	N/A	10	71.43%
Jibb 2018 [33]	1	1	1	1	1	0	1	0	1	0	0	1	1	1	1	11	73.33%
Total items scored yes (out of 31)	13	12	26	30	26	9	21	13	24	2	1	14	20	25	11		

Abbreviated items of the scale: 1—Valid measurement instruments, 2—reliable measurement instruments, 3—procedures adequate to study's aims, 4—procedures adequate to the development stage of the product, 5—procedures adequate to the participants' characteristics, 6—triangulation, 7—analysis adequate to the study's aims and variables measurement scale, 8—combination of user's and experts' evaluation, 9—representativeness of participants (potential users and/or experts) 10—experience of the investigator that conducted the usability evaluation, 11—investigator conducting usability assessment external to the development of the product/service, 12*—assessment in real context or close to real context, 13—number of participants (potential users and/or experts), 14—representativeness of the tasks to perform on the usability evaluation, 15*—continuous and prolonged use of the product. N/A—not-applicable.

4. Discussion

This systematic review found 31 manuscripts assessing the usability of a total of 32 pain-related apps, 25 of which were for pain assessment and seven for pain management. The lower number of mobile apps devoted to pain management may reflect the complexity of pain management, which requires multicomponent interventions but may also suggest that this field requires further development. This is the first systematic review that assesses the methodological quality of studies on the usability assessment of pain-related apps using a scale that is specific for the methodological quality of usability studies and that has been tested for its reliability and validity [23], and results suggest that several important methodological aspects regarding the assessment of usability are not being considered when developing pain-related apps. The complex nature of usability assessment is reflected in the low methodological quality of many studies with 12 (39%) out of 31 manuscripts scoring less than 50% in the methodological quality assessment. With this work, we aim to highlight the need for good practices in the assessment of usability more in line with existing recommendations [12,27].

Many of the studies included in the present systematic review fail to use valid (*n* = 18 out of 31) and reliable (*n* = 19 out of 31) measurement instruments. However, validity and reliability are fundamental characteristics of any measurement instrument and an indicator of their quality [60]. Similarly, studies fail to use triangulation of methods (*n* = 22 out of 31), despite the claims that a sound methodology of usability assessment requires the use of combined approaches [18]. Further, when using qualitative approaches for data collection, scarce details were provided in terms of the researchers involved and the procedures used for data collection and data analysis. We strongly suggest that authors follow the existing criteria for reporting qualitative research [61]. A few studies (*n* = 6) do not provide the total score for the instrument used, namely fail to provide the total score for the SUS or state in the methods to have assessed the number of errors, but do not provide this indicator in the results. We highlight the need for systematic planning and report of usability assessment procedures. The full description of all procedures employed for usability assessment may be included as an appendix section if the word limit of some journals prevents the authors from comprehensively reporting procedures and results.

Interestingly, only one manuscript reported having used older adults to assess the usability of a mobile app. Considering that pain prevalence tends to increase with age, mobile apps can have the potential to help health professionals reach a higher number of older adults with pain at lower costs. However, the specificities of this group, including a high number of painful body sites, increased comorbidities, lower digital literacy when compared to younger groups, and potential cognitive deficits in subgroups of older adults, suggest that the design and content of mobile apps need to be specifically designed and tested for this age group.

Other important methodological aspects that most manuscripts did not report was on the experience or training of the researcher involved in the assessment of usability ($n = 29$ out of 31) and whether this person was external to the team developing the app ($n = 30$ out of 31). Nevertheless, most studies employed procedures of usability assessment, such as think-aloud ($n = 8$ out of 31) and focus groups and interviews ($n = 18$ out of 31), that greatly depend on the ability of the researcher to conduct the assessment. Furthermore, if the researcher has a vested interest in the application, this can unintentionally bias the results towards a more favorable outcome. This has been shown for intervention studies, where beliefs and expectations have been found to bias the results towards a higher probability of a type I error (i.e., false-positive result) [62]. The lack of methodological detail on the reports of studies on usability has already been highlighted [23,63]. The exponential growth and the enormous potential of mobile apps to change the paradigm of health interventions, by increasing the access of individuals to health services at lower costs, requires a rigorous and methodologically sound assessment.

