BMJ Open

Here's something I prepared earlier: A systematic review of the time to publication of cross-sectional reviews of smartphone health apps

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-039817
Article Type:	Original research
Date Submitted by the Author:	27-Apr-2020
Complete List of Authors:	Larsen, Mark; Black Dog Institute, Nicholas, Jennifer; Orygen; University of Melbourne, Centre for Youth Mental Health Han, J; Black Dog Institute, Lemon, Christopher; St Vincent's Hospital; UNSW Sydney, Faculty of Medicine Okun, Kelsi; Stanford University School of Humanities and Science Tye, Michelle; Black Dog Institute Wong, David; The University of Manchester, Centre for Health Informatics Wong, Iana; Black Dog Institute Wong, Quincy; Western Sydney University, School of Psychology Huckvale, Kit; Black Dog Institute,
Keywords:	Telemedicine < BIOTECHNOLOGY & BIOINFORMATICS, BIOTECHNOLOGY & BIOINFORMATICS, PUBLIC HEALTH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Here's something I prepared earlier: A systematic review of the time to publication of cross-sectional reviews of smartphone health apps

Mark E. Larsen, DPhil[1]
Jennifer Nicholas, PhD[2,3,4]
Jin Han, BSc, PhD[1]
Christopher Lemon, MPhil[5,6]
Kelsi Okun[7]
Michelle Tye, PhD[1]
David Wong, DPhil[8]
Iana Wong, MSc[1]
Quincy J. J. Wong, PhD[9]
Kit Huckvale, PhD[1]

[1] Black Dog Institute, UNSW Sydney, Australia
[2] Orygen, Victoria, Australia
[3] Centre for Youth Mental Health, The University of Melbourne, Australia
[4] Center for Behavioral Intervention Technologies, Northwestern University, Chicago, IL, United States
[5] St Vincent's Hospital, 390 Victoria Street, Sydney, Australia
[6] Faculty of Medicine, UNSW Sydney, Australia
[7] School of Humanities and Sciences, Stanford University, Palo Alto, CA, United States
[8] Division of Informatics, Imaging & Data Sciences, University of Manchester, United Kingdom
[9] School of Psychology, Western Sydney University, Australia

Corresponding author:
Mark Larsen
Black Dog Institute, UNSW Sydney, Hospital Road, Randwick NSW 2031, Australia
+61 (0)2 9382 8508
mark.larsen@blackdog.org.au

Abstract

Objectives: Across a range of health conditions, apps are increasingly valued as tools for supporting the delivery and co-ordination of healthcare. Research-led cross-sectional reviews of apps are a potential resource to inform app selection in face of uncertainties around content quality, safety, and privacy. However, these peer-reviewed publications only capture a snapshot of highly dynamic app stores and marketplaces. To determine the extent to which marketplace dynamics might impact the interpretation of app reviews, the current study sought to quantify the lag between the reported time of app assessment and publication of the results of these studies.

Design: Searches were conducted on MEDLINE, Embase and PsycINFO to identify published cross-sectional reviews of health, fitness, or wellness apps. Publication timeline metadata were extracted, allowing the primary outcome measure, the delay between app store search and manuscript publication, to be calculated. A secondary measure, the time between search and manuscript submission, was also calculated where possible.

Results: After screening, 136 relevant cross-sectional app review studies were analysed. The median time to publication was 431 days (approximately 14 months, range: 42-1054 days). The median time to submission was 269 days (approximately 9 months, range: 5-874 days). Studies which downloaded apps typically took longer to publish (p=0.010), however the number of apps reviewed did not impact the time to publication (p=0.964). Studies which recommended specific apps were not published more rapidly (p=0.998).

Conclusions: Most health app reviews present data that is at least a year out of date at the time of publication. Given the high rate of turnover of health apps in public marketplaces, it may not be appropriate, therefore, for these reviews to be presented as a resource concerning specific products for commissioners, clinicians and the public. Alternative sources of information may be better calibrated to the dynamics of the app marketplace.

Keywords

eHealth, mobile health, mHealth, health app, smartphone, cross sectional study, systematic review, research methodology

Strengths and limitations of this study

- This review considers over a decade of published cross-sectional reviews of health apps.
- The age of the review findings, at the time of publication, was determined and compared to the observed rate of change of the app stores.
- The time to journal submission was also calculated, where possible, providing an indication of the quickest possible time for results to be made publicly available to inform decisions.
- Heterogeneity across reviewed clinical and technical domains may impact publication time.

Introduction

Rationale

Smartphone applications (apps) are increasingly valued as tools for supporting the delivery and coordination of healthcare. Across a range of health conditions, there is growing evidence that app-based self-care interventions can be effective at reducing symptoms,[1] supporting self-management,[2,3] and promoting health behaviour change.[4,5] In 2017, half of surveyed Australian primary care doctors reported recommending apps to their patients at least once a month.[6] Across both physical[7] and mental health,[8] consumers either indicate interest in using health apps or report having already attempted to integrate apps into their health management. At a systems level, there is growing interest in the potential for digital health to enable value-based care that offers potential resource savings compared to face-to-face therapies. Examples of established initiatives include the Australian Federal Government's e-Mental Health Strategy,[9] which seeks to increase the accessibility and reach of mental health support while decreasing load on traditional services, and the state of California's Technology Suite Collaborative, which is harnessing digital technology to expand the capacity and capability of the county mental health systems, again to decrease burden on traditional care pathways.[10] Most recently, the National Health Service in England stated its intention for 'digitally-enabled primary and outpatient care [to] go mainstream' across the entire health system as part of its long term plan.[11]

Quality remains a key concern for healthcare providers seeking to integrate health apps into routine care. Content quality, safety, and privacy deficits have been identified in a wide range of health app categories.[12-15] Despite recent progress in clarifying regulatory requirements around 'software as a medical device' in the USA and Europe, only a small fraction of available apps either fall into a category that requires formal regulatory review or have been subject to experimental evaluation. Indeed, the number of health apps evaluated through randomised studies within the research literature is dwarfed by the numbers available to consumers.[16] These apps are typically made available within the same commercial marketplaces as apps for navigation, social media, and finance. This combination of prevalent quality issues and potentially large numbers of options presented without technical differentiation represents a major challenge for healthcare systems, clinicians, and consumers trying to select high quality, clinically-appropriate apps.

Research-led cross-sectional reviews of published apps ('app reviews') that critically appraise aspects of app quality and safety are a potential resource for consumers and practitioners, to inform rational app selection. Indeed, many app reviews either state, as aims, an intention to guide health professionals and consumers to the best apps for a given health condition or make recommendations targeting clinicians in discussion.[17] There is now a substantial collection of such reviews; our searches identified at least 149 such studies published between 2008 and 2019. Evidence of the potential impact on practice and policy of these cross-sectional studies include citations in clinical guidelines,[18] professional guidance concerning health app use[19,20] and design,[21] health system policy documents,[22,23] and expert[24] and intergovernmental[25] reports. Tools commonly used in app reviews, such as the Mobile App Rating Scale

(MARS),[26] were developed with the explicit goal of providing an app evaluation resource for use by health professionals (as well as researchers.)

However, there is a critical, and widely acknowledged limitation of these reviews: they are static snapshots of a dynamic environment. Within app stores, app updates, additions, removals, and search result list changes are common and unpredictable. In 2016, the dynamic nature of the two leading commercial app environments was quantified.[27] Tracking the search results for depression, bipolar disorder, and suicide prevention apps each day over 9 months, findings indicated that half of the Google Play search results change approximately every 4 months. Moreover, across both platforms, an app for depression became unavailable to download every 2.9 days.[27] These dynamic changes highlight the potential for information contained in cross-sectional reviews to become out-of-date, limiting its validity if used for the purposes of selecting and recommending specific health apps.

Objective

In order to explore the extent to which marketplace dynamics might impact the interpretation of research-led app reviews, the current study sought to quantify the lag between the reported time of app assessment and publication of the results of these studies. We assessed the impact on time-to-publication of specific features of the review process likely to act as a proxy for researcher workload, such as whether assessment involved downloading and reviewing app content. Finally, given that some app reviews explicitly state their intention to influence professional and patient behaviour, for example by recommending specific apps for use, we tested the hypothesis that these studies would be published more rapidly.

Methods

Literature search

We aimed to identify reported studies that performed cross-sectional analysis, assessments, or reviews of smartphone health and wellbeing apps. To identify studies, we developed a bespoke literature search strategy. Working separately, two reviewers (KH and JN) first performed exploratory searches of articles published in 2018 indexed by the MEDLINE citation database. Each reviewer used these searches to try to devise, respectively, a specificity-maximising and sensitivity-maximising search strategy (detailed in Supplementary File 1.)

In order to evaluate the performance of these alternative strategies, the results of each search (specificity-maximising n=78, sensitivity-maximising n=220) were independently screened by two reviewers (ML and JN). After reconciling any differences, screening yielded a binary partition of relevant/non-relevant studies for each search strategy. Subsequent comparison of these results indicated the overall suitability of a specificity-maximising approach. Individual discrepancies in included/excluded studies were also reviewed, yielding qualitative judgements about the likely contributions of different search terms to the observed results.

We used this information to devise a unified search strategy based on the original specificity-maximising approach. This strategy was then re-run on the original sample of 2018 citations to confirm that no relevant citations were omitted. In a final step, we broadened the search to include all years and selectively removed terms from the strategy to ascertain their impact on the final result set. Terms that did not alter the overall result count were discarded. Search results were also reviewed to ensure that studies already known to the reviewers were captured by the strategy. The final search strategy is detailed in Box 1.

(apps or applications or (app adj development)).ti.

AND

(smartphone? or mobile? or cell? or cellular? or (smart adj phone?) or iphone?).ab,ti.

AND

(review or (cross adj sectional) or content or quality or survey or assessment).ti. or (mobile adj2 rating adj scale?).ab. or (google adj play).ab.

AND

(appstore? or store? or marketplace? or (market adj place?) or apple or google or android or download$ or (app$ adj rating adj scale?)).ab.

Box 1. Optimised specificity-maximising search strategy.

Searches were run on MEDLINE, Embase and PsycINFO on 30 April 2019 and included all studies published between 2008 (on the basis that this was the year in which the first commercial app store was launched) and the search date. Search results were combined and deduplicated before screening.

Eligibility criteria

Search result titles and abstracts were reviewed against a standard set of inclusion criteria. Studies were retained if they i) focussed on a topic relating to health, fitness, or wellness (irrespective of whether the intended app users were consumers, carers, clinicians, researchers or some combination of these); ii) involved a cross-sectional search of an app store or library intended to generate a set of apps for subsequent examination; and iii) applied one or more methods to this set to evaluate either the metadata associated with each app (such as app store descriptions), the contents of each app, or both.

Study selection

Each study was screened by two out of three reviewers (of KH, ML and JN), working independently, with any discrepancies resolved by the third reviewer. Inter-rater agreement during initial screening of studies returned by the original search strategy (n=78) was calculated using Fleiss' Kappa at 0.78, indicating substantial agreement between reviewers.

Data extraction

Following screening, the full text of each included study was obtained for data extraction. Table 1 details the data elements that were extracted, if available. Coding aimed to quantitate the time taken to publish each study and identify proxy measures of the effort required for its completion (e.g. number of apps included in the study, and whether apps were downloaded as part of the review process.) A final parameter concerning whether study authors identified specific apps in their results or discussion (e.g. to recommend for, or caution against, use) was collected to investigate whether inclusion of such recommendations influenced publication speed. The data extraction schema was developed through an initial pilot phase in which n=60 studies were reviewed by three reviewers (KH, ML and JN) to identify relevant data items and confirm the feasibility of extraction.

Category	Item	Description
Publication timeline	Earliest search date	Earliest stated date that app stores were searched.
	Latest search date	If a date range is specified, the latest stated date that app stores were searched.
	Updated search date	Date of any subsequent searches or updates to the original search.
	Submission date	Date of manuscript submission to journal.
	Publication date	Earliest identified date at which the accepted, peer-reviewed manuscript is made available to the public – which may be an online-first/electronic pre-print. Pre-prints prior to manuscript acceptance were not considered.
	Dates imputed	A Boolean variable coded as: FALSE if both the search date and publication date were specified precisely, or TRUE if either date was imputed. Imputation was based on the mid-point of the specified date range, for example if a search month is specified rather than a search date, then the 15th day of the month was the imputed search date.
Review parameters	Number of apps reviewed	Number of apps reported for analysis, after any screening or filtering for relevance.
	Apps downloaded	Ordinal variable coded as ‘no apps downloaded’ (e.g. analysis was based on only app store metadata), ‘some apps downloaded’ (e.g. a targeted or random sample), or ‘all apps downloaded’.
	Apps recommended	A Boolean variable coded as: TRUE if individual apps were named and described in a manner which suggests or recommends their

		use, or FALSE otherwise.

Table 1. Data extracted from each reported study.

Extraction was completed in a two-phase process. In the first phase, we attempted to automatically extract study metadata, app store search and publication history dates. We used a heuristic text matching strategy to locate and excerpt relevant text from study full text, published study metadata and citation database records. Any matched text was used to pre-populate a standardised data extraction form for subsequent review.

In the second phase, each study was reviewed manually to verify automatically identified data and populate coding items not suitable for automation. Only studies which reported at least one search date (i.e. earliest search date) and a publication date were retained for analysis. Each study was reviewed by a pair of reviewers (from JH, KH, CL, ML, JN, KO, MT, IW, QW), each working independently. Any discrepancies were resolved by a third reviewer (from KH, ML and JN) not involved in the original review. Because data extraction included items with non-categorical assignments, we assessed inter-rater agreement using raw agreement (the proportion of scored data items where each reviewer pair assigned the same value). Overall agreement was 0.79 (n=1273/1618 data items) versus 0.83 (n=538/646) for those items extracted in the pilot phase. Considering extraction of study publication dates alone, being the data items intended to inform the primary analysis of study, agreement was 0.84 (n=478/570) versus 0.90 (n=207/230) extracted in the pilot phase.

