Issues in reporting of systematic review methods in health app-focused reviews: A scoping review

Rebecca Grainger
University of Otago Wellington, New Zealand

Hemakumar Devan
Centre for Health, Activity and Rehabilitation Research (CHARR), School of Physiotherapy, University of Otago, Wellington, New Zealand

Bahram Sangelaji
Jean Hay-Smith
University of Otago Wellington, New Zealand

Abstract
No guidelines exist for the conduct and reporting of manuscripts with systematic searches of app stores for, and then appraisal of, mobile health apps (‘health app-focused reviews’). We undertook a scoping review including a systematic literature search for health app-focused reviews describing systematic app store searches and app appraisal, for apps designed for patients or clinicians. We created a data extraction template which adapted data elements from the PRISMA guidelines for systematic literature reviews to data elements operationalised for health app-focused reviews. We extracted the data from included health app-focused reviews to describe: (1) which elements of the adapted ‘usual’ methods of systematic review are used; (2) methods of app appraisal; and (3) reporting of clinical efficacy and recommendations for app use. From 2798 records, the 26 included health app-focused reviews showed incomplete or unclear reporting of review protocol registration; use of reporting guidelines; processes of screening apps; data extraction; and appraisal tools. Reporting of clinical efficacy of apps or recommendations for app use were infrequent.

Corresponding author:
Rebecca Grainger, Department of Medicine, University of Otago Wellington, PO Box 7343, 23a Mein St, Newtown, Wellington South 6242, New Zealand.
Email: rebecca.grainger@otago.ac.nz

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
The reporting of methods in health app-focused reviews is variable and could be improved by developing a consensus reporting standard for health app-focused reviews.

**Keywords**
apps, ehealth, methodology, mhealth, systematic review

**Background and significance**

Smartphone use is ubiquitous, with user numbers expected to reach 5 billion in 2020.\(^1\) Given the rise in smartphone ownership and internet access, mobile health apps are widely used as a self-management tool for people living with long-term health conditions\(^2\)–\(^4\) Furthermore, clinicians are increasingly using mobile apps in clinical practice for patient education and to facilitate self-management,\(^5\)\(^,\)\(^6\) and for supporting clinical care such as clinical decision making, diagnosis, disease activity monitoring, and prescribing.

The interest in mobile health apps is evident from the increase in publication rate for manuscripts describing systematic search of apps stores for, and appraisal of health apps (which we have called ‘health app-focused reviews’) in the last 5 years.\(^2\)\(^,\)\(^7\) Many of these health app-focused reviews have ‘systematic’ in the title or include typical features of a systematic literature review in the reporting of methods. The consistent finding from these health app-focused reviews is that most existing self-management apps from app stores are: developed with minimal end-user input and lack evidence-based content, designed for commercial use with minimal or no clinician involvement, not tested for clinical efficacy.\(^2\)\(^,\)\(^4\) These results have led to creation of guidelines by government agencies for mobile app developers\(^8\) and critical appraisal checklists for clinicians for recommending mobile apps in clinical practice.\(^9\)

These health app-focused reviews, which may describe app features, functionality, efficacy and effectiveness, are critical for several reasons. First, they inform clinicians and health systems about apps that can be recommended for patient use in a patient-empowered health system. Second, they inform researchers and developers about the gaps for high-quality health apps to direct future research or development. Most importantly, they provide a high-quality trusted resource for informed patients to choose the highest quality app to support optimising health.

While these health app-focused reviews can inform evidence-based practice, there is no existing methodological framework to guide such reviews. Adequate conduct and reporting are essential for maximising transparency of search and appraisal methods and to foster confidence in the reliability of health app-focused review findings.\(^10\) A clearly conducted and reported health app-focused review will provide sufficient details about the app store search, app identification, app data extraction and evaluation of apps so that a reader can be confident that all apps relevant to the health condition have been identified and appraised. For example, clear statement of the country and date on which an app store was searched and language of apps included are essential for a reader to know if the apps retrieved are likely to be relevant to them. There are standardised reporting guidelines tailored for systematic reviews and meta-analysis of randomised control trials (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA\(^11\)) and qualitative studies (Enhancing transparency in reporting the synthesis of qualitative research; ENTREQ\(^12\)). The PRISMA reporting guidelines may be a framework for the potential items for reporting guidelines for health app-focused reviews but it is not clear if all items are relevant or how to operationalise the items that are relevant. A logical next step is to develop a reporting guideline for health app-focused reviews.
As a first step to develop reporting guidelines, it is appropriate to review existing health app-focused reviews. This provides a basis for developing a reporting standard for publishing health app-focused reviews for people managing aspects of their health and for clinicians. The main purpose of this scoping review is to provide a narrative synthesis on reporting of systematic search of app stores and app appraisal methods used in health app-focused reviews, designed to support people to self-manage aspects of diagnosed health conditions or for clinicians for use in clinical care.