In terms of the usability domains assessed, all but three manuscripts assessed "satisfaction" [34–36]. Besides, most manuscripts did not report on the measurement of effectiveness, efficiency, and satisfaction, because they fail to use a combination of methods that allows assessing these three domains [27]. Only eight out of the 31 manuscripts included [19,29,32,37,38,54–56] assessed all the three domains of usability, despite existing recommendations to include measures of efficiency, effectiveness, and user satisfaction since a narrower selection of usability measures may lead to unreliable conclusions about the overall usability of the app [64]. Furthermore, there was an inconsistency between what was reported in the methodology section and the results presented. For example, some studies reported to have collected the number of errors (task error rate) or to have used a specific instrument but did not report on these results in the manuscript results section [37], [31,39,41,42].

mHealth solutions have the potential to foster self-assessment and self-management for patients suffering from pain and have a positive impact on their overall functioning and quality of life. The International Association for the Study of Pain has highlighted the mobile apps as a new advantage in the field of pain, highlighting the potential of technology to improve access to health care, contain costs, and improve clinical outcomes, but also calling for the need of studies measuring their efficacy, feasibility, usability, and compliance, and for the involvement of the scientific community reviewing the quality of existing solutions [65].

There are a few limitations to this systematic review. First, the protocol for this systematic review was not registered in a public database. Secondly, the grey literature was not searched. In addition, it lacks the analysis of the impact of the quality of the usability procedures on the results of the usability assessment and the quality of the resulting mobile app. However, the diversity of the procedures used in the manuscripts included in the systematic review makes this analysis difficult.

5. Conclusions

This systematic review found 31 manuscripts assessing the usability of a total of 32 pain-related apps, 25 of which were for pain assessment and seven for pain management. A detailed methodological analysis of these manuscripts revealed that several important methodological aspects regarding the assessment of usability for pain-related applications are not being considered. Future developments should be planned and implemented in line with existing guidelines.

References

1. Henschke, N.; Kamper, S.J.; Maher, C.G. The epidemiology and economic consequences of pain. *Mayo Clin. Proc.* **2015**, *90*, 139–147. [CrossRef] [PubMed]

2. Wells, N.; Pasero, C.; McCaffery, M. Improving the quality of care through pain assessment and management. In *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*; Hughes, R.G., Ed.; AHRQ Publication: Rockville, MD, USA, 2008.

3. Breivik, H.; Borchgrevink, P.C.; Allen, S.M.; Rosseland, L.A.; Romundstad, L.; Hals, E.K.; Kvarstein, G.; Stubhaug, A. Assessment of pain. *Br. J. Anaesth.* **2008**, *101*, 17–24. [CrossRef] [PubMed]

4. Vital Wave Consulting. *"mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World"*; UN Foundation-Vodafone Foundation Partnership: Washington, DC, USA; Berkshire, UK, 2009.

5. Research2Guidance, mHealth Economics 2016 – Current Status and Trends of the mHealth App Market. 2016. Available online: https://research2guidance.com/product/mhealth-app-developer-economics-2016/ (accessed on 20 September 2019).

6. Pain management-Apple. Available online: https://www.apple.com/us/search/Pain-management?src=serp (accessed on 20 September 2019).

7. Pain management app - Android Apps on Google Play. Available online: https://play.google.com/store/search?q=pain%20management%20app&c=apps&hl=en_US (accessed on 20 September 2019).

8. Pain assessment app - Android Apps on Google Play. Available online: https://play.google.com/store/search?q=Pain%20assessment%20app&c=apps&hl=en_US (accessed on 20 September 2019).

9. Lalloo, C.; Jibb, L.A.; Rivera, J.; Agarwal, A.; Stinson, J.N. There's a pain app for that': Review of patient-targeted smartphone applications for pain management. *Clin. J. Pain* **2015**, *31*, 557–563. [CrossRef] [PubMed]

10. Wallace, L.S.; Dhingra, L.K. A systematic review of smartphone applications for chronic pain available for download in the United States. *J. Opi. Manag.* **2014**, *10*, 63–68.