Data analysis

The primary reported outcome is time to publication (TTP), calculated as the difference in days between the earliest search date and the date of publication. This window was justified on the basis that it reflects a conservative upper bound on the 'staleness' of information contained in any review at the earliest time it becomes accessible to a public audience. A secondary measure, the time to submission (TTS), was calculated as the difference between the earliest search date and the submission date. Descriptive statistics are reported for both TTP and TTS.

Data concerning review parameters (number of apps reviewed, whether apps were downloaded, and whether apps were recommended) are presented descriptively. Due to non-normalcy, non-parametric tests were used to assess the relationship between these parameters and the primary outcome, TTP. The correlation between TTP and the number of apps reviewed was measured using Spearman's correlation coefficient (ρ). The impact of downloading some or all apps on TTP was assessed using a Kruskal-Wallis test, with follow-up tests to identify differences between specific groups. Whether studies which recommended specific apps were published more rapidly was assessed using a Wilcoxon rank sum test.

Two sensitivity analyses were specified *a priori*. We anticipated that studies would exist where date information was reported only partially, for example, reporting only the month and year in which app searches were performed. In these cases (n=93), we imputed the date as the 15th of the stated month

or, if the authors reported a range of actual dates, selected a single date lying in the middle of this range. The first sensitivity analysis assessed the consequences of partial date reporting by comparing TTP for those studies where date imputation was and was not required.

A second sensitivity analysis aimed to explore the consequences of assuming that no important app changes occurred between the earliest and most recent app search/update dates. To do this, TTP was recalculated using the *last available* search date for each study (i.e. the latest reported value of 'earliest search date', 'latest search date', and 'updated search date') and compared to the original TTP measure. For both sensitivity analyses, TTP values were compared using Wilcoxon rank sum tests.

Finally, a *post hoc* exploration of the relationship between the two effort-related review parameters (number of apps reviewed and whether apps were downloaded) was conducted. The number of apps reviewed was compared across the sub-groups based on whether none, some, or all of the apps were downloaded in the review process and tested with a Kruskal-Wallis test. An additional metric, the review time per app, was defined as the time to publication divided by the number of apps reviewed. This metric was again compared across downloaded sub-groups, and compared using a Kruskal-Wallis test. All analyses were conducted using MATLAB 8.6. As this was an analysis of previously published literature, patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Results

Search and selection

Searches performed on 30 April 2019 yielded 439 study reports. After deduplication, screening and full-text review (summarised in Figure 1), 136 reports were included in the final analysis.

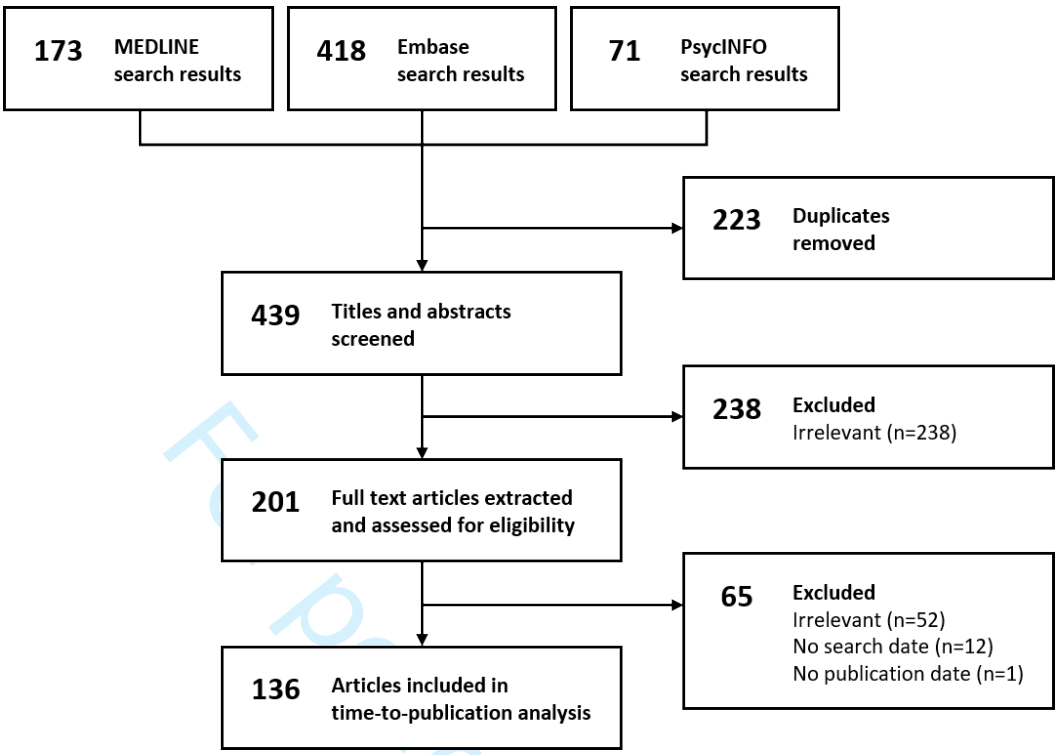


Figure 1. Study selection flowchart.

Time to publication and submission

The median time to publication, TTP, was 431 days (approximately 14 months, range: 42-1054 days) from the earliest search date. 100 papers reported a submission date, however in 8 cases the submission date was prior to the search date, therefore these papers were therefore excluded from the analysis. From the 92 remaining studies, the median time to submission, TTS, was 269 days (approximately 9 months, range: 5-874 days). The distribution of TTP and TTS are shown in Figure 2.

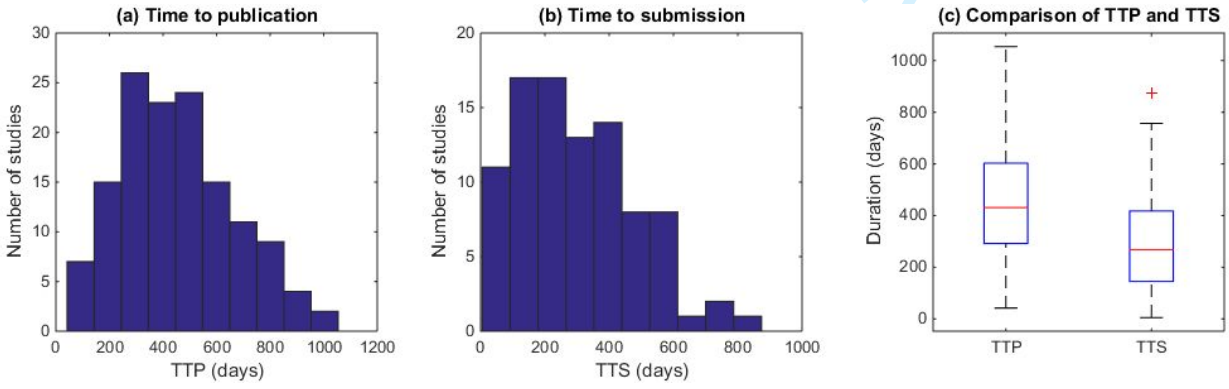


Figure 2. Frequency distributions for (a) the time to publication (TTP), and (b) time to submission (TTS), with (c) a side-by-side comparison.

Review parameters

The median number of apps reviewed in the 136 included studies was 52, although there was large variation between studies (range: 4-1806 apps). A near-zero correlation was found between the number of apps reviewed and TTP (ρ=-0.004, p=0.964).

Authors typically downloaded the apps for review, rather than relying on app store descriptions: 72.1% (n=98/136) papers indicated all apps were downloaded for review vs 17.6% (n=24/136) where no apps were downloaded. A targeted or random sample of apps were downloaded in 9.6% (n=13/136) of studies, and in 1 study it was not possible to determine whether or not apps were downloaded. There was a significant difference in TTP between the subgroups (p=0.010). Figure 3 shows the distribution of TTP for each sub-group. Follow-up tests identified a longer publication time when all apps compared to no apps were downloaded (median TTP: 476 days vs 292 days).

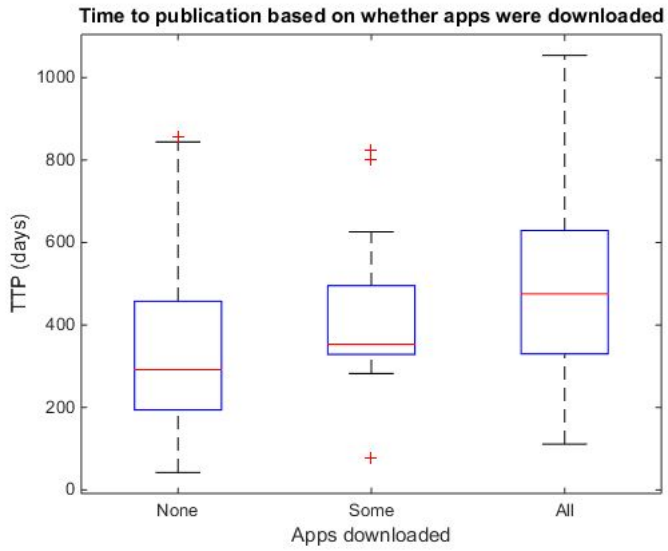


Figure 3. Distribution in the time to publication based on whether apps were downloaded as part of the review process.

Specific apps were named and recommended for use in 15.4% (n=21/136) of the reviewed studies. Studies which included recommendations for specific named apps were published marginally more quickly than other studies, but this difference was not significant (median: 425 vs 440 days, p=0.998, see Figure 4.)

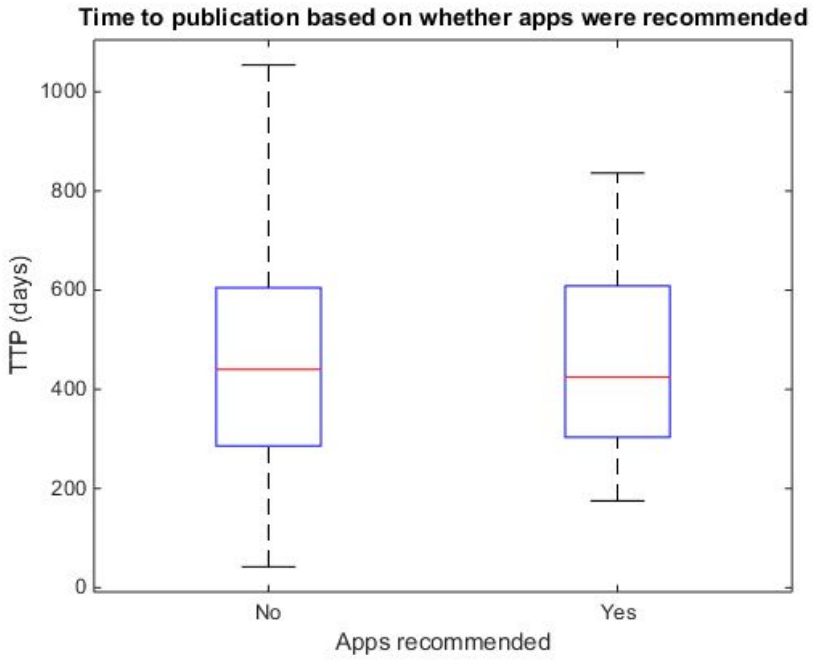


Figure 4. Distribution in the time to publication based on whether specific apps were named and recommended.

Sensitivity analyses

Two-thirds of the studies (92/136, 67.6%) did not specify an exact app store search date, and two (2/136, 1.5%) were not associated with a precise publication date. In combination, dates were imputed in 93 of the studies (68.4%). The difference in time to publication was not significantly different between papers with precise or imputed dates (457 vs 430 days respectively, p=0.648),

Approximately a quarter of the app reviews (37/136, 27.2%) reported a latest search date, and 5.1% (7/136) reported an updated search date. Using the latest of the three reported dates, the median time to publication reduced to 387 days, which did not reach significance for difference from the primary outcome (p=0.063).

Post-hoc analysis

As there appeared to be a relationship between whether apps were downloaded and the time to publication, but no such relationship for the number of apps reviewed in the studies, we conducted a *post-hoc* analysis to examine whether there was a relationship between these two review features. Figure 5 shows the variation in the number of apps reviewed, based on whether apps were downloaded as part of the review process, and a significant difference was found (p<0.001). Follow-up tests identified significantly more apps were reviewed when no or some apps were downloaded, compared with all apps being downloaded (median number of apps: 130, 175, and 43 respectively).

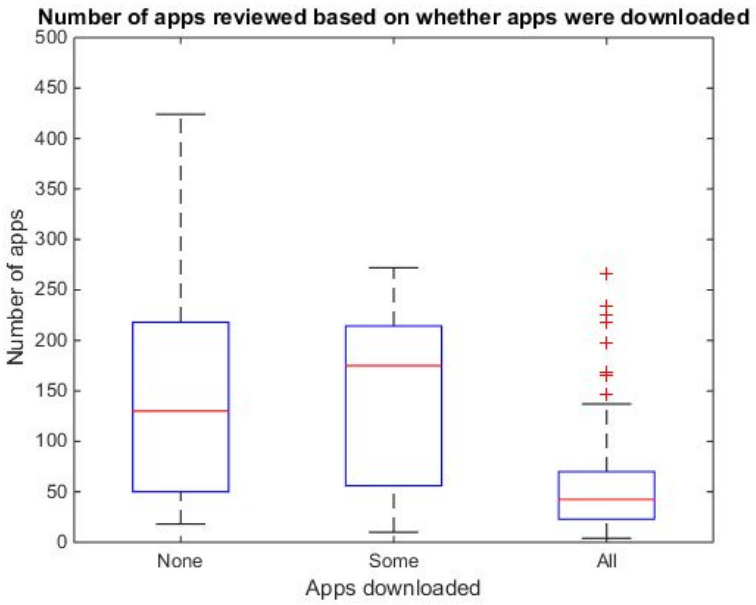


Figure 5. Variation in the number of apps reviewed, depending on whether the apps were downloaded as part of the review process. The y-axis has been truncated, and not all outlier values are shown.