**Objectives**

The key aims of our review were:

1. What parts of the ‘usual’ methods of systematic review are used and adapted in health app-focused reviews (e.g. how are app search and selection methods used and adapted).
2. In particular, how are methods of critical appraisal adapted? Usually it is the methods of individual studies that are appraised, but in app reviews it is the content of the apps, and how is this done (including application of guidance on assessing the quality of apps, such as the Mobile App Rating Scale (MARS))
3. What recommendations are made based on health app-focused reviews and how the ‘strength’ of those recommendations is estimated (e.g. any equivalent to GRADE).

**Materials and methods**

A scoping review methodology was adopted based on the Arksey and O’Malley framework stages as the overarching aim was to explore the current state of reporting of methods of health app-focused reviews. An a priori review protocol was developed (available on request) and registered with PROSPERO International Registry of Prospective Systematic reviews (CRD42018102786). The review was reported based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses – Extension for Scoping reviews (PRISMA-ScR) with checklist provided in Supplemental Appendix 1.

**Primary search**

Literature for the scoping review was identified by a systematic keyword search (Table 1) in the following databases: Medline (via Ovid), AMED, PsycINFO, Cochrane Library (via Ovid), PubMed, CINAHL (via EBSCO) and Scopus. The search was undertaken with reference librarian support, limited to 1 January 2012 to 1 August 2018, with the date of last search 9 August 2018. The search was limited to this period to allow for some maturity of the process of app review literature, since mobile apps first become widespread after the launch of the iPhone in 2007. Combining ‘app’ terms AND ‘review’ terms in a preliminary search suggested sufficient articles would be found from this period. All references were imported into Endnote (Version X9, Clarivate Analytics) then uploaded to Covidence for screening.

**Secondary search**

The same search terms were used to search the journal database of Journal of Medical Internet Research (JMIR) (frequently publishes health app-focused reviews and not all its journals are indexed in MEDLINE). The review registries PROSPERO and The Joanna Briggs Institute (JBI) were both manually searched. Identified records were exported into Endnote and uploaded to Covidence.
Inclusion and exclusion criteria

To identify health app-focused reviews that were most likely to be systematic in nature the following criteria were applied:

1. The article was ‘systematic’ defined by presence of the term ‘systematic review’ in the title OR an a priori systematic search procedure (along with inclusion and exclusion criteria) was described in the methods section of the article OR the search process was described based on PRISMA standard flow diagram for inclusion and exclusion of apps. (2) The article had a systematic search for health apps in at least one app store described in the text of the article. (3) There was a specific health condition named as focus of the article. (4) The focus of the article was apps for either people with any acute or chronic health condition where a specific disease diagnosis has been made (e.g. inflammatory bowel disease, diabetes, epilepsy or gout) or for clinicians where the app may support direct clinical care (which may include clinical decision making, diagnosis, disease activity measures, patient education and prescribing, and potentially other aspects, as identified by the review).

The exclusion criteria for articles were: (1) not in English language, (2) focus on apps for general wellbeing or general health, or for supporting aspects of mental or psychological health, and (3) focus on apps for supporting general nutrition outside a named health condition.

Article screening

After automatic removal of duplicates in Covidence, one author (BS) manually removed remaining duplicates. Two authors (BS, RG) independently reviewed all titles and abstracts against eligibility criteria. The full text of all eligible articles then underwent independent review by two authors (BS, RG) to confirm if inclusion criteria were met. At each stage, any disagreements were resolved by consensus discussion between those two authors and if decision could not be reached, articles were discussed with a third author (JHS or HD). When an included article had an author who was a member of the research team, that author was excluded from making decisions about or data extraction from that article.