11. Liew, M.S.; Zhang, J.; See, J.; Ong, Y.L. Usability challenges for health and wellness mobile apps: mixed-methods study among mhealth experts and consumers. *JMIR mHealth uHealth* **2019**, *7*, e12160. [CrossRef] [PubMed]

12. ISO 9241-11. Ergonomics of Human-System Interaction — Part 11: Usability: Definitions and Concepts. 2018. Available online: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&ved=2ahUKEwjCuYCa9qLnAhWNJaYKHQF9Cd0QFjAEegQIBhAB&url=https%3A%2F%2Finfostore.saiglobal.com%2Fpreview%2Fis%2Fen%2F2018%2Fi.s.eniso9241-11-2018.pdf%3Fsku%3D1980667&usg=AOvVaw1kP3G6g9MDGCVd_WazNF6c (accessed on 20 September 2019).

13. Lewis, J.R. Usability: Lessons learned . . . and yet to be learned. *Int. J. Hum. Comput. Interact.* **2014**, *30*, 663–684. [CrossRef]

14. Klaassen, B.; van Beijnum, B.; Hermens, H. Usability in telemedicine systems-A literature survey. *Int. J. Med. Inform.* **2016**, *93*, 57–69. [CrossRef]

15. Martins, A.I.; Queirós, A.; Silva, A.G.; Rocha, N.P. Usability evaluation methods: A systematic review. In *Human Factors in Software Development and Design*; Daeed, S., Bajwa, I.S., Mahmood, Z., Eds.; IGI Global: Pennsylvania, PA, USA, 2015.

Int. J. Environ. Res. Public Health **2020**, *17*, 785

10 of 12

16. ISO 9241-210:2019. Ergonomics of Human-System Interaction — Part 210: Human-Centred Design for Interactive Systems. 2019. Available online: https://www.iso.org/standard/77520.html (accessed on 20 September 2019).

17. Middleton, B.; Bloomrosen, M.; Dente, M.A.; Hashmat, B.; Koppel, R.; Overhage, J.M.; Payne, T.H.; Rosenbloom, S.T.; Weaver, C.; Zhang, J. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: Recommendations from AMIA. *J. Am. Med. Inform. Assoc.* **2013**, *20*. [CrossRef]

18. Martins, A.I.; Queirós, A.; Rocha, N.P.; Santos, B.S. Avaliação de usabilidade: Uma revisão sistemática da literatura. *RISTI Rev. Iber. Sist. e Tecnol. Inf.* **2013**, *11*, 31–43. [CrossRef]

19. Reynoldson, C.; Stones, C.; Allsop, M.; Gardner, P.; Bennett, M.I.; Closs, S.J.; Jones, R.; Knapp, P. Assessing the quality and usability of smartphone apps for pain self-management. *PAIN Med.* **2014**, *15*, 898–909. [CrossRef]

20. Devan, H.; Farmery, D.; Peebles, L.; Grainger, R. Evaluation of self-management support functions in apps for people with persistent pain: Systematic review. *JMIR mHealth uHealth* **2019**, *7*, e13080. [CrossRef] [PubMed]

21. Stoyanov, S.R.; Hides, L.; Kavanagh, D.J.; Zelenko, O.; Tjondronegoro, D.; Mani, M. Mobile app rating scale: A new tool for assessing the quality of healh mobile apps. *JMIR mHealth uHealth* **2015**, *3*, e27. [CrossRef] [PubMed]

22. Shah, U.M.; Chiew, T.K. A systematic literature review of the design approach and usability evaluation of the pain management mobile applications. *Symmetry* **2019**, *11*, 400. [CrossRef]

23. Silva, A.; Simões, P.; Santos, R.; Queirós, A.; Rocha, N.; Rodrigues, M. A scale to assess the methodological quality of studies assessing usability of eHealth products: A delphi study followed by validity and reliability testing. *J. Med. Inter. Res.* **2019**, *21*, e14829.