When the time to publication was normalised by the number of apps included in the review, the review time per app was 8.2 days. There were significant differences between studies based on whether apps were downloaded ($p<0.001$) – with those studies where all apps were downloaded taking significantly longer (median 12.1 days/app) compared with the no downloaded (2.9 days/app) and some downloaded (4.2 days/app) apps.

Discussion

This study aimed to quantify the extent to which data presented in cross-sectional app reviews is up-to-date by examining the delay between the selection of apps for review and the time of publication. By the time that most app reviews become available for use, a considerable period of time has elapsed (median: 431 days, 14.2 months). This measure still exceeds one year (387 days, 12.7 months) when a more lenient measure of the recency of findings at the time of publication is used. This delay is not wholly attributable to factors outside the control of researchers, such as the peer review process: when the estimated time to submission was calculated to provide a crude metric of the time to conduct the study, excluding journal peer-review and editing processes, there was still substantial time between app search and manuscript submission (269 days, 8.8 months). The time taken to publish app reviews was influenced by the nature of the analysis, with findings indicating that reviews that downloaded apps for analysis took significantly longer to publish than those that did not. Surprisingly, the number of apps reviewed did not influence publication time, however a post-hoc analysis indicated papers that downloaded apps reviewed significantly fewer apps than those that did not.

Given previous research indicating a high rate of turnover in the app marketplace, with 50% of mental health search results changing within approximately three months,[27] the observed delay in publication raises questions about the validity of study findings at the time they become available to the research, clinical, and broader community, particularly where such reviews focus on recommendations concerning specific products. Reviews may recommend the use of apps which are no longer supported by the developers, have been withdrawn from the app stores or, conversely, have been updated substantially since the review. Contrary to our hypothesis, we did not find that studies making specific recommendations had a shorter time to publication than other reviews. Recommendations for specific apps in published reviews cannot, therefore, automatically be considered reliable. The delay in publication may also mean that more recent, potentially high quality, apps are not made known to the research or clinical communities.

There are a number of potential strategies that could mitigate this issue. The first is for review authors to refresh their results prior to publication. Pre-publication update is a standard practice in systematic reviews. For example, the Cochrane Collaboration will not publish reviews unless the most recent search date is less than twelve months (and ideally less than six months) old.[28] As part of efforts to improve the quality of cross-sectional app reviews,[29] editors and peer reviewers should consider at least asking for justification where there is a long period between search and publication. The practicality of this solution must nevertheless consider review-specific factors that may affect the time to publication, the feasibility of update, and whether the review *intends* to guide clinical and public uses.

The second potential strategy is for authors to adjust the numbers of apps incorporated in their reviews. The relationship between downloading of apps and the number of apps reviewed may indicate that study authors hold some shared perceptions about what represents a 'publishable unit' of work. This may be achieved by either downloading a smaller number of apps, or by reviewing the app store descriptions for a larger number of apps. Both appear to result in publication in approximately one year (± three months). While both have merit, it seems likely that studies that scrutinise app content directly are likely to yield richer insights than those relying on summary information presented in app stores for the purposes of marketing. However, with the longer time to publication associated with downloading, it may be appropriate to focus on a smaller sample of the most popular, most used, or top-ranked apps, which can be published more quickly.

A third possible strategy is to remove app assessment and review from the academic sphere, to organisations whose resources are not subject to the constraints of the academic publication process and are, at least in principle, resourced to be able to respond to app dynamics such as update and withdrawal. Indeed, continuous app scanning and review approaches have now been adopted by a number of health organisations, including the NHS Apps Library[30] and the American Psychiatric Association.[31] However, despite the intention to provide continuous review, it is unclear how often reviews and recommendations that appear within these portals are actually updated. Further, even within app portals, the large number of available apps often necessitates that thoroughness be balanced with expediency,[32] potentially still limiting broad utility of such resources.

Importantly, we do not suggest that our findings imply that cross-sectional assessments of health apps have no utility. App reviews may not be the optimal source of timely information about the function and quality of specific apps, but research-based methods are well suited to identifying and providing unbiased estimates of the nature and extent of thematic issues affecting specific populations of apps. Such insights can and have guided systemic responses to app quality problems. Research-led studies have arguably been important both in identifying *new* issues affecting apps, particularly where identification of issues involves complex exploratory and technical analysis,[33] and in devising systematic strategies for their identification, such as MARS. Unless specifically resourced to do so, it seems unlikely that continuous scanning programmes and app portals will be able to fulfill this discovery function.

Limitations

While this study examined app reviews across all health domains, to characterise the publication delay generally for the mobile health field of research, it did not examine differences across specific sub-domains. It is possible that different outcomes would be observed for different health conditions, for example due to resourcing availability/constraints across clinical domains, or due to differences in self-management approaches for different conditions. These differences were not considered in the current study due to substantial observed heterogeneity in the scope of reviews: some considered only technical domains of app quality (for example, data privacy), some considered broad categories (for example, mental health or physical health), and some considered specific conditions. Furthermore, the databases selected for the literature searches may not have provided complete coverage across all reviews focused on health and wellbeing domains (for example, those reported in allied health publications) or technical domains (for example, those focusing on data privacy and security).

The time to submission was calculated in addition to the primary outcome, time to publication, to provide an approximate measure of the time to conduct the review process. However, this is only a crude estimate as it cannot account for manuscripts submitted to multiple journals prior to acceptance.

Finally, given the research question addressed by this study, we acknowledge the time taken to conduct and publish this review. However, given the relatively slower pace of research compared to app development, these findings from 2019 are likely to still hold relevance.

Conclusions

The majority of health app reviews present data that is at least a year out of data at the time of publication. Given the high rate of observed turnover of health apps in public marketplaces, it may not be appropriate, therefore, for these reviews to be presented as a resource concerning specific products for commissioners, clinicians and the public. Authors of such reviews should, where possible, take steps to minimise the delay to publication, update their results prior to publication and consider whether making specific product recommendations is appropriate. App reviews may nevertheless fulfill important functions to identify novel and thematic issues and guide policy and systemic responses to health app quality and safety. App users should consider alternative sources of information about apps

that are better calibrated to the dynamics of the app marketplace, such as continuous scanning services offered by dedicated health app portals.

Statement of contributorship

ML and KH conceived the study and are guarantors. JN and KH devised the search strategies and performed the searches. ML, JN and KH performed initial screening and ML collated the results. KH developed and executed the automatic data extraction process. KH prepared the manual data extraction tools, allocated reviewer tasks and collated results. All authors performed data extraction. ML performed principal data analysis. ML, JN and KH wrote the first draft of the manuscript. All authors provided comments to and/or contributed revisions to the draft manuscript. All authors reviewed and approved the final manuscript version.

Declaration of interests

KH, ML, JN are authors of studies that were included in this review. They declare no other competing financial or non-financial interests. No specific funding was received to undertake this study. All other authors declare no competing financial or non-financial interests.

References

1. Kitsiou S, Pare G, Jaana M, Gerber B. Effectiveness of mhealth interventions for patients with diabetes: An overview of systematic reviews. *PLoS One* 2017; **12**(3): e0173160.

2. Coorey GM, Neubeck L, Mulley J, Redfern J. Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: Systematic review with meta-synthesis of quantitative and qualitative data. *Eur J Prev Cardiol* 2018; **25**(5): 505-21.

3. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: A systematic review. *J Med Internet Res* 2016; **18**(5): e97.

4. Han M, Lee E. Effectiveness of mobile health application use to improve health behavior changes: A systematic review of randomized controlled trials. *Healthc Inform Res* 2018; **24**(3): 207-26.

5. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. *J Med Internet Res* 2016; **18**(11): e287.

6. Byambasuren O, Beller E, Glasziou P. Current knowledge and adoption of mobile health apps among australian general practitioners: Survey study. *JMIR Mhealth Uhealth* 2019; **7**(6): e13199.

7. Boyle L, Grainger R, Hall RM, Krebs JD. Use of and beliefs about mobile phone apps for diabetes self-management: Surveys of people in a hospital diabetes clinic and diabetes health professionals in new zealand. *JMIR Mhealth Uhealth* 2017; **5**(6): e85.

8. Rubanovich CK, Mohr DC, Schueller SM. Health app use among individuals with symptoms of depression and anxiety: A survey study with thematic coding. *JMIR Ment Health* 2017; **4**(2): e22.

9. Australian Government Department of Health and Ageing. E-Mental health strategy for Australia. Canberra, Australia: Commonwealth of Australia, 2012.

10. California Mental Health Services Authority. "The Technology Suite" Driving access to behavioral health care thru innovation. 2018.

11. NHS England. The NHS Long Term Plan. London, United Kingdom: HM Government, 2019.

12. Huckvale K, Adomaviciute S, Prieto JT, Leow MK, Car J. Smartphone apps for calculating insulin dose: a systematic assessment. *BMC Med* 2015; **13**: 106.

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

14. Larsen M, Nicholas J, Christensen H. A Systematic Assessment of Smartphone Tools for Suicide Prevention. *PLoS One* 2016; **11**(4): e0152285.

15. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile Apps for Bipolar Disorder: A Systematic Review of Features and Content Quality. *J Med Internet Res* 2015; **17**(8): e198.

16. Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res* 2013; **15**(11): e247.

17. Thornton L, Quinn C, Birrell L, et al. Free smoking cessation mobile apps available in australia: A quality review and content analysis. *Australian and New Zealand Journal of Public Health* 2017; **41**(6): 625-30.

18. Goodwin GM, Haddad PM, Ferrier IN, et al. Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacol* 2016; **30**(6): 495-553.

19. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist. *Clinical Medicine Journal (London)* 2015; **15**(6): 519-21.

20. Torous JB, Chan SR, Gipson SY-MT, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatric Services* 2018; **69**(5): 498-500.

21. Stephens H, Uccellini M, McKay F. Guidelines for creating healthy living apps. Melbourne, Australia: Dialog Consultking, 2015.

22. Centers for Disease Control and Prevention. Best practices user guide: Health communications in tobacco prevention and control. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2018.

23. Rose KJ, Petrut C, L'Heveder R, de Sabata S. IDF Europe's position on mobile applications in diabetes. *Diabetes Res Clin Pract* 2019; **149**: 39-46.

24. Ronchi E, Woskie LR, Milstein JA. Mobile technology-based services for global health and wellness: opportunities and challenges. Boston, USA: Organisation for Economic Cooperation and Development (OECD), 2017.

25. World Bank Group. Empowering households and individuals to co-produce positive health outcomes for dignified, person-centered care amidst demographic change Washington, DC, USA: World Bank, 2018.

26. 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 and uHealth* 2015; **3**(1): e27-e.

27. Larsen ME, Nicholas J, Christensen H. Quantifying App Store Dynamics: Longitudinal Tracking of Mental Health Apps. *JMIR Mhealth Uhealth* 2016; **4**(3): e96.

28. Tariq S, Akhtar N, Afzal H, et al. A novel co-training-based approach for the classification of mental illnesses using social media posts. *IEEE Access* 2019; **7**: 166165-72.

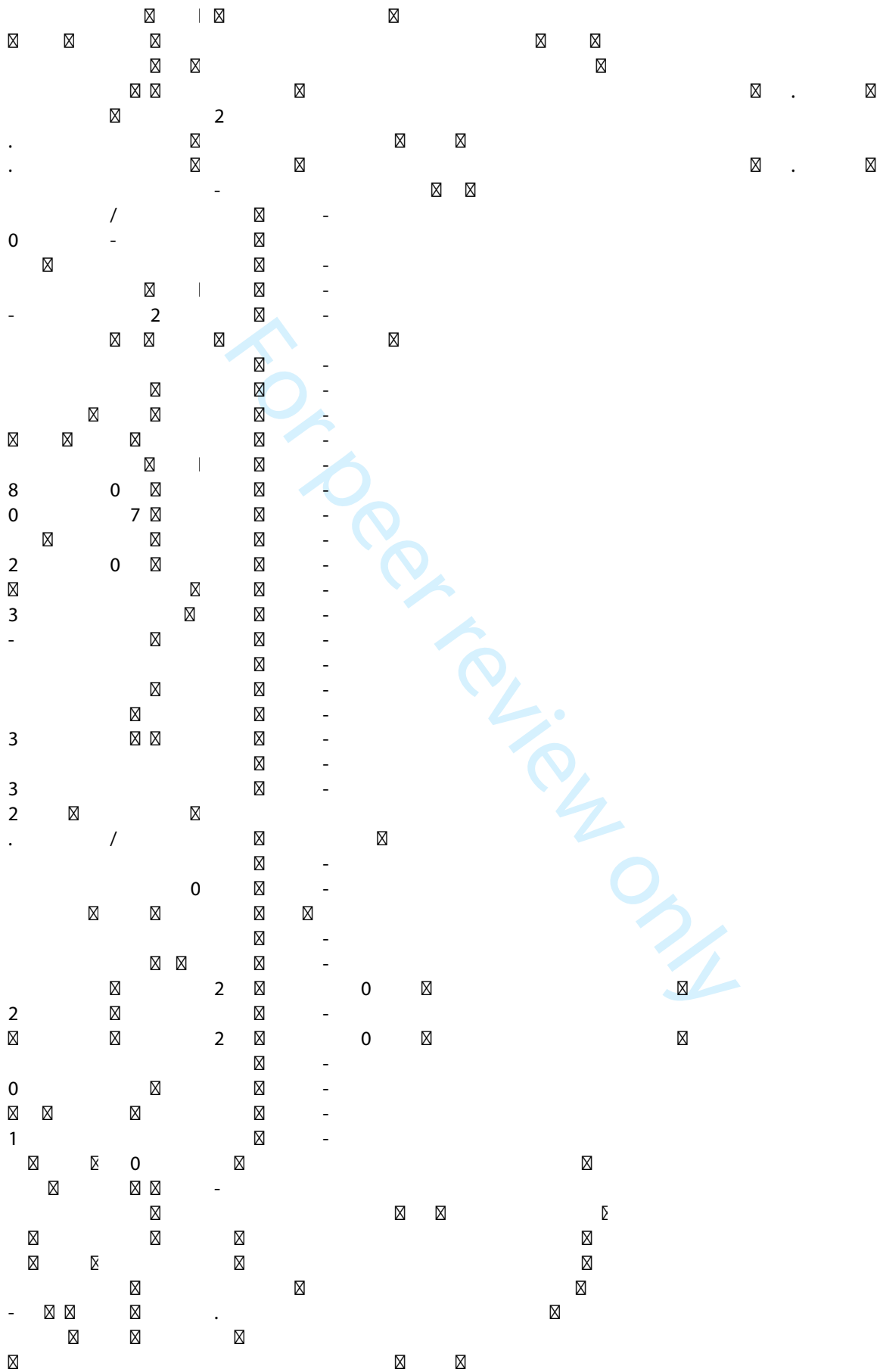
29. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. *Telemed J E Health* 2015; **21**(2): 97-104.