Development of data extraction template

All authors contributed to development of the data extraction template (available on request). A draft template was created in Word (v16.24, Microsoft) based on standard data elements or sections in PRISMA guidelines. The data extraction template was refined by iterative discussions, and then piloted by extracting data from four included articles. For each article, two authors independently extracted the data. Further discussions with all authors (videoconference and email) led to iterative updating of the data extraction template. When disagreement occurred, decision rules were developed for data extraction. The final data extraction template data elements in a health app-focused review are presented in Table 2.
One author extracted data for every included article (BS) and transferred these to an Excel spreadsheet. The other authors (RG, HD, JHS) each independently extracted data from one third of the included articles. RG entered these data into the Excel spreadsheet. All discrepancies were discussed between the authors (BS and RG for all, with HD or JHS) and agreed by consensus.

Data analysis and synthesis

Reporting of each data element was categorised as ‘yes’ or ‘no’, with some items having categories of ‘unclear’ or ‘not applicable’. Data were reported as frequencies.
Results

Database search and screening

Flow of references is shown in Figure 1 including reasons for exclusion at full text review. Database searching identified 2274 records, and secondary searching 856 of which 332 were duplicate records. From 2798 records, 2615 were excluded after abstract and title screening. Of the remaining 183, full text was not obtainable for 6, leaving 177 eligible full text articles. After full text review, 26 articles were included for final synthesis. The majority of articles (n = 77) were excluded as they were not app reviews. Of the 58 articles that were app reviews, reasons for exclusion were mainly due to not specific to a health condition (n = 35) (Figure 1).
Overview of included articles

Over half (58%, 15/26) of included articles were published in the last 2 years of the search period (2017/18) (Table 3). Fourteen out of 26 articles (54%) stated systematic review in the title. Almost half of the included articles (45%, 12/26) were published in journals in the field of health informatics. Health app-focused reviews addressing a wide range of acute and chronic health conditions were included.

A priori review registration and registration guideline

Of the 26 included articles, one reported a priori protocol registration and four stated reporting was based on the PRISMA guideline for systematic reviews.7,30,37,39

Table 3. Description of included articles.

| First author | Year of publication | Journal | Health condition |
|--------------|---------------------|---------|------------------|
| Adam20       | 2018                | World of Urology | Prostate Cancer  |
| Athilingam21 | 2018                | JMIR Cardiologya | Heart Failure    |
| Bhattarai22  | 2018                | Pain Medicine   | Arthritic Pain   |
| Garcia-Cruz23| 2018                | Actas Urologicas Espanolas | Men's sexual disorders |
| Jupp24       | 2018                | Pediatric Blood and cancer | Cancer/Oncology |
| Lee25        | 2018                | JMIR mHealth and uHealtha | Chronic Kidney Disease |
| Luo26        | 2018                | Journal of Clinical Rheumatology | Rheumatoid Arthritis |
| Rey27        | 2018                | Neurologica | Parkinson’s Disease |
| Svendsen28   | 2018                | Journal of Dermatological Treatment | Psoriasis |
| Choi29       | 2017                | Health Informatics Journala | Osteoarthritis |
| Grainger30   | 2017                | JMIR mHealth and uHealtha | Rheumatoid Arthritis |
| Martinez31   | 2017                | JMIR Diabetesa | Diabetes |
| Rossi32      | 2017                | mHealtha | Diabetes |
| Sobnath33    | 2017                | JMIR mHealth and uHealtha | Chronic Obstructive Pulmonary Disease |
| Xie34        | 2017                | JMIR mHealth and uHealtha | Cardiovascular Disease |
| Arnhold35    | 2016                | JMIRa | Diabetes |
| Brouard36    | 2016                | Annals of Medicine | Cancer |
| Brzan37      | 2016                | Journal of Medical Systemsa | Diabetes |
| Con38        | 2016                | JMIR mHealth and uHealtha | Inflammatory Bowel Disease |
| Machado7     | 2016                | Best Practice and Research Clinical Rheumatology | Low back pain |
| Fijacko39    | 2015                | Journal of Medical Systemsa | Type 2 Diabetes |
| Lee40        | 2015                | British Journal of Sports Medicine | Concussion |
| Wurzer41     | 2015                | Burns | Burns |
| Vega42       | 2014                | PLOS One | Pain |
| Wallace43    | 2014                | Journal of Opioid Management | Chronic Pain |
| El-Gayar44   | 2013                | Journal of Diabetes Science and Technology | Diabetes |