24. Islam, M.R.; Mazumder, T.A. Mobile application and its global impact. *Int. J. Eng. Technol. IJET IJENS* **2010**, *10*, 104–111.

25. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Medica* **2012**, *22*, 276–282. [CrossRef]

26. Silva, A.; Queirós, A.; Caravau, H.; Ferreira, A.; Rocha, N. Systematic review and evaluation of pain-related mobile applications. In *Encyclopedia of E-Health and Telemedicine*; Cruz-Cunha, M.M., Miranda, I.M., Martinho, R., Rijo, R., Eds.; IGI Global: Hershey, PA, USA, 2016; pp. 383–400.

27. Ratwani, R.M.; Zachary Hettinger, A.; Kosydar, A.; Fairbanks, R.J.; Hodgkins, M.L. A framework for evaluating electronic health record vendor user-centered design and usability testing processes. *J. Am. Med. Inform. Assoc.* **2017**, *24*, e35–e39. [CrossRef]

28. Weir, J.P. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J. Strength Cond. Res.* **2005**, *19*, 231–240.

29. Dantas, T.; Santos, M.; Queirós, A.; Silva, A.G. Mobile applications in the management of headache. *Proc. Comput. Sci.* **2016**, *100*, 369–374. [CrossRef]

30. Spyridonis, F.; Gawronski, J.; Ghinea, G.; Frank, A.O. An interactive 3-D application for pain management: Results from a pilot study in spinal cord injury rehabilitation. *Comput. Meth. Prog. Biomed.* **2012**, *108*, 356–366. [CrossRef]

31. Spyridonis, F.; Hansen, J.; Grønli, T.-M.; Ghinea, G. PainDroid: An android-based virtual reality application for pain assessment. *Multimed. Tools Appl.* **2014**, *72*, 191–206. [CrossRef]

32. Jibb, L.A.; Cafazzo, J.A.; Nathan, P.C.; Seto, E.; Stevens, B.J.; Nguyen, C.; Stinson, J.N. Development of a mhealth real-time pain self-management app for adolescents with cancer: An iterative usability testing study. *J. Pediatr. Oncol. Nurs.* **2017**, *34*, 283–294. [CrossRef] [PubMed]

33. Jibb, L.A.; Stevens, B.J.; Nathan, P.C.; Seto, E.; Cafazzo, J.A.; Johnston, D.L.; Hum, V.; Stinson, J.N. Perceptions of adolescents with cancer related to a pain management app and its evaluation: Qualitative study nested within a multicenter pilot feasibility study. *JMIR mHealth uHealth* **2018**, *6*, e80. [CrossRef] [PubMed]

34. De Knegt, N.C.; Lobbezoo, F.; Schuengel, C.; Evenhuis, H.M.; Scherder, E.J.A. Self-reporting tool on pain in people with intellectual disabilities (stop-id!): A usability study. *Altern. Commun.* **2016**, *32*, 1–11. [CrossRef]

35. Kaltenhauser, A.; Schacht, I. Quiri: Chronic pain assessment for patients and physicians. *MobileHCI'18* **2018**. [CrossRef]

36. Jaatun, E.A.A.; Haugen, D.F.; Dahl, Y.; Kofod-Petersen, A. An improved digital pain body map. *Healthcom* **2013**. [CrossRef]

37. Diana, C.; Cristina, B.; Azucena, G.-P.; Luis, F.; Ignacio, M. Experience-sampling methodology with a mobile device in fibromyalgia. *Int. J. Telemed. Appl.* **2012**. [CrossRef]

38. Yen, P.-Y.; Lara, B.; Lopetegui, M.; Bharat, A.; Ardoin, S.; Johnson, B.; Mathur, P.; Embi, P.J.; Curtis, J.R. Usability and workflow evaluation of 'RhEumAtic disease activitY' (READY). A mobile application for rheumatology patients and providers. *Appl. Clin. Inform.* **2016**, *7*, 1007–1024. [CrossRef]