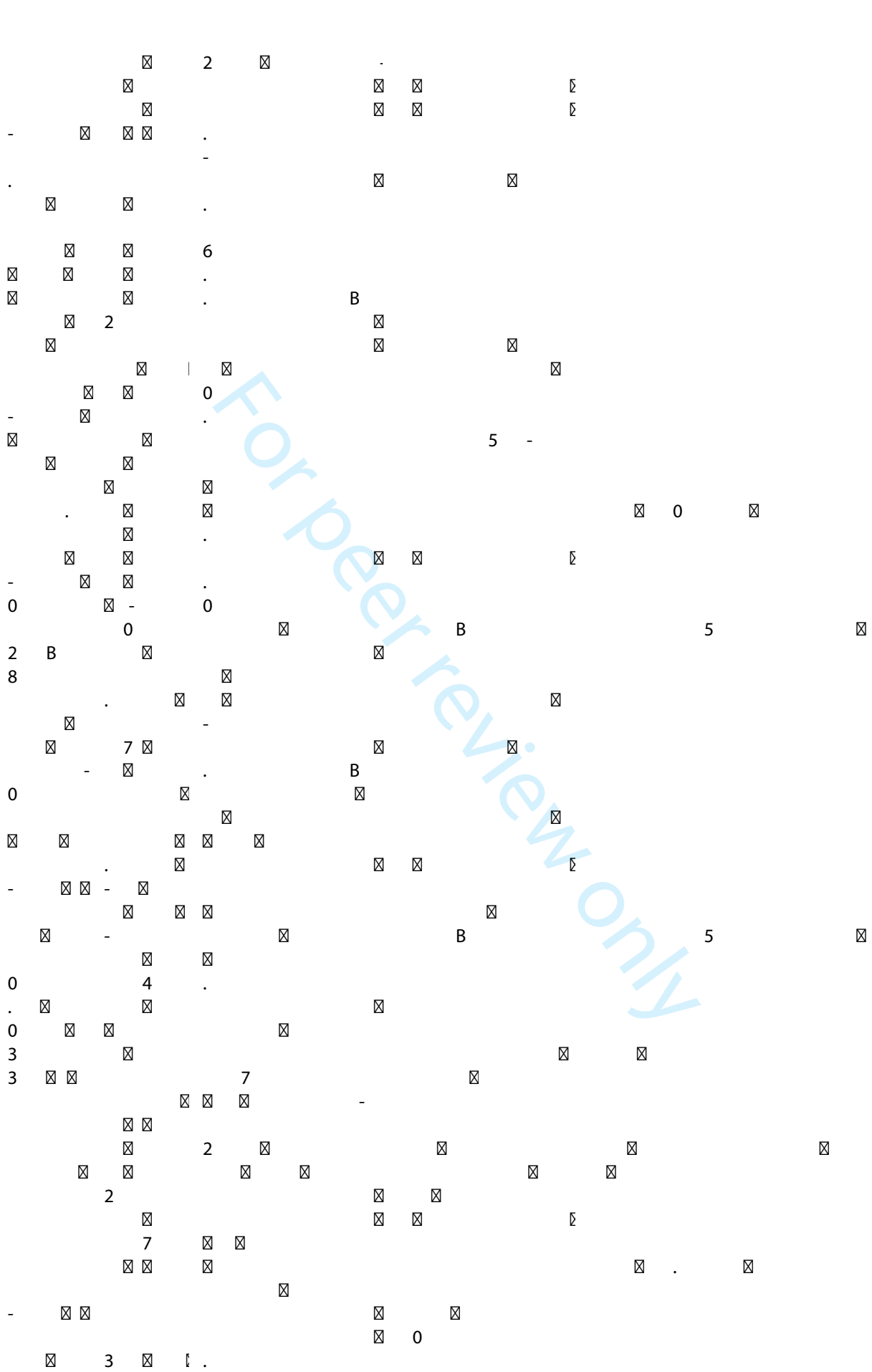
30. NHS. NHS Apps Library. https://www.nhs.uk/apps-library/ (accessed 25 February 2020).

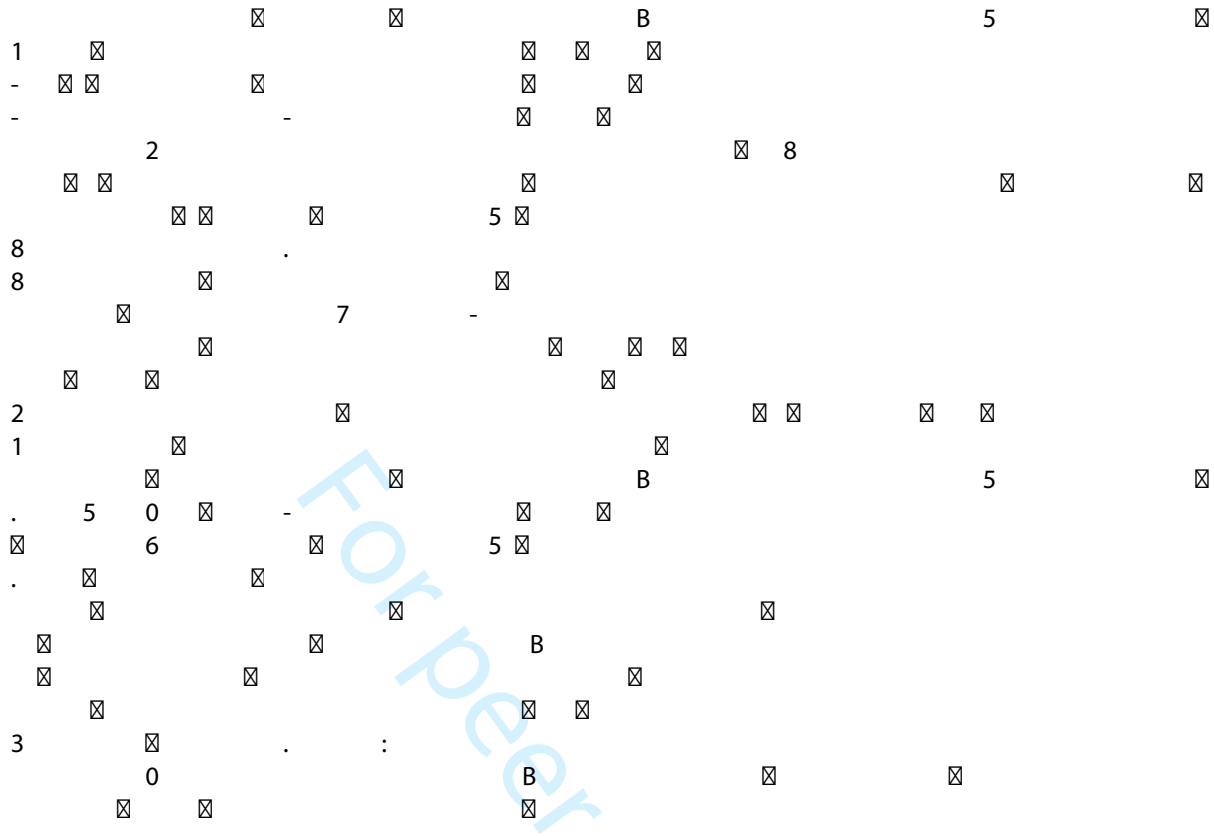
31. American Psychiatric Association. App evaluation model. https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model (accessed 25 February 2020).

32. Meek T. Tom's digital disruptors: testing times for apps. Digital Health; 2015.

33. Huckvale K, Torous J, Larsen ME. Assessment of the Data Sharing and Privacy Practices of Smartphone Apps for Depression and Smoking Cessation. *JAMA Netw Open* 2019; **2**(4): e192542.







PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	N/A
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	N/A
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	N/A
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	9
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	9
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	S1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	N/A
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	N/A
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	12
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	13
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	17

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

Here's something I prepared earlier: A review of the time to publication of cross-sectional reviews of smartphone health apps

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-039817.R1
Article Type:	Original research
Date Submitted by the Author:	09-Nov-2020
Complete List of Authors:	Larsen, Mark; Black Dog Institute, Nicholas, Jennifer; Orygen; University of Melbourne, Centre for Youth Mental Health Han, J; Black Dog Institute, Lemon, Christopher; St Vincent's Hospital; UNSW Sydney, Faculty of Medicine Okun, Kelsi; Stanford University School of Humanities and Science Tye, Michelle; Black Dog Institute Wong, David; The University of Manchester, Centre for Health Informatics Wong, Iana; Black Dog Institute Wong, Quincy; Western Sydney University, School of Psychology Huckvale, Kit; Black Dog Institute,
Primary Subject Heading:	Health informatics
Secondary Subject Heading:	Public health
Keywords:	Telemedicine < BIOTECHNOLOGY & BIOINFORMATICS, BIOTECHNOLOGY & BIOINFORMATICS, PUBLIC HEALTH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Here's something I prepared earlier: A review of the time to publication of cross-sectional reviews of smartphone health apps

Mark E. Larsen, DPhil[1]
Jennifer Nicholas, PhD[2,3,4]
Jin Han, BSc, PhD[1]
Christopher Lemon, MPhil[5,6]
Kelsi Okun[7]
Michelle Tye, PhD[1]
David Wong, DPhil[8]
Iana Wong, MSc[1]
Quincy J. J. Wong, PhD[9]
Kit Huckvale, PhD[1]

[1] Black Dog Institute, UNSW Sydney, Australia
[2] Orygen, Victoria, Australia
[3] Centre for Youth Mental Health, The University of Melbourne, Australia
[4] Center for Behavioral Intervention Technologies, Northwestern University, Chicago, IL, United States
[5] St Vincent's Hospital, 390 Victoria Street, Sydney, Australia
[6] Faculty of Medicine, UNSW Sydney, Australia
[7] School of Humanities and Sciences, Stanford University, Palo Alto, CA, United States
[8] Division of Informatics, Imaging & Data Sciences, University of Manchester, United Kingdom
[9] School of Psychology, Western Sydney University, Australia

Corresponding author:
Mark Larsen
Black Dog Institute, UNSW Sydney, Hospital Road, Randwick NSW 2031, Australia
+61 (0)2 9382 8508
mark.larsen@blackdog.org.au

Abstract

Objectives: Across a range of health conditions, apps are increasingly valued as tools for supporting the delivery and co-ordination of healthcare. Research-led cross-sectional reviews of apps are a potential resource to inform app selection in face of uncertainties around content quality, safety, and privacy. However, these peer-reviewed publications only capture a snapshot of highly dynamic app stores and marketplaces. To determine the extent to which marketplace dynamics might impact the interpretation of app reviews, the current study sought to quantify the lag between the reported time of app assessment and publication of the results of these studies.

Design: Searches were conducted on MEDLINE, Embase and PsycINFO to identify published cross-sectional reviews of health, fitness, or wellness apps. Publication timeline metadata were extracted, allowing the primary outcome measure, the delay between app store search and manuscript publication, to be calculated. A secondary measure, the time between search and manuscript submission, was also calculated where possible.

Results: After screening, 136 relevant cross-sectional app review studies were analysed. The median time to publication was 431 days (approximately 14 months, range: 42-1054 days). The median time to submission was 269 days (approximately 9 months, range: 5-874 days). Studies which downloaded apps typically took longer to publish (p=0.010), however the number of apps reviewed did not impact the time to publication (p=0.964). Studies which recommended specific apps were not published more rapidly (p=0.998).

Conclusions: Most health app reviews present data that is at least a year out of date at the time of publication. Given the high rate of turnover of health apps in public marketplaces, it may not be appropriate, therefore, for these reviews to be presented as a resource concerning specific products for commissioners, clinicians and the public. Alternative sources of information may be better calibrated to the dynamics of the app marketplace.

Keywords

eHealth, mobile health, mHealth, health app, smartphone, cross sectional study, research methodology

Strengths and limitations of this study

- This review considers over a decade of published cross-sectional reviews of health apps.
- The age of the review findings, at the time of publication, was determined and compared to the observed rate of change of the app stores.
- The time to journal submission was also calculated, where possible, providing an indication of the quickest possible time for results to be made publicly available to inform decisions.
- Heterogeneity across reviewed clinical and technical domains may impact publication time.