*Journals listed in field of ‘Medical Informatics’ in InCites Journal Citation Reports (Clarivate Analytics,45 or published in subsidiary journal of JMIR.46*
Reporting of search process

The majority (85%, 22/26) searched for apps from two or more app stores. The two most commonly searched app stores were Google Play (Android) and iOS (Apple). Other key aspects of reporting of the search process are presented in Figure 2. Most did not report the country of app store search or if any secondary search (e.g. using the website find.io) was conducted. App stores were searched with 1 to 9 keywords, with 17 of the included articles using 2 or more keywords relevant to the topic of interest. However, the duplicates removal process (i.e. same app found in more than one place) was not clearly reported in half of the included articles. While some articles clearly reported the exact date of searching app stores, some articles did not report the date (15%, 4/26) or the date was unclear (38%, 10/26) (e.g. stating the month of app store search only).

Screening criteria for app selection

The number of reviewers involved in conducting app screening and selection was poorly reported in over half of articles (58%, 15/26) and not reported in two articles (Figure 3). It was unclear if screening was done independently in the majority of included articles (83%, 20/24). Of the 13 articles that clearly reported the method used for screening apps for eligibility in the app stores, three articles22,38,39 used a two-step process involving screening of app store description and then downloading the apps for further screening. The other 10 studies screened apps mainly based on app store description. In articles with two reviewers assessing apps for eligibility, the process of reaching consensus to include apps was not reported in most (74%, 14/19). The reporting of using both lite and full versions of apps and the process of including multiplatform apps was usually unclear (Figure 3). Most of the included articles (73%, 19/26) excluded non-English apps.

Process of data extraction

While over half of the included articles (61%, 14/23) reported the number of data extractors per app, it was unclear if the data extractors were independent (Figure 4). Furthermore, most articles (67%, 16/24) did not report the process of resolving differences between data extractors. The
devices used (i.e. phone model) for downloading the apps and the operating systems were also not reported in most of the included articles (Figure 4).

**Appraisal of app inclusion of best practice content, and quality or usability**

A third of the included articles (32%, 8/25) evaluated the quality of the app contents against best practice content items (e.g. clinical guidelines) relevant to the health condition (Figure 5). Of the eight articles\(^7,22,30,34,37–40\) that evaluated the content quality, four\(^22,38,34,40\) clearly reported the adaptations made to the guidelines. Seven articles\(^7,20–22,24,30,35\) evaluated the quality and/or usability of the
included apps using tools such as the Mobile Apps Rating Scale (MARS). About half of the included articles (46%, 12/26) used customised tools to evaluate content quality or usability, of which, only five articles clearly reported the source of those critical appraisal tools.

Clinical recommendation and reporting of clinical efficacy of apps

Most articles (77%, 20/26) did not recommend apps for clinical use or report of clinical efficacy of the evaluated apps (Figure 6). One article recommended apps for clinical use and reported the clinical efficacy of apps. There were no examples of strength of recommendations (like GRADE).
Discussion

In the 26 included health app-focused reviews, many of the typical procedures of a systematic literature review as operationalised for app store search and app appraisal were not reported, or not clearly reported. Absences, lack of clarity or completeness of reporting were observed in: a priori review protocol registration or following a reporting guideline; the processes of screening apps in the app stores; data extraction details from the app store description or which type of device the app was downloaded to for data extraction; and appraisal tools used for assessing the clinical content, app quality or app usability. Further, the majority of articles did not explicitly report clinical efficacy of apps or recommendations for use in clinical setting.

As health apps are rapidly increasing in number and developed by multiple stakeholders, it is important that health apps are curated and appraised.47 Health app-focused reviews conducted with methodological rigour and published in the peer-reviewed literature should be the mechanism for identifying best in class apps for a health condition. Our study would suggest that this potential is not being met.