39. De la Vega, R.; Roset, R.; Castarlenas, E.; Sanchez-Rodriguez, E.; Sole, E.; Miro, J. Development and testing of painometer: A smartphone app to assess pain intensity. *J. Pain* **2014**, *15*, 1001–1007. [CrossRef]

40. De la Vega, R.; Roset, R.; Galan, S.; Miro, J. Fibroline: A mobile app for improving the quality of life of young people with fibromyalgia. *J. Heal. Psychol.* **2018**, *23*, 67–78. [CrossRef]

41. Suso-Ribera, C.; Castilla, D.; Zaragoza, I.; Victoria Ribera-Canudas, M.; Botella, C.; Garcia-Palacios, A. Validity, reliability, feasibility, and usefulness of pain monitor: A multidimensional smartphone app for daily monitoring of adults with heterogenous chronic pain. *Clin. J. Pain* **2018**, *34*, 900–908. [CrossRef] [PubMed]

42. Spyridonis, F.; Gronli, T.-M.; Hansen, J.; Ghinea, G. Evaluating the usability of a virtual reality-based android application in managing the pain experience of wheelchair users. *IEEE Eng. Med. Biol. Soc.* **2012**, *2012*, 2460–2463.

43. Vanderboom, E.; Vincent, A.; Luedtke, C.A.; Rhudy, L.M.; Bowles, K.H. Feasibility of interactive technology for symptom monitoring in patients with fibromyalgia. *Pain Manag. Nurs.* **2014**, *15*, 557–564. [CrossRef]

44. Bedson, J.; Hill, J.; White, D.; Chen, Y.; Wathall, S.; Dent, S.; Cooke, K.; van der Windt, D. Development and validation of a pain monitoring app for patients with musculoskeletal conditions (The Keele pain recorder feasibility study). *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 24. [CrossRef] [PubMed]

45. Stefke, A.; Wilm, F.; Richer, R.; Gradl, S.; Eskofier, B.; Forster, C.; Namer, B. MigraineMonitor - Towards a system for the prediction of migraine attacks using electrostimulation. *Curr. Dir. Biomed. Eng.* **2018**, *4*, 629–632. [CrossRef]

46. Turner-Bowker, D.M.; Saris-Baglama, R.N.; Smith, K.J.; DeRosa, M.A.; Paulsen, C.A.; Hogue, S.J. Heuristic evaluation and usability testing of a computerized patient-reported outcomes survey for headache sufferers. *Telemed. J. E. Health.* **2011**, *17*, 40–45. [CrossRef]

47. Huguet, A.; McGrath, P.J.; Wheaton, M.; Mackinnon, S.P.; Rozario, S.; Tougas, M.E.; Stinson, J.N.; MacLean, C. Testing the feasibility and psychometric properties of a mobile diary (myWHI) in adolescents and young adults with headaches. *JMIR mHealth uHealth* **2015**, *3*, 2. [CrossRef]

48. Minen, M.T.; Jalloh, A.; Ortega, E.; Powers, S.W.; Sevick, M.A.; Lipton, R.B. User design and experience preferences in a novel smartphone application for migraine management: A think aloud study of the RELAXaHEAD application. *Pain Med.* **2019**, *20*, 369–377. [CrossRef]

49. Stinson, J.N.; Lindsay, A.J.; Nguyen, C.; Nathan, P.C.; Maloney, A.M.; Dupuis, L.L.; Gerstle, J.T.; Alman, B.; Hopyan, S.; Strahlendorf, C.; et al. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. *J. Med. Inter. Res.* **2013**, *15*, e51. [CrossRef]

50. Hochstenbach, L.M.J.; Zwakhalen, S.M.G.; Courtens, A.M.; van Kleef, M.; de Witte, L.P. Feasibility of a mobile and web-based intervention to support self-management in outpatients with cancer pain. *Eur. J. Oncol. Nurs.* **2016**, *23*, 97–105. [CrossRef]

51. Fortier, M.A.; Chung, W.W.; Martinez, A.; Gago-Masague, S.; Sender, L. Pain buddy: A novel use of m-health in the management of children's cancer pain. *Comput. Biol. Med.* **2016**, *76*, 202–214. [CrossRef]