Introduction

Rationale

Smartphone applications (apps) are increasingly valued as tools for supporting the delivery and coordination of healthcare. Across a range of health conditions, there is growing evidence that app-based self-care interventions can be effective at reducing symptoms,[1] supporting self-management,[2,3] and promoting health behaviour change.[4,5] In 2017, half of surveyed Australian primary care doctors reported recommending apps to their patients at least once a month.[6] Across both physical[7] and mental health,[8] consumers either indicate interest in using health apps or report having already attempted to integrate apps into their health management. At a systems level, there is growing interest in the potential for digital health to enable value-based care that offers potential resource savings compared to face-to-face therapies. Examples of established initiatives include the Australian Federal Government's e-Mental Health Strategy,[9] which seeks to increase the accessibility and reach of mental health support while decreasing load on traditional services, and the state of California's Technology Suite Collaborative, which is harnessing digital technology to expand the capacity and capability of the county mental health systems, again to decrease burden on traditional care pathways.[10] Most recently, the National Health Service in England stated its intention for 'digitally-enabled primary and outpatient care [to] go mainstream' across the entire health system as part of its long term plan.[11]

Quality remains a key concern for healthcare providers seeking to integrate health apps into routine care. Content quality, safety, and privacy deficits have been identified in a wide range of health app categories.[12-15] Despite recent progress in clarifying regulatory requirements around 'software as a medical device' in the USA and Europe, only a small fraction of available apps either fall into a category that requires formal regulatory review or have been subject to experimental evaluation. Indeed, the number of health apps evaluated through randomised studies within the research literature is dwarfed by the numbers available to consumers.[16] These apps are typically made available within the same commercial marketplaces as apps for navigation, social media, and finance. This combination of prevalent quality issues and potentially large numbers of options presented without technical differentiation represents a major challenge for healthcare systems, clinicians, and consumers trying to select high quality, clinically-appropriate apps.

Research-led cross-sectional reviews of published apps ('app reviews') that critically appraise aspects of app quality and safety are a potential resource for healthcare practitioners patients and the public when choosing an appropriate health app. Indeed, many app reviews either state, as aims, an intention to guide health professionals and consumers to the best apps for a given health condition or make recommendations targeting clinicians in discussion.[17] Evidence of the potential impact on clinical practice and policy of these cross-sectional studies include citations in clinical guidelines,[18] professional guidance concerning health app use[19,20] and design,[21] health system policy documents,[22,23] and expert[24] and intergovernmental[25] reports. Tools commonly used in app reviews, such as the Mobile App Rating Scale (MARS),[26] were developed with the explicit goal of providing an app evaluation resource for use by health professionals (as well as researchers.)

There is now a substantial collection of such reviews; our searches identified at least 149 such studies published between 2008 and 2019. However, there is a critical, and widely acknowledged limitation of these reviews: they are static snapshots of a volatile environment. Within app stores, app updates, additions, removals, and search result list changes are common and unpredictable, and may be further compounded by different app listings and availability in different jurisdictions. In 2016, the dynamic nature of the two leading commercial app environments was quantified.[27] Tracking the search results for depression, bipolar disorder, and suicide prevention apps each day over 9 months, findings indicated that half of the Google Play search results change approximately every 4 months. Moreover, across both platforms, an app for depression became unavailable to download every 2.9 days.[27] These dynamic changes highlight the potential for information contained in cross-sectional reviews to become out-of-date, limiting its validity if used for the purposes of selecting and recommending specific health apps.

Objective

In order to explore the extent to which marketplace dynamics might impact the interpretation of research-led app reviews, the current study sought to quantify the lag between the reported time of app assessment and publication of the results of these studies. We assessed the impact on time-to-publication of specific features of the review process likely to act as a proxy for researcher workload, such as whether assessment involved downloading and reviewing app content. Finally, given that some app reviews explicitly state their intention to influence professional and patient behaviour, for example by recommending specific apps for use, we tested the hypothesis that these studies would be published more rapidly. This review focuses on published reviews of health and wellbeing apps which could be downloaded onto a smartphone (typically, but not exclusively, native apps via the Apple App Store or Google Play Store), without limitation on app functionality. Therefore, within this study, no constraints were placed on what review authors defined as health and wellbeing apps, as long as the review focused on a topic related to fitness, wellness, or health, and no restrictions were placed on health domain. Apps available through curated third-party lists or libraries were also considered.

Methods

Literature search

We aimed to identify reported studies that performed cross-sectional analysis, assessments, or reviews of smartphone health and wellbeing apps. To identify studies, we developed a bespoke literature search strategy. Working separately, two reviewers (KH and JN) first performed exploratory searches of articles published in 2018 indexed by the MEDLINE citation database. Each reviewer used these searches to try to devise, respectively, a specificity-maximising and sensitivity-maximising search strategy (detailed in Supplementary File 1.)

In order to evaluate the performance of these alternative strategies, the results of each search (specificity-maximising n=78, sensitivity-maximising n=220) were independently screened by two

reviewers (ML and JN). After reconciling any differences, screening yielded a binary partition of relevant/non-relevant studies for each search strategy. Subsequent comparison of these results indicated the overall suitability of a specificity-maximising approach. Individual discrepancies in included/excluded studies were also reviewed, yielding qualitative judgements about the likely contributions of different search terms to the observed results.

We used this information to devise a unified search strategy based on the original specificity-maximising approach. This strategy was then re-run on the original sample of 2018 citations to confirm that no relevant citations were omitted. In a final step, we broadened the search to include all years and selectively removed terms from the strategy to ascertain their impact on the final result set. Terms that did not alter the overall result count were discarded. Search results were also reviewed to ensure that studies already known to the reviewers were captured by the strategy. The final search strategy is detailed in Box 1.

(apps or applications or (app adj development)).ti.

AND

(smartphone? or mobile? or cell? or cellular? or (smart adj phone?) or iphone?).ab,ti.

AND

(review or (cross adj sectional) or content or quality or survey or assessment).ti. or (mobile adj2 rating adj scale?).ab. or (google adj play).ab.

AND

(appstore? or store? or marketplace? or (market adj place?) or apple or google or android or download$ or (app$ adj rating adj scale?)).ab.

Box 1. Optimised specificity-maximising search strategy.

Searches were run on MEDLINE, Embase and PsycINFO on 30 April 2019 and included all studies published between 2008 (on the basis that this was the year in which the first commercial app store was launched) and the search date. Search results were combined and deduplicated before screening.

Eligibility criteria

Search result titles and abstracts were reviewed against a standard set of inclusion criteria. Studies were retained if they i) focussed on a topic relating to health, fitness, or wellness (irrespective of whether the intended app users were consumers, carers, clinicians, researchers or some combination of these); ii) involved a cross-sectional search of an app store or library intended to generate a set of apps for subsequent examination; and iii) applied one or more methods to this set to evaluate either the metadata associated with each app (such as app store descriptions), the contents of each app, or both.

Study selection

Each study was screened by two out of three reviewers (of KH, ML and JN), working independently, with any discrepancies resolved by the third reviewer. Inter-rater agreement during initial screening of studies returned by the original search strategy (n=78) was calculated using Fleiss' Kappa at 0.78, indicating substantial agreement between reviewers.

Data extraction

Following screening, the full text of each included study was obtained for data extraction. Table 1**Error! Reference source not found.** details the data elements that were extracted, if available. Coding aimed to quantitate the time taken to publish each study and identify proxy measures of the effort required for its completion (e.g. number of apps included in the study, and whether apps were downloaded as part of the review process.) A final parameter concerning whether study authors identified specific apps in their results or discussion (e.g. to recommend for, or caution against, use) was collected to investigate whether inclusion of such recommendations influenced publication speed. The data extraction schema was developed through an initial pilot phase in which n=60 studies were reviewed by three reviewers (KH, ML and JN) to identify relevant data items and confirm the feasibility of extraction.

Table 1. Data extracted from each included study.

Category	Item	Description
Publication timeline	Earliest search date	The earliest date authors report searching the app stores.
	Latest search date	If app store searches were conducted over a period of time, the latest date authors report searching the app stores.
	Updated search date	If subsequent app store searches were performed, for example to update the initial search results, the latest date authors report conducting the updated search.
	Submission date	Date of manuscript submission to journal.
	Publication date	Earliest identified date at which the accepted, peer-reviewed manuscript is made available to the public – which may be an online-first/electronic pre-print. Pre-prints prior to manuscript acceptance were not considered.
	Dates imputed	A Boolean variable coded as: FALSE if both the search date and publication date were specified precisely, or TRUE if either date was imputed. Imputation was based on the mid-point of the specified date range, for example if a search month is specified rather than a search date, then the 15th day of the month was the

		imputed search date.
Review parameters	Number of apps reviewed	Number of apps reported for analysis, after any screening or filtering for relevance.
	Apps downloaded	Ordinal variable coded as 'no apps downloaded' (e.g. analysis was based on only app store metadata), 'some apps downloaded' (e.g. a targeted or random sample), or 'all apps downloaded'.
	Apps recommended	A Boolean variable coded as: TRUE if individual apps were named and described in a manner which suggests or recommends their use, or FALSE otherwise.

Extraction was completed in a two-phase process. In the first phase, we attempted to automatically extract study metadata, app store search and publication history dates. We used a heuristic text matching strategy to locate and excerpt relevant text from study full text, published study metadata and citation database records. Any matched text was used to pre-populate a standardised data extraction form for subsequent review.

In the second phase, each study was reviewed manually to verify automatically identified data and populate coding items not suitable for automation. Only studies which reported at least one search date (i.e. earliest search date) and a publication date were retained for analysis. Each study was reviewed by a pair of reviewers (from JH, KH, CL, ML, JN, KO, MT, IW, QW), each working independently. Any discrepancies were resolved by a third reviewer (from KH, ML and JN) not involved in the original review. Because data extraction included items with non-categorical assignments, we assessed inter-rater agreement using raw agreement (the proportion of scored data items where each reviewer pair assigned the same value). Overall agreement was 0.79 (n=1273/1618 data items) versus 0.83 (n=538/646) for those items extracted in the pilot phase. Considering extraction of study publication dates alone, being the data items intended to inform the primary analysis of study, agreement was 0.84 (n=478/570) versus 0.90 (n=207/230) extracted in the pilot phase.

Data analysis

The primary reported outcome is time to publication (TTP), calculated as the difference in days between the earliest search date and the date of publication. This window was justified on the basis that it reflects a conservative upper bound on the 'staleness' of information contained in any review at the earliest time it becomes accessible to a public audience. A secondary measure, the time to submission (TTS), was calculated as the difference between the earliest search date and the submission date. Descriptive statistics are reported for both TTP and TTS.

Data concerning review parameters (number of apps reviewed, whether apps were downloaded, and whether apps were recommended) are presented descriptively. Due to non-normalcy, non-parametric tests were used to assess the relationship between these parameters and the primary outcome, TTP. The correlation between TTP and the number of apps reviewed was measured using Spearman's

correlation coefficient (ρ). The impact of downloading some or all apps on TTP was assessed using a Kruskal-Wallis test, with follow-up tests to identify differences between specific groups. Whether studies which recommended specific apps were published more rapidly was assessed using a Wilcoxon rank sum test.

Two sensitivity analyses were specified *a priori*. We anticipated that studies would exist where date information was reported only partially, for example, reporting only the month and year in which app searches were performed. In these cases (n=93), we imputed the date as the 15th of the stated month or, if the authors reported a range of actual dates, selected a single date lying in the middle of this range. The first sensitivity analysis assessed the consequences of partial date reporting by comparing TTP for those studies where date imputation was and was not required.

A second sensitivity analysis aimed to explore the consequences of assuming that no important app changes occurred between the earliest and most recent app search/update dates. To do this, TTP was recalculated using the *last available* search date for each study (i.e. the latest reported value of 'earliest search date', 'latest search date', and 'updated search date') and compared to the original TTP measure. For both sensitivity analyses, TTP values were compared using Wilcoxon rank sum tests.

A *post hoc* exploration of the relationship between the two effort-related review parameters (number of apps reviewed and whether apps were downloaded) was conducted. The number of apps reviewed was compared across the sub-groups based on whether none, some, or all of the apps were downloaded in the review process and tested with a Kruskal-Wallis test. An additional metric, the review time per app, was defined as the time to publication divided by the number of apps reviewed. This metric was again compared across downloaded sub-groups, and compared using a Kruskal-Wallis test.

Finally, a second *post hoc* analysis was undertaken to examine whether the time to publication has changed over time. This may reflect, for example, that the methodology for app reviews has developed and normalised in recent years. To investigate this effect, a linear regression of the time to publication against the earliest search date was performed.

All analyses were conducted using MATLAB 8.6.

Patient and Public Involvement

As this was an analysis of previously published literature, patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Results

Search and selection

Searches of the published literature were performed on 30 April 2019 and yielded 439 study reports. After deduplication, screening and full-text review (summarised in Figure 1), 136 reports were included in the final analysis (see Supplementary File 2).

Time to publication and submission

The median time to publication, TTP, was 431 days (approximately 14 months, range: 42-1054 days) from the earliest search date. 100 papers reported a submission date, however in 8 cases the submission date was prior to the search date. Logically the search date should precede the submission, therefore these 8 papers were excluded from the analysis as the accuracy of the reported dates is uncertain. From the 92 remaining studies, the median time to submission, TTS, was 269 days (approximately 9 months, range: 5-874 days). The distribution of TTP and TTS are shown in Figure 2.

Review parameters

The median number of apps reviewed in the 136 included studies was 52, although there was large variation between studies (range: 4-1806 apps). A near-zero correlation was found between the number of apps reviewed and TTP (ρ=-0.004, p=0.964).

Authors typically downloaded the apps for review, rather than relying on app store descriptions: 72.1% (n=98/136) papers indicated all apps were downloaded for review vs 17.6% (n=24/136) where no apps were downloaded. A targeted or random sample of apps were downloaded in 9.6% (n=13/136) of studies, and in 1 study it was not possible to determine whether or not apps were downloaded. There was a significant difference in TTP between the subgroups (p=0.010). Figure 3 shows the distribution of TTP for each sub-group. Follow-up tests identified a longer publication time when all apps compared to no apps were downloaded (median TTP: 476 days vs 292 days).

Specific apps were named and recommended for use in 15.4% (n=21/136) of the reviewed studies. Studies which included recommendations for specific named apps were published marginally more quickly than other studies, but this difference was not significant (median: 425 vs 440 days, p=0.998, see Figure 4.)