There were few articles that described robust screening processes operationalised for app store searches as an example for other reviewers. One example included two-step screening process based on app store description and subsequently downloading the apps to devices for detailed screening. Using this two-step approach increased rigour of screening against selection criteria and confirmed apps were available for download. Screening on app store alone increases the chance of including apps where the version is no longer usable in current operating system or only available with login. Handling of multiplatform apps during screening was poorly reported, and even when the number of multiplatform apps were reported, it was usually not clear which app version(s) were assessed. This is important as the usability evaluation of an app can vary depending on the operating system.13 Addressing these methodological shortcomings will require researchers to develop robust protocols and describe the methods of app screening in detail.

Only a small proportion of health app-focused reviews reported the source and any adaptation made to guidelines, frameworks or instruments used to evaluate app quality or usability and clinical quality of app content. Bespoke tools adapted from clinical guidelines were often used for assessing clinical quality of app content.7,22,30 Thus, it seems unlikely any ‘standard’ clinical content quality instrument could be developed and therefore explicit reporting of the criteria and source is necessary to increase confidence in recommending health apps. In a systematic review of 23 studies evaluating the quality of mHealth apps, Nouri et al noted high variability in quality assessment tools or criteria and identified seven classes of quality criteria, with 38 subclasses.48 The authors concluded this likely reflects different domains for evaluating quality depending on the use-case for the apps but also emphasised the requirement for careful and non-overlapping definition of quality criteria. Health app quality can now be assessed using the validated MARS13 – published in 2015 and cited almost 550 times. Only a third of the articles published since 2016 assessed app quality with the MARS although use in a higher proportion of health app-focused reviews is likely as maturing methodology and author rigour in app store search and app appraisal methodology occurs. There may be other domains for inclusion in health app-focused reviews. A recent scoping review found multiple safety concerns in health apps and recommended mandatory reporting of health app safety.49 Another recent study operationalised Health on the Net (HON) Foundation principles, designed for evaluating health web pages, and used these to evaluate popular medical apps.50 Future health app-focused reviews that aim to evaluate app content quality, usability or clinical quality appraisal should utilise the most appropriate, rigorous and contemporaneous evaluation criteria.

Some lack of rigour is partly attributable to lack of reporting guidelines for the conduct and reporting of health app-focused reviews. Another contributor is the fast-changing environment in
which health app-focused reviews are conducted. For example, only one study\(^7\) reported a priori review registration but it seems one review registration database (i.e. PROSPERO) no longer accepts health app-focused review protocols (personal communication to authors CC 6th January 2018 and DD 18th January 2018). However, a fast-changing environment is a further reason for explicit reporting as any health app-focused review will need to be reliably repeated.

Another example of unstable environment is searching app stores; unlike electronic journal databases, app stores are not designed for systematic searching and exporting of data.\(^5\) There are challenges in removing duplicate apps from app store searches (Android and iOS) and the number of hits from using a key word search in the Google Play store can provide endless results. To manage this issue, authors have to decide how to limit searches. One included study screened the first 5 pages of results and deemed this sufficient as no relevant apps were retrieved beyond the first five pages.\(^22\) An app review excluded from our final sample has noted reported the endless Google play app list as a limitation in their review.\(^52\) Further, app store searches can provide inconsistent results within a day and thus it is impossible to replicate the search strategy reliably. This is a key methodological challenge to be addressed for rigour in conduct and reporting of health app-focused reviews.

While this scoping review was planned, conducted, and reported according to relevant guidelines,\(^18\) there are limitations to this work. Included health app-focused reviews attested to being systematic, aiming to identify all relevant health apps available in a structured and repeatable manner. Our definition of ‘systematic’ (as defined in our inclusion criteria) may not have been sufficiently rigorous to select only systematic reviews. If authors used any of three key elements – the article had ‘systematic review’ in the title OR an a priori systematic search procedure was described in the methods section of the article OR the search process was described based on PRISMA standard flow diagram – then these articles met the inclusion criterion whether or not the authors were intentionally systematic. In contrast, we were very rigorous about defining and recording decision rules for our methodology to ensure consistent decision making about inclusion/exclusion, data extraction, and rating a data element as yes, no, unclear or uncertain. It is possible that the decision rules were unreasonably rigorous and therefore identified excessive numbers of data elements that were ‘not clear’. It is also possible that authors conducted health app-focused reviews in a more systematic manner but reported without precision or detail, contributing to the ‘no’ and ‘unclear’ ratings. Thus, the health app-focused reviews may be conducted more robustly than it appears from the sample in our research. While we registered a study protocol a priori, there was one departure from our study protocol. During the record identification stage, the secondary search of the Journal of Medical Internet Research (JMIR) databases retrieved a large number of potentially eligible articles for screening. Therefore secondary searches of other journals were not undertaken as planned due to efficiency requirements. This protocol deviation may have led to failure to identify potentially eligible high-quality health app-focused reviews. If included there may have been some increase in the number of data elements rated ‘yes’. A further caution is that including only health app-focused reviews for named/diagnosed health conditions excluded health app-focused reviews of other areas of health or wellness (e.g. diet, exercise, mental health apps) where conduct and reporting may be more rigorous.