52. Neubert, T.-A.; Dusch, M.; Karst, M.; Beissner, F. Designing a tablet-based software app for mapping bodily symptoms: Usability evaluation and reproducibility analysis. *JMIR mHealth uHealth* **2018**, *6*, e127. [CrossRef] [PubMed]

53. Fledderus, M.; Schreurs, K.M.; Bohlmeijer, E.T.; Vollenbroek-Hutten, M.M. Development and pilot evaluation of an online relapse-prevention program based on acceptance and commitment therapy for chronic pain patients. *JMIR Hum. Factor.* **2015**, *2*, e1. [CrossRef] [PubMed]

54. Birnie, A.; Nguyen, C.; Do Amaral, T.; Baker, L.; Campbell, F.; Lloyd, S.; Ouellette, C.; von Baeyer, C.; Chitra Ted Gerstle, L.J.; Stinson, J. A parent-science partnership to improve postsurgical pain management in young children: Co-development and usability testing of the Achy Penguin smartphone-based app. *Can. J. Pain* **2018**, *2*, 280–291. [CrossRef]

55. Sun, T.; Dunsmuir, D.; Miao, I.; Devoy, G.M.; West, N.C.; Görges, M.; Lauder, G.R.; Ansermino, J.M. In-hospital usability and feasibility evaluation of Panda, an app for the management of pain in children at home. *Paediatr. Anaesth.* **2018**, *28*, 897–905. [CrossRef]

56. Caon, M.; Angelini, L.; Ledermann, K.; Martin-Sölch, C.; Khaled, O.A.; Mugellini, E. My pain coach: A mobile system with tangible interface for pain assessment. *Adv. Intell. Syst. Comput.* **2019**, *824*, 1372–1381.

Int. J. Environ. Res. Public Health **2020**, *17*, 785

12 of 12

57. Cardos, R.A.I.; Soflau, R.; Gherman, A.; Sucală, M.; Chiorean, A. A mobile intervention for core needle biopsy related pain and anxiety: A usability study. *J. Evid. Based Psychother.* **2017**, *17*, 21–30. [CrossRef]

58. Boceta, J.; Samper, D.; de la Torre, A.; Sanchez-de la Rosa, R.; Gonzalez, G. Usability, acceptability, and usefulness of an mhealth app for diagnosing and monitoring patients with breakthrough cancer pain. *JMIR Can.* **2019**, *5*, e10187. [CrossRef]

59. Docking, R.E.; Lane, M.; Schofield, P.A. Usability testing of the iphone app to improve pain assessment for older adults with cognitive impairment (prehospital setting): A qualitative study. *Pain Med.* **2018**, *19*, 1121–1131. [CrossRef]

60. Streiner, D.L.; Norman, G.R.; Cairney, J. *Health Measurement Scales*, 5th ed.; Oxford University Press: Oxford, UK, 2015.

61. Tong, A.; Sainsbury, P.; Craig, J. Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *Int. J. Qual. Heal. Care* **2007**, *19*, 349–357. [CrossRef]

62. Lewis, S.C.; Warlow, C.P. How to spot bias and other potential problems in randomised controlled trials. *J. Neurol. Neurosurg. Psych.* **2004**, *75*, 181–187. [CrossRef] [PubMed]

63. Ellsworth, M.A.; Dziadzko, M.; Horo, J.C.O.; Farrell, A.M.; Zhang, J.; Herasevich, V. An appraisal of published usability evaluations of electronic health records via systematic review. *J. Am. Med. Inform. Assoc.* **2017**, *24*, 218–226. [CrossRef] [PubMed]

64. Frokjaer, E.; Hertzum, M.; Hornbæk, K. Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated? In Proceedings of the Conference on Human Factors in Computing Systems, Hague, The Netherlands, 1–6 April 2000; 2000; pp. 345–352.

65. Vardeh, D.; Edwards, R.R.; Jamison, R.N.; Eccleston, C. There's an app for that: Mobile technology is a new advantage in managing chronic pain. *Pain Clin. Updat.* **2013**, *21*, 1–7.