Sensitivity analyses

Two-thirds of the studies (92/136, 67.6%) did not specify an exact app store search date, and two (2/136, 1.5%) were not associated with a precise publication date. In combination, dates were imputed in 93 of the studies (68.4%). The difference in time to publication was not significantly different between papers with precise or imputed dates (457 vs 430 days respectively, p=0.648),

Approximately a quarter of the app reviews (37/136, 27.2%) reported a latest search date, and 5.1% (7/136) reported an updated search date. Using the latest of the three reported dates, the median time to publication reduced to 387 days, which did not reach significance for difference from the primary outcome (p=0.063).

Post-hoc analysis

As there appeared to be a relationship between whether apps were downloaded and the time to publication, but no such relationship for the number of apps reviewed in the studies, we conducted a *post-hoc* analysis to examine whether there was a relationship between these two review features. Figure 5 shows the variation in the number of apps reviewed, based on whether apps were downloaded as part of the review process, and a significant difference was found (p<0.001). Follow-up tests identified significantly more apps were reviewed when no or some apps were downloaded, compared with all apps being downloaded (median number of apps: 130, 175, and 43 respectively).

When the time to publication was normalised by the number of apps included in the review, the review time per app was 8.2 days. There were significant differences between studies based on whether apps were downloaded (p<0.001) – with those studies where all apps were downloaded taking significantly longer (median 12.1 days/app) compared with the no downloaded (2.9 days/app) and some downloaded (4.2 days/app) apps.

The second *post hoc* analysis examined the change in time to publication over time, as shown in Figure 6. The time to publication has increased by 4.6 days each year, however this is not significant (p=0.69).

Discussion

This study aimed to quantify the extent to which data presented in cross-sectional app reviews is up-to-date by examining the delay between the selection of apps for review and the time of publication. By the time that most app reviews become available for use, a considerable period of time has elapsed (median: 431 days, 14.2 months). This measure still exceeds one year (387 days, 12.7 months) when a more lenient measure of the recency of findings at the time of publication is used. This delay is not wholly attributable to factors outside the control of researchers, such as the peer review process: when the estimated time to submission was calculated to provide a crude metric of the time to conduct the study, excluding journal peer-review and editing processes, there was still substantial time between app search and manuscript submission (269 days, 8.8 months). The time taken to publish app reviews was influenced by the nature of the analysis, with findings indicating that reviews that downloaded apps for analysis took significantly longer to publish than those that did not. Surprisingly, the number of apps reviewed did not influence publication time, however a post-hoc analysis indicated papers that downloaded apps reviewed significantly fewer apps than those that did not.

Given previous research indicating a high rate of turnover in the app marketplace, with 50% of mental health search results changing within approximately three months,[27] the observed delay in publication

raises questions about the validity of study findings at the time they become available to the research, clinical, and broader community, particularly where such reviews focus on recommendations concerning specific products. Reviews may recommend the use of apps which are no longer supported by the developers, have been withdrawn from the app stores or, conversely, have been updated substantially since the review. Contrary to our hypothesis, we did not find that studies making specific recommendations had a shorter time to publication than other reviews. Recommendations for specific apps in published reviews cannot, therefore, automatically be considered reliable. The delay in publication may also mean that more recent, potentially high quality, apps are not made known to the research or clinical communities.

Our finding of substantial delays between initial assessment and publication provides a counterpoint to recent commentaries from academic clinicians discussing how to introduce apps into clinical practice that have emphasised a role for this kind of research to guide clinicians by identifying unsafe and poor quality apps[28] while simultaneously identifying deficiencies in alternatives to framework-based, structured app reviews such as certification programs[28] and user reviews[29]. Our data show that cross-sectional app reviews are also subject to important limitations concerning how up-to-date the information they contain is (and perhaps can be, given the academic publication process.) Because academic reviews are not designed to be continuously updated, healthcare professionals cannot assume that conclusions concerning the quality and safety of specific apps are still valid.

There are a number of potential strategies that could mitigate this issue. The first is for review authors to refresh their results prior to publication. Pre-publication update is a standard practice in systematic reviews. For example, the Cochrane Collaboration will not publish reviews unless the most recent search date is less than twelve months (and ideally less than six months) old.[30] As part of efforts to improve the quality of cross-sectional app reviews,[31] editors and peer reviewers should consider at least asking for justification where there is a long period between search and submission. The practicality of this solution must nevertheless consider review-specific factors that may affect the time to publication, the feasibility of update, and whether the review *intends* to guide clinical and public uses.

The second potential strategy is for authors to adjust the numbers of apps incorporated in their reviews. The relationship between downloading of apps and the number of apps reviewed may indicate that study authors hold some shared perceptions about what represents a 'publishable unit' of work. This may be achieved by either downloading a smaller number of apps, or by reviewing the app store descriptions for a larger number of apps. Both appear to result in publication in approximately one year (± three months). While both have merit, it seems likely that studies that scrutinise app content directly are likely to yield richer insights than those relying on summary information presented in app stores for the purposes of marketing. However, with the longer time to publication associated with downloading, it may be appropriate to focus on a smaller sample of the most popular, most used, or top-ranked apps, which can be published more quickly. Authors should also consider how cross-platform apps should be handled. Apps which are available for both Android and iOS may share common features and functionality, however some aspects may be unique to one platform. There may therefore be a trade-off between comprehensively reviewing all versions and streamlining the review of a single version.

A third possible strategy is to remove app assessment and review from the academic sphere, to organisations whose resources are not subject to the constraints of the academic publication process and are, at least in principle, resourced to be able to respond to app dynamics such as update and withdrawal. Indeed, continuous app scanning and review approaches have now been adopted by a number of health organisations, including the NHS Apps Library[32] and the American Psychiatric Association.[33] However, despite the intention to provide continuous review, it is unclear how often reviews and recommendations that appear within these portals are actually updated. Further, even within app portals, the large number of available apps often necessitates that thoroughness be balanced with expediency,[34] potentially still limiting broad utility of such resources.

In parallel with the development of the academic literature regarding the quality of health apps, different jurisdictions have developed regulatory frameworks to govern the distribution of apps, particularly those which may be considered to be medical devices. While these frameworks differ across jurisdictions, harmonisation of quality criteria may help further refine and improve the wide range of quality assessment methodologies employed across the literature.[35]

Importantly, we do not suggest that our findings imply that cross-sectional assessments of health apps have no utility. App reviews may not be the optimal source of timely information about the function and quality of specific apps, but research-based methods are well suited to identifying and providing unbiased estimates of the nature and extent of thematic issues affecting specific populations of apps. Such insights can and have guided systemic responses to app quality problems. Research-led studies have arguably been important both in identifying *new* issues affecting apps, particularly where identification of issues involves complex exploratory and technical analysis,[36] and in devising systematic strategies for their identification, such as MARS. Unless specifically resourced to do so, it seems unlikely that continuous scanning programmes and app portals will be able to fulfill this discovery function.

Limitations

While this study examined app reviews across all health domains, to characterise the publication delay generally for the mobile health field of research, it did not examine differences across specific sub-domains. It is possible that different outcomes would be observed for different health conditions, for example due to resourcing availability/constraints across clinical domains, or due to differences in self-management approaches for different conditions. These differences were not considered in the current study due to substantial observed heterogeneity in the scope of reviews: some considered only technical domains of app quality (for example, data privacy), some considered broad categories (for example, mental health or physical health), and some considered specific conditions. Furthermore, the databases selected for the literature searches may not have provided complete coverage across all reviews focused on health and wellbeing domains (for example, those reported in allied health publications) or technical domains (for example, those focusing on data privacy and security). However, the databases selected are likely to capture the papers most likely to have an impact on clinical practice.

The time to submission was calculated in addition to the primary outcome, time to publication, to provide an approximate measure of the time to conduct the review process. However, this is only a crude estimate as it cannot account for manuscripts submitted to multiple journals prior to acceptance.

Some journals offer a fee-for-service option to expedite the peer review process, which would be expected to result in a quicker time to publication. It is possible that authors using this facility may also conduct the reviews in a shorter period of time, resulting in quicker time to submission. We observed no markers of whether articles had been expedited, so it was not possible to assess the impact this publication model.

Finally, given the research question addressed by this study, we acknowledge the time taken to conduct and publish this review. However, the body of peer-reviewed, published academic literature is more stable and develops at a slower pace than the highly dynamic app stores. Furthermore, our findings show that the time to publication has been stable for the past decade, therefore it is unlikely that the findings reported here have lost relevancy since the literature search was conducted.

Conclusions

The majority of health app reviews present data that is at least a year out of data at the time of publication. Given the high rate of observed turnover of health apps in public marketplaces, it may not be appropriate, therefore, for these reviews to be presented as a resource concerning specific products for commissioners, clinicians and the public. Authors of such reviews should, where possible, take steps to minimise the delay to publication, update their results prior to publication and consider whether making specific product recommendations is appropriate. App reviews may nevertheless fulfill important functions to identify novel and thematic issues and guide policy and systemic responses to health app quality and safety. App users should consider alternative sources of information about apps that are better calibrated to the dynamics of the app marketplace, such as continuous scanning services offered by dedicated health app portals.

Statement of contributorship

ML and KH conceived the study and are guarantors. JN and KH devised the search strategies and performed the searches. ML, JN and KH performed initial screening and ML collated the results. KH developed and executed the automatic data extraction process. KH prepared the manual data extraction tools, allocated reviewer tasks and collated results. ML, JN, JH, CL, KO, MT, DW, IW, QW and KH performed data extraction. ML performed principal data analysis. ML, JN and KH wrote the first draft of the manuscript. ML, JN, JH, CL, KO, MT, DW, IW, QW and KH provided comments to and/or contributed revisions to the draft manuscript. ML, JN, JH, CL, KO, MT, DW, IW, QW and KH reviewed and approved the final manuscript version.

Competing interests

KH, ML, JN are authors of studies that were included in this review. They declare no other competing financial or non-financial interests. All other authors declare no competing financial or non-financial interests.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability statement

Data are available upon reasonable request.

References

1. Kitsiou S, Pare G, Jaana M, Gerber B. Effectiveness of mhealth interventions for patients with diabetes: An overview of systematic reviews. *PLoS One* 2017; **12**(3): e0173160.

2. Coorey GM, Neubeck L, Mulley J, Redfern J. Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: Systematic review with meta-synthesis of quantitative and qualitative data. *Eur J Prev Cardiol* 2018; **25**(5): 505-21.

3. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: A systematic review. *J Med Internet Res* 2016; **18**(5): e97.

4. Han M, Lee E. Effectiveness of mobile health application use to improve health behavior changes: A systematic review of randomized controlled trials. *Healthc Inform Res* 2018; **24**(3): 207-26.

5. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. *J Med Internet Res* 2016; **18**(11): e287.

6. Byambasuren O, Beller E, Glasziou P. Current knowledge and adoption of mobile health apps among australian general practitioners: Survey study. *JMIR Mhealth Uhealth* 2019; **7**(6): e13199.

7. Boyle L, Grainger R, Hall RM, Krebs JD. Use of and beliefs about mobile phone apps for diabetes self-management: Surveys of people in a hospital diabetes clinic and diabetes health professionals in new zealand. *JMIR Mhealth Uhealth* 2017; **5**(6): e85.

8. Rubanovich CK, Mohr DC, Schueller SM. Health app use among individuals with symptoms of depression and anxiety: A survey study with thematic coding. *JMIR Ment Health* 2017; **4**(2): e22.

9. Australian Government Department of Health and Ageing. E-Mental health strategy for Australia. Canberra, Australia: Commonwealth of Australia, 2012.

10. California Mental Health Services Authority. "The Technology Suite" Driving access to behavioral health care thru innovation. 2018.

11. NHS England. The NHS Long Term Plan. London, United Kingdom: HM Government, 2019.

12. Huckvale K, Adomaviciute S, Prieto JT, Leow MK, Car J. Smartphone apps for calculating insulin dose: a systematic assessment. *BMC medicine* 2015; **13**: 106.

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

14. Larsen M, Nicholas J, Christensen H. A Systematic Assessment of Smartphone Tools for Suicide Prevention. *PLoS One* 2016; **11**(4): e0152285.

15. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile Apps for Bipolar Disorder: A Systematic Review of Features and Content Quality. *Journal of medical Internet research* 2015; **17**(8): e198.

16. Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res* 2013; **15**(11): e247.

17. Thornton L, Quinn C, Birrell L, et al. Free smoking cessation mobile apps available in australia: A quality review and content analysis. *Australian and New Zealand Journal of Public Health* 2017; **41**(6): 625-30.

18. Goodwin GM, Haddad PM, Ferrier IN, et al. Evidence-based guidelines for treating bipolar disorder: Revised third edition recommendations from the British Association for Psychopharmacology. *J Psychopharmacol* 2016; **30**(6): 495-553.

19. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist. *Clinical Medicine Journal (London)* 2015; **15**(6): 519-21.

20. Torous JB, Chan SR, Gipson SY-MT, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatric Services* 2018; **69**(5): 498-500.

21. Stephens H, Uccellini M, McKay F. Guidelines for creating healthy living apps. Melbourne, Australia: Dialog Consultking, 2015.

22. Centers for Disease Control and Prevention. Best practices user guide: Health communications in tobacco prevention and control. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2018.

23. Rose KJ, Petrut C, L'Heveder R, de Sabata S. IDF Europe's position on mobile applications in diabetes. *Diabetes Res Clin Pract* 2019; **149**: 39-46.

24. Ronchi E, Woskie LR, Milstein JA. Mobile technology-based services for global health and wellness: opportunities and challenges. Boston, USA: Organisation for Economic Cooperation and Development (OECD), 2017.

25. World Bank Group. Empowering households and individuals to co-produce positive health outcomes for dignified, person-centered care amidst demographic change Washington, DC, USA: World Bank, 2018.

26. 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 and uHealth* 2015; **3**(1): e27-e.

27. Larsen ME, Nicholas J, Christensen H. Quantifying App Store Dynamics: Longitudinal Tracking of Mental Health Apps. *JMIR Mhealth Uhealth* 2016; **4**(3): e96.

28. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. *npj Digital Medicine* 2020; **3**(1): 14.

29. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. *The Lancet Digital Health* 2019; **1**(2): e52-e4.

30. Tariq S, Akhtar N, Afzal H, et al. A novel co-training-based approach for the classification of mental illnesses using social media posts. *IEEE Access* 2019; **7**: 166165-72.

31. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association* 2015; **21**(2): 97-104.

32. NHS. NHS Apps Library. https://www.nhs.uk/apps-library/ (accessed 25 February 2020).

33. American Psychiatric Association. App evaluation model. https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model (accessed 25 February 2020).

34. Meek T. Tom's digital disruptors: testing times for apps. Digital Health; 2015.

35. 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. *International Journal of Technology Assessment in Health Care* 2018; **34**(5): 464-75.

36. Huckvale K, Torous J, Larsen ME. Assessment of the Data Sharing and Privacy Practices of Smartphone Apps for Depression and Smoking Cessation. *JAMA Netw Open* 2019; **2**(4): e192542.

Figure Captions

Figure 1. Study selection flowchart.

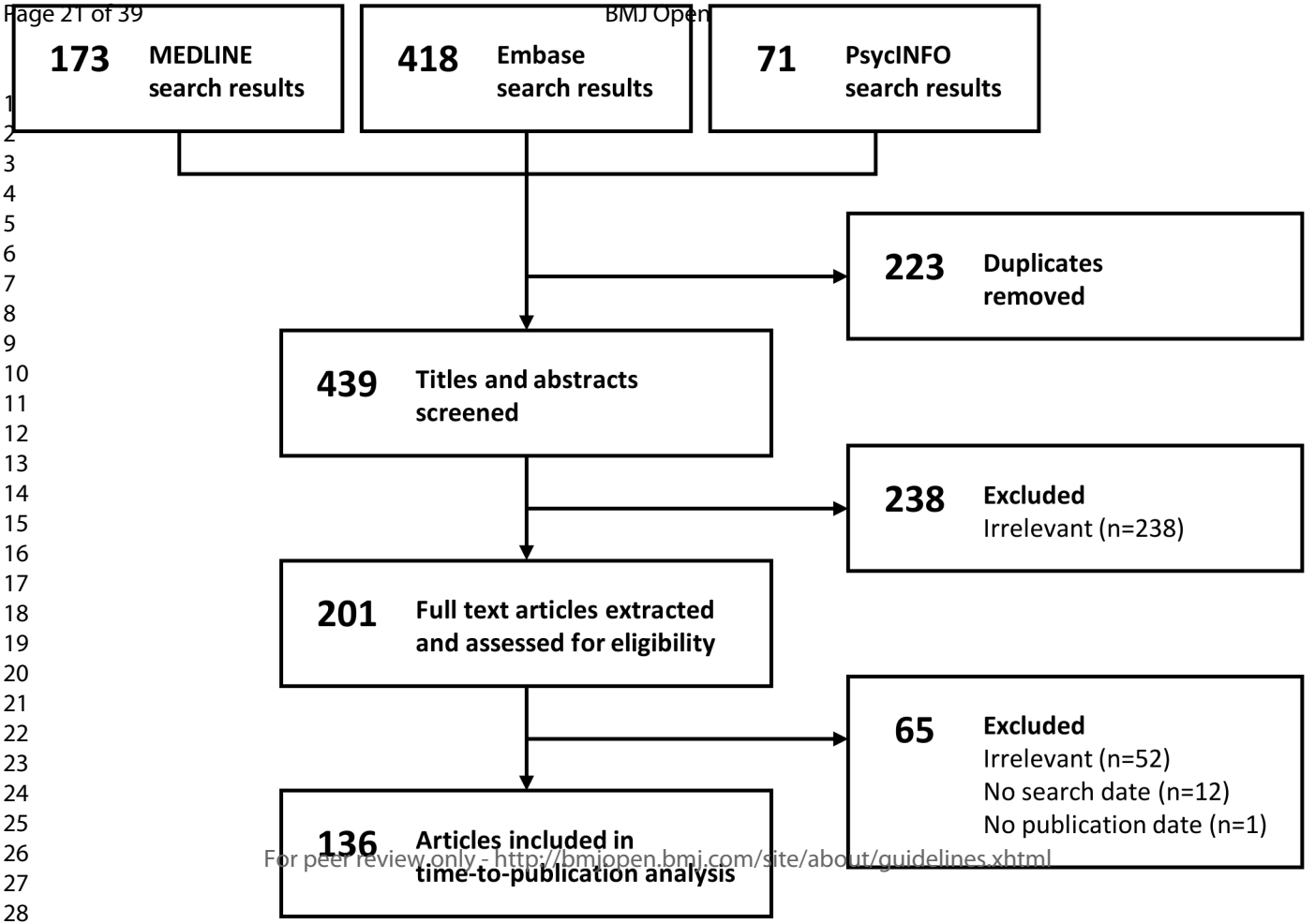
Figure 2. Frequency distributions for (a) the time to publication (TTP), and (b) time to submission (TTS), with (c) a side-by-side comparison.

Figure 3. Distribution in the time to publication based on whether apps were downloaded as part of the review process.

Figure 4. Distribution in the time to publication based on whether specific apps were named and recommended.

Figure 5. Variation in the number of apps reviewed, depending on whether the apps were downloaded as part of the review process. The y-axis has been truncated, and not all outlier values are shown.

Figure 6. Variation in the time to publication based on the app store search date.



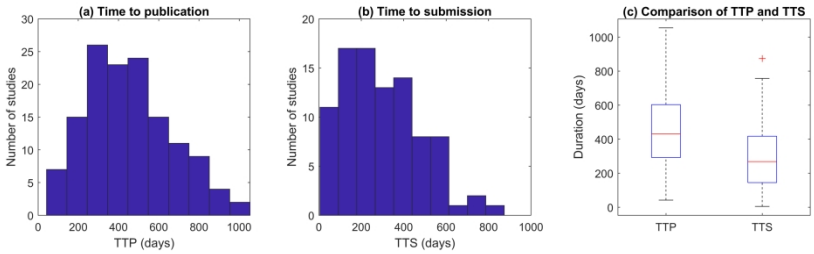


Figure 2. Frequency distributions for (a) the time to publication (TTP), and (b) time to submission (TTS), with (c) a side-by-side comparison.

317x79mm (300 x 300 DPI)

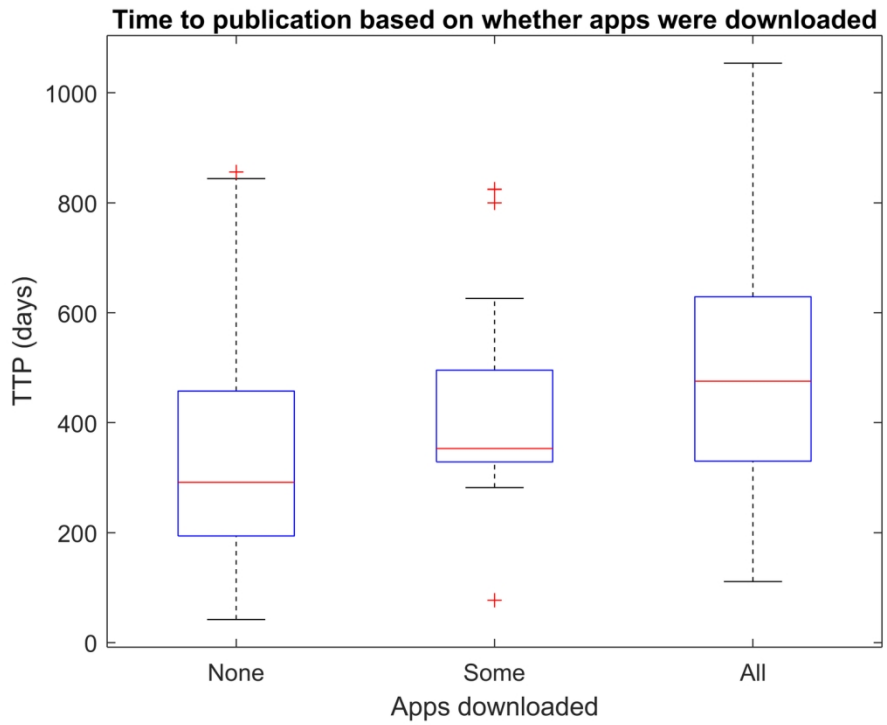


Figure 3. Distribution in the time to publication based on whether apps were downloaded as part of the review process.

148x111mm (300 x 300 DPI)

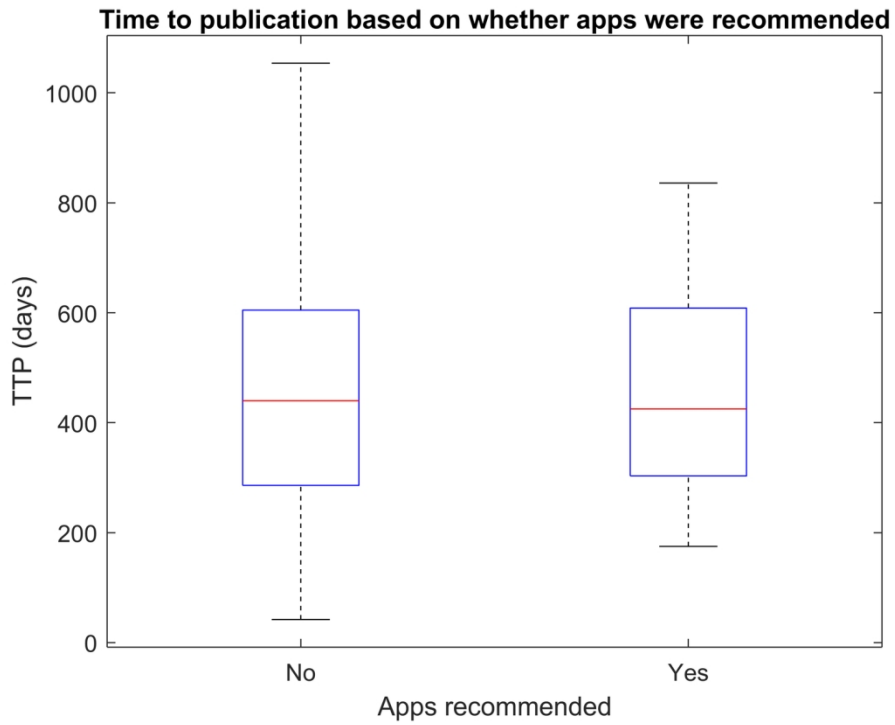


Figure 4. Distribution in the time to publication based on whether specific apps were named and recommended.

148x111mm (300 x 300 DPI)

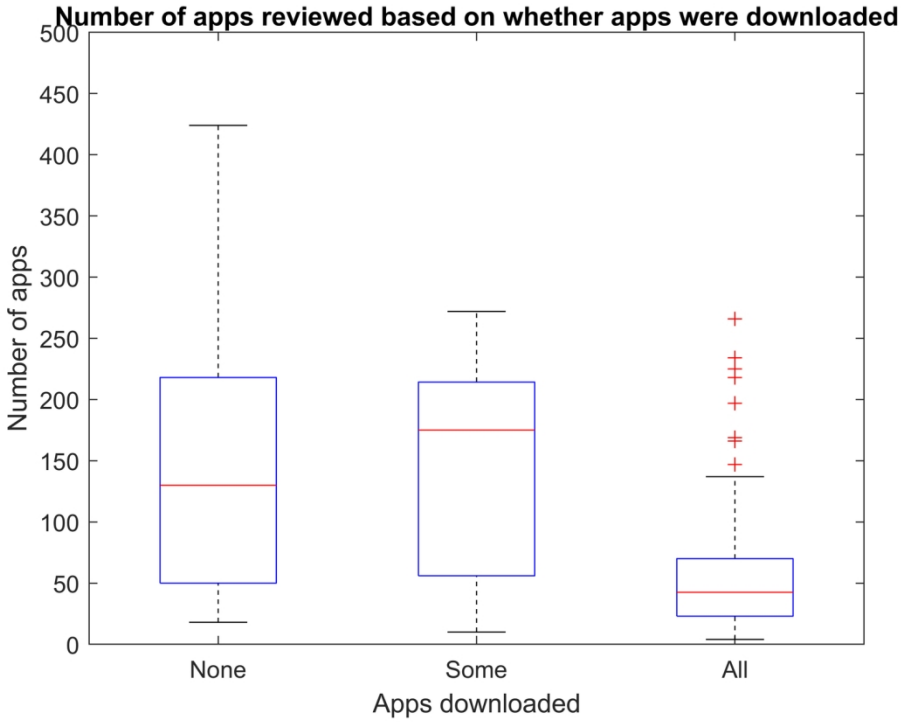


Figure 5. Variation in the number of apps reviewed, depending on whether the apps were downloaded as part of the review process. The y-axis has been truncated, and not all outlier values are shown.