Our operationalisation of PRISMA guidelines (Table 2) to health app-focused reviews may not have considered all important aspects of conduct and reporting. For example, we did not extract data on type of device (computer or mobile device) on which app store screening was undertaken. Any future development of health app-focused review methods could use our data extraction template, modified from the PRISMA literature search and synthesis guideline as a basis for conduct and reporting. Future researchers should also consider and extend our proposed methods for their current app environment to be as rigorous in conduct and reporting as possible.
People with health conditions and clinicians need ways to find apps to support and manage health conditions or augment clinical care, and have confidence these apps are fit for purpose. Some authors have provided useful summaries of how to find apps recommending use of online vendors (apps stores), online repositories (app repositories, online communities, news stories), and peer-reviewed literature. A recently published framework for finding apps recommends the peer-reviewed literature as first source to identify both high-quality evaluations of single apps and apps identified in app store search. While criteria for reporting evaluations of mobile health interventions in clinical research are already agreed to support the likelihood of finding a high-quality evaluation of a single app, our work suggests this is currently less likely for health app-focused reviews.

In a previous overview of the quality of health app-focused reviews, the variability in quality was observed. BinDhim et al. identified health app-focused reviews that discussed quality or quality-related aspects of health apps for consumers published between 2008 and 2013. The 10 identified articles were scored on 8 bespoke criteria with a mean score of 5 out of 8 and range 2 to 7. Similar to our study, 80% of articles did not include country of search, and 40% did not clearly explain if apps were evaluated based only on app store description or on app download. A more recent study evaluated 91 health app-focused review articles published between 2010 and 2016 based on 53 items that included the eight criteria from BinDhim et al and additional items based on scope of study, appraisal methods and on Cochrane review methodology. Overall, the assessment of reporting appropriate methodological steps had remarkable similarities to our findings. In line with the findings of these previous studies, our study confirms that the quality of reporting of health app-focused reviews, that is app store searches and app appraisal, could be improved.

The goal of high-quality health app-focused reviews in peer-reviewed literature can only be fully realized in co-operation with major app stores/vendors and regulators as infrastructure and regulatory factors are outside the control of those searching for apps. Firstly app stores need stable search environments, not driven by algorithms to promote apps in a non-transparent nature. Secondly app stores should require standardized app description fields for medical or health apps that explicitly report process and results of content, usability and efficacy evaluation. A coordinated approach will be needed to achieve this. These efforts would considerably advance the efforts to realize the potential for mHealth and improve health outcomes.

Conclusion

Of the 26 health app-focused reviews we identified as having features of a systematic approach, most did not report or incompletely reported many of the typical procedures of a systematic review as operationalized for health app-focused reviews. To improve the quality of the health app-focused review literature key actions include: (1) Identify an appropriate repository for registration of health app-focused reviews, or develop such a registry for health app-focused reviews. (2) Develop a consensus on best practice methods for the conduct of health app-focused review. (3) Agree and achieve transparency in reporting by authors, probably via consensus reporting standards.

Author Contributions

RG conceived and designed the study and acquired analyzed and interpreted the data; and drafted the work and revised it critically for intellectual content; and approved the final version; and agreed to be accountable for all aspects of the work. HD made substantial contributions to the conception and design of the study and acquired, analyzed and interpreted the data; and drafted the work and revised it critically for intellectual content; and approved the final version; and agreed to be accountable for all aspects of the work. BS made substantial contributions to design of the study and acquired, analyzed and interpreted the data; and revised the work critically for intellectual content; and approved the final version; and agreed to be accountable for all aspects of the work.
aspects of the work. JHS made substantial contributions to the conception and design of the study and acquired, analyzed and interpreted the data; and revised the work critically for intellectual content; and approved the final version; and agreed to be accountable for all aspects of the work.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: One author (RG) has a published article included in the data set, but was not involved in data extraction from this article. The authors have no other known conflicts of interest.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Rebecca Grainger https://orcid.org/0000-0001-9201-8678

Supplemental material

Supplemental material for this article is available online.

References

1. Statistica. Number of smartphone users worldwide from 2014 to 2020 (in billions). https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/, 2017 (accessed 13 July 2019).
2. Hood M, Wilson R, Corsica J, et al. What do we know about mobile applications for diabetes self-management? A review of reviews. *J Behav Med* 2016; 39(6): 981–994.
3. Slater H, Dear BF, Merolli MA, et al. Use of eHealth technologies to enable the implementation of musculoskeletal models of care: evidence and practice. *Best Pract Res Clin Rheumatol* 2016; 30(3): 483–502.
4. Marzano L, Bardill A, Fields B, et al. The application of mHealth to mental health: opportunities and challenges. *Lancet Psychiatry* 2015; 2(10): 942–948.
5. Sinclair C, Holloway K, Riley G, et al. Online mental health resources in rural Australia: clinician perceptions of acceptability. *J Med Internet Res* 2013; 15(9): e193.
6. Whitfield G and Williams C. If the evidence is so good – why doesn’t anyone use them? A national survey of the use of computerised cognitive behaviour therapy. *Behav Cogn Psychoth* 2004; 32(1): 57–65.
7. Machado GC, Pinheiro MB, Lee H, et al. Smartphone apps for the self-management of low back pain: a systematic review. *Best Pract Res Clin Rheumatol* 2016; 30(6): 1098–1109.
8. Federal Trade Commission. Mobile health apps interactive tool, https://www.ftc.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool (2016, accessed 13 July 2019).
9. National Institute for Health and Care Excellence. Evidence standards framework for digital health technologies, https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies (2019, accessed 16 August 2019).
10. Moher D, Schulz KF, Simera I, et al. Guidance for developers of health research reporting guidelines. *Plos Med* 2010; 7(2): e1000217.
11. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; 339: b2535–b2535.
12. Tong A, Flemming K, McNees E, et al. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Med Res Methodol* 2012; 12(1): 181.
13. Stoyanov SR, Hides L, Kavanagh DJ, et al. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015; 3(1): e27.
14. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; 64(4): 383–394.
15. Arksey H and O’Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Method* 2005; 8(1): 19–32.

16. Levac D, Colquhoun H and O’Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010; 5(1): 69.

17. PROSPERO international registry of prospective systematic reviews, https://www.crd.york.ac.uk/prospero/ (accessed 13 July 2019).

18. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018; 169(7): 467–473.

19. Covidence. https://www.covidence.org/home (accessed 13 July 2019).

20. Adam A, Hellig JC, Perera M, et al. ‘Prostate Cancer Risk Calculator’ mobile applications (Apps): a systematic review and scoring using the validated user version of the Mobile Application Rating Scale (uMARS). *World J Urol* 2018; 36(4): 565–573.

21. Athilingam P and Jenkins B. Mobile phone apps to support heart failure self-care management: integrative review. *JMIR Cardiol* 2018; 2(1): e10057.

22. Bhattarai P, Newton-John T and Phillips JL. Quality and usability of arthritic pain self-management apps for older adults: a systematic review. *Pain Med* 2017; 19(3): 471–484.

23. García-Cruz E, Romero-Otero J, Fode M, Alcaraz A, et al. The digital environment in men’s sexual disorders: a systematic review. *Actas Urológicas Españolas Engl Ed* 2018; 42(6): 365–374.

24. Jupp JC, Sultani H, Cooper CA, et al. Evaluation of mobile phone applications to support medication adherence and symptom management in oncology patients. *Pediatr Blood Cancer* 2018; 65(11): e27278.

25. Lee Y-L, Cui Y-Y, Tu M-H, et al. Mobile health to maintain continuity of patient-centered care for chronic kidney disease: content analysis of apps. *JMIR Mhealth Uhealth* 2018; 6(4): e10173.

26. Luo D, Wang P, Lu F, et al. Mobile apps for individuals with rheumatoid arthritis. *J Clin Rheumatology* 2019; 25(3): 133–141.

27. Linares-del Rey M, Vela-Desojo L and la Cuerda CR. Mobile phone applications in Parkinson’s disease: a systematic review. *Neurologia* 2018 (online). Neurologica English Edition 2019; 38–54.

28. Sangseng S. Mobile apps for chronic obstructive pulmonary disease: systematic review. *JMIR Mhealth Uhealth* 2017; 5(12): e195.

29. Martínez M, Park S, Maison I, et al. iOS appstore-based phone apps for diabetes management: potential for use in medication adherence. *JMIR Diabetes* 2017; 2(2): e12.

30. Rossi M and Bigi S. mHealth for diabetes support: a systematic review of apps available on the Italian market. *Mhealth* 2017; 3(5): 16–16.

31. Sobnath DD, Philip N, Kayyali R, et al. Features of a mobile support app for patients with chronic obstructive pulmonary disease: literature review and current applications. *JMIR Mhealth Uhealth* 2017; 5(2): e17.

32. Con D and Cruz P. Mobile phone apps for inflammatory bowel disease self-management: a systematic assessment of content and tools. *JMIR Mhealth Uhealth* 2016; 4(1): e13.
39. Fijacko N, Brzan P and Stiglic G. Mobile applications for type 2 diabetes risk estimation: a systematic review. *J Med Syst* 2015; 39(10): 124.

40. Lee H, Sullivan JS, Schneider AG, et al. Smartphone and tablet apps for concussion road warriors (team clinicians): a systematic review for practical users. *Brit J Sport Med* 2015; 49(8): 499–505.

41. Wurzer P, Parvizi D, Lumenta DB, et al. Smartphone applications in burns. *Burns* 2015; 41(5): 977–989.

42. de la Vega R and Miró J. mHealth: a strategic field without a solid scientific soul. A systematic review of pain-related apps. *Plos One*. 2014; 9(7): e101312.

43. Wallace S and Dhingra K. A systematic review of smartphone applications for chronic pain available for download in the United States. *J Opioid Manag*. 2014; 10(1): 63–68.

44. El-Gayar O, Timsina P, Nawar N, et al. Mobile applications for diabetes self-management: status and potential. *J Diabetes Sci Technology* 2013; 7(1): 247–262.

45. InCites Journal Citation Reports, https://clarivate.com/products/journal-citation-reports/. (accessed 13 July 2019).

46. Journal of Medical Internet Research, https://www.jmir.org/ (accessed 13 July 2019).

47. Wyatt JC. How can clinicians, specialty societies and others evaluate and improve the quality of apps for patient use? *BMC Med*. 2018; 16(1): 225.

48. Nouri R, Kalhori SR, Ghazisaeedi M, et al. Criteria for assessing the quality of mHealth apps: a systematic review. *J Am Med Inform Assn* 2018; 25(8): 1089–1098.

49. Akbar S, Coiera E and Magrabi F. Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. *J Am Med Inform Assn* 2020; 27(2): 330–340.

50. Ben-Mussa A and Paget AM. Popular apps on the medical category targeting patients and the general public in the United Kingdom: do they conform to the Health of the Net Foundation Principles? *Health Inform J* 2018; 24(3): 259–276.

51. Grundy QH, Wang Z and Bero LA. Challenges in assessing mobile health app quality a systematic review of prevalent and innovative methods. *Am J Prev Med* 2016; 51(6): 1051–1059.

52. Bright T and Pallawela D. Validated smartphone-based apps for ear and hearing assessments: a review. *JMRI Rehabil Assist Technol* 2016; 3(2): e13.

53. Paglialonga A, Lugo A and Santoro E. An overview on the emerging area of identification, characterization, and assessment of health apps. *J Biomed Inform* 2018; 83: 97–102.

54. Scott IA, Scuffham P, Gupta D, et al. Going digital: a narrative overview of the effects, quality and utility of mobile apps in chronic disease self-management. *Aust Health Rev* 2020; 44(1): 62–82.

55. Eysenbach C-E. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. *J Med Internet Res* 2011; 13(4): e126.

56. Agarwal S, LeFevre AE, Lee J, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ* 2016; 352: i1174.

57. BinDhim NF, Hawkey A and Trevena L. A systematic review of quality assessment methods for smartphone health apps. *Telemed E-health* 2015; 21(2): 97–104.