148x111mm (300 x 300 DPI)

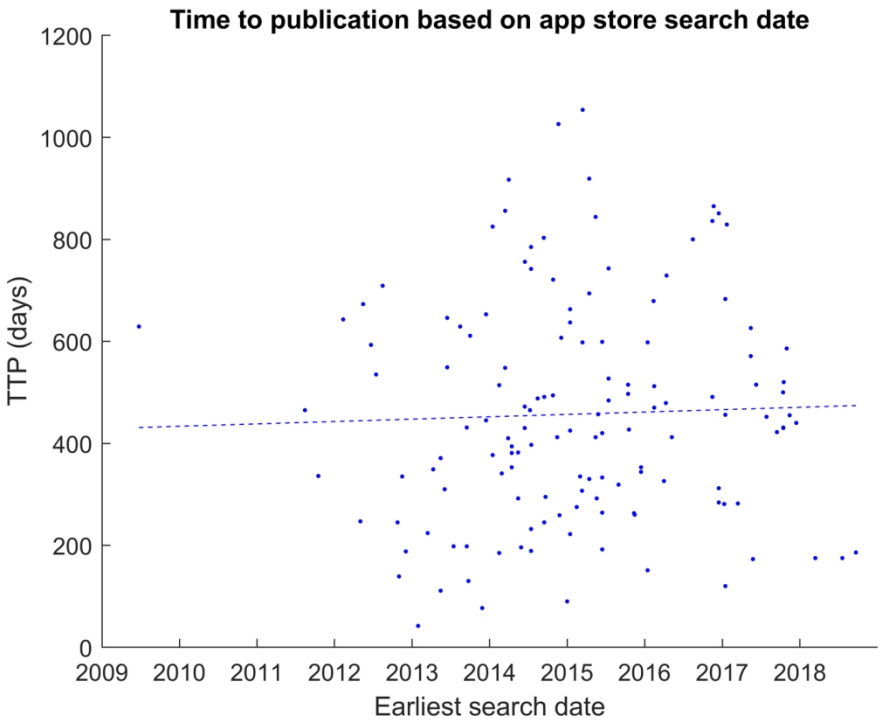


Figure 6. Variation in the time to publication based on the app store search date.

148x111mm (300 x 300 DPI)

S1 Alternative pilot search strategy

The final search strategy is reported in Box 1 of the manuscript.

1. Mobile Applications/
2. Cell Phone/
3. 1 or 2
4. exp "Review"/ or exp "Systematic Review"/ or review.mp.
5. assessment.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
6. 4 or 5
7. 3 and 6
8. Humans/
9. 7 and 8
10. limit 9 to yr="2007 -Current"
11. limit 10 to yr="2018"

S2 List of included studies

Authors	Title	Journal
O'Loughlin K., Neary M., Adkins E.C., Schueller S.M.	Reviewing the data security and privacy policies of mobile apps for depression.	Internet Interventions
Quevedo Rodriguez A., Wagner A.M.	Mobile phone applications for diabetes management: A systematic review.	Endocrinologia, Diabetes y Nutricion
Mousavi Jazayeri S.M.H., Jamshidnezhad A.	Top mobile applications in pediatrics and children's health: Assessment and intelligent analysis tools for a systematic investigation.	Malaysian Journal of Medical Sciences
Middelweerd, Anouk, Mollee, Julia S, van der Wal, C. Natalie, Brug, Johannes, te Velde, Saskia J	Apps to promote physical activity among adults: A review and content analysis.	The International Journal of Behavioral Nutrition and Physical Activity
Singer, Mathias Van, Chatton, Anne, Khazaal, Yasser	Quality of smartphone apps related to panic disorder.	Frontiers in Psychiatry
Bennett, Melanie E, Toffey, Kristin, Dickerson, Faith, Himelhoch, Seth, Katsafanas, Emily, Savage, Christina L. G	A review of android apps for smoking cessation.	Journal of Smoking Cessation
Bardus, Marco, van Beurden, Samantha B, Smith, Jane R, Abraham, Charles	A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management.	The International Journal of Behavioral Nutrition and Physical Activity
Jeon E, Park HA, Min YH, Kim HY	Analysis of the information quality of korean obesity-management smartphone applications.	Healthcare Informatics Research
Mendiola MF, Kalnicki M, Lindenauer S	Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings.	JMIR MHealth and UHealth
Payne HE, Moxley VB, MacDonald E	Health Behavior Theory in Physical Activity Game Apps: A Content Analysis.	JMIR Serious Games
Ramo DE, Popova L, Grana R, Zhao S, Chavez K	Cannabis Mobile Apps: A Content Analysis.	JMIR MHealth and UHealth
Mani M, Kavanagh DJ, Hides L, Stoyanov SR	Review and Evaluation of Mindfulness-Based iPhone Apps.	JMIR MHealth and UHealth

Authors	Title	Journal
Hale K, Capra S, Bauer J	A Framework to Assist Health Professionals in Recommending High-Quality Apps for Supporting Chronic Disease Self-Management: Illustrative Assessment of Type 2 Diabetes Apps.	JMIR MHealth and UHealth
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.	Interactive Journal of Medical Research
Chen J, Cade JE, Allman-Farinelli M	The Most Popular Smartphone Apps for Weight Loss: A Quality Assessment.	JMIR MHealth and UHealth
Mangone ER, Lebrun V, Muessig KE	Mobile Phone Apps for the Prevention of Unintended Pregnancy: A Systematic Review and Content Analysis.	JMIR MHealth and UHealth
Con D, De Cruz P	Mobile Phone Apps for Inflammatory Bowel Disease Self-Management: A Systematic Assessment of Content and Tools.	JMIR MHealth and UHealth
Iribarren SJ, Schnall R, Stone PW, Carballo-Dieguez A	Smartphone Applications to Support Tuberculosis Prevention and Treatment: Review and Evaluation.	JMIR MHealth and UHealth
Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S	Review and Analysis of Existing Mobile Phone Apps to Support Heart Failure Symptom Monitoring and Self-Care Management Using the Mobile Application Rating Scale (MARS).	JMIR MHealth and UHealth
Zaidan S, Roehrer E	Popular Mobile Phone Apps for Diet and Weight Loss: A Content Analysis.	JMIR MHealth and UHealth
Pereira-Azevedo N, Osorio L, Cavadas V, Fraga A, Carrasquinho E, Cardoso de Oliveira E, Castelo-Branco M, Roobol MJ	Expert Involvement Predicts mHealth App Downloads: Multivariate Regression Analysis of Urology Apps.	JMIR MHealth and UHealth
Rivera J, McPherson A, Hamilton J, Birken C, Coons M, Iyer S, Agarwal A, Lalloo C, Stinson J	Mobile Apps for Weight Management: A Scoping Review.	JMIR MHealth and UHealth
Franco RZ, Fallaize R, Lovegrove JA, Hwang F	Popular Nutrition-Related Mobile Apps: A Feature Assessment.	JMIR MHealth and UHealth
Wilson H, Stoyanov SR, Gandabhai S, Baldwin A	The Quality and Accuracy of Mobile Apps to Prevent Driving After Drinking Alcohol.	JMIR MHealth and UHealth
Geuens J, Swinnen TW, Westhovens R, de Vlam K, Geurts L, Vanden Abeele V	A Review of Persuasive Principles in Mobile Apps for Chronic Arthritis Patients: Opportunities for Improvement.	JMIR MHealth and UHealth
Huang ET, Williams H, Hocking JS, Lim MS	Safe Sex Messages Within Dating and Entertainment Smartphone Apps: A Review.	JMIR MHealth and UHealth

Authors	Title	Journal
Chen E, Mangone ER	A Systematic Review of Apps using Mobile Criteria for Adolescent Pregnancy Prevention (mCAPP).	JMIR MHealth and UHealth
Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J	Mobile Phone Apps to Improve Medication Adherence: A Systematic Stepwise Process to Identify High-Quality Apps.	JMIR MHealth and UHealth
Sullivan RK, Marsh S, Halvarsson J, Holdsworth M, Waterlander W, Poelman MP, Salmond JA, Christian H, Koh LS, Cade JE, Spence JC, Woodward A, Maddison R	Smartphone Apps for Measuring Human Health and Climate Change Co-Benefits: A Comparison and Quality Rating of Available Apps.	JMIR MHealth and UHealth
Grainger R, Townsley H, White B, Langlotz T, Taylor WJ	Apps for People With Rheumatoid Arthritis to Monitor Their Disease Activity: A Review of Apps for Best Practice and Quality.	JMIR MHealth and UHealth
Christmann CA, Hoffmann A, Bleser G	Stress Management Apps With Regard to Emotion-Focused Coping and Behavior Change Techniques: A Content Analysis.	JMIR MHealth and UHealth
Ginossar T, Shah SF, West AJ, Bentley JM, Caburnay CA, Kreuter MW, Kinney AY	Content, Usability, and Utilization of Plain Language in Breast Cancer Mobile Phone Apps: A Systematic Analysis.	JMIR MHealth and UHealth
Formagini TD, Ervilha RR, Machado NM, Andrade BA, Gomide HP, Ronzani TM	A review of smartphone apps for smoking cessation available in Portuguese.	Cadernos de Saude Publica
Bright T, Pallawela D	Validated Smartphone-Based Apps for Ear and Hearing Assessments: A Review.	JMIR Rehabilitation And Assistive Technologies
Cheng F, Xu J, Su C, Fu X, Bricker J	Content Analysis of Smartphone Apps for Smoking Cessation in China: Empirical Study.	JMIR MHealth and UHealth
Tinschert P, Jakob R, Barata F, Kramer JN, Kowatsch T	The Potential of Mobile Apps for Improving Asthma Self-Management: A Review of Publicly Available and Well-Adopted Asthma Apps.	JMIR MHealth and UHealth
Matava C, Leo AM, Alam F	Mobile Apps for Teaching Intubation: Scoping Review and Critical Analysis in eLearning.	JMIR medical education
Muthing J, Jaschke T, Friedrich CM	Client-Focused Security Assessment of mHealth Apps and Recommended Practices to Prevent or Mitigate Transport Security Issues.	JMIR MHealth and UHealth
Lalloo C, Shah U, Birnie KA, Davies-Chalmers C, Rivera J, Stinson J, Campbell F	Commercially Available Smartphone Apps to Support Postoperative Pain Self-Management: Scoping Review.	JMIR MHealth and UHealth
Sudol NT, Adams-Piper E, Perry R, Lane F, Chen KT	In Search of Mobile Applications for Patients With Pelvic Floor Disorders.	Female Pelvic Medicine & Reconstructive Surgery

Authors	Title	Journal
Frie K, Hartmann-Boyce J, Jebb S, Albury C, Nourse R, Aveyard P	Insights From Google Play Store User Reviews for the Development of Weight Loss Apps: Mixed-Method Analysis.	JMIR MHealth and UHealth
Wallace SL, Mehta S, Farag S, Kelley RS, Chen KT	In Search of Mobile Applications for Urogynecology Providers.	Female Pelvic Medicine & Reconstructive Surgery
Metelmann B, Metelmann C, Schuffert L, Hahnenkamp K, Brinkrolf P	Medical Correctness and User Friendliness of Available Apps for Cardiopulmonary Resuscitation: Systematic Search Combined With Guideline Adherence and Usability Evaluation.	JMIR MHealth and UHealth
Park JYE, Li J, Howren A, Tsao NW, De Vera M	Mobile Phone Apps Targeting Medication Adherence: Quality Assessment and Content Analysis of User Reviews.	JMIR MHealth and UHealth
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 and UHealth
Yang G, Long J, Luo D, Xiao S, Kaminga AC	The Characteristics and Quality of Mobile Phone Apps Targeted at Men Who Have Sex With Men in China: A Window of Opportunity for Health Information Dissemination?.	JMIR MHealth and UHealth
Abroms L.C., Padmanabhan N., Thaweethai L., Phillips T.	iPhone apps for smoking cessation: A content analysis.	American Journal of Preventive Medicine
Cowan L.T., van Wagenen S.A., Brown B.A., Hedin R.J., Seino-Stephan Y., Hall P.C., West J.H.	Apps of Steel: Are Exercise Apps Providing Consumers With Realistic Expectations?: A Content Analysis of Exercise Apps for Presence of Behavior Change Theory.	Health Education and Behavior
Muessig K.E., Pike E.C., Legrand S., Hightow-Weidman L.B.	Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review.	Journal of medical Internet research
Azar K.M.J., Lesser L.I., Laing B.Y., Stephens J., Aurora M.S., Burke L.E., Palaniappan L.P.	Mobile applications for weight management: Theory-based content analysis.	American Journal of Preventive Medicine
Abroms L.C., Lee Westmaas J., Bontemps-Jones J., Ramani R., Mellerson J.	A content analysis of popular smartphone apps for smoking cessation.	American Journal of Preventive Medicine
Cantudo Cuenca M.R., Cantudo Cuenca M.D., Morillo Verdugo R.	Availability and medical professional involvement in mobile healthcare applications related to pathophysiology and pharmacotherapy of HIV/AIDS.	European Journal of Hospital Pharmacy
Haffey F., Brady R.R.W., Maxwell S.	Smartphone apps to support hospital prescribing and pharmacology education: A review of current provision.	British Journal of Clinical Pharmacology
Stevens D.J., Jackson J.A., Howes N., Morgan J.	Obesity surgery smartphone apps: A review.	Obesity Surgery

Authors	Title	Journal
Wallace L.S., Dhingra L.K.	A systematic review of smartphone applications for chronic pain available for download in the United States.	Journal of Opioid Management
Cantudo-Cuenca Ma.R., Robustillo-Cortes Ma.A., Cantudo-Cuenca Ma.D., Morillo-Verdugo R.	A better regulation is required in viral hepatitis smartphone applications.	Farmacia Hospitalaria
Choi J., Noh G.Y., Park D.J.	Smoking cessation apps for smartphones: content analysis with the self-determination theory.	Journal of medical Internet research
Arnhold M., Quade M., Kirch W.	Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older.	Journal of medical Internet research
Huckvale K., Car M., Morrison C., Car J.	Apps for asthma self-management: A systematic assessment of content and tools.	BMC Medicine
Connor K., Brady R.R.W., De Beaux A., Tulloh B.	Contemporary hernia smartphone applications (apps).	Hernia
Bailey S.C., Belter L.T., Pandit A.U., Carpenter D.M., Carlos E., Wolf M.S.	The availability, functionality, and quality of mobile applications supporting medication selfmanagement.	Journal of the American Medical Informatics Association
Mobasheri M.H., Johnston M., King D., Leff D., Thiruchelvam P., Darzi A.	Smartphone breast applications - What's the evidence?.	Breast
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.	Clinical Journal of Pain
Stevens D.J., McKenzie K., Cui H.W., Noble J.G., Turney B.W.	Smartphone apps for urolithiasis.	Urolithiasis
Wurzer P., Parvizi D., Lumenta D.B., Giretzlehner M., Branski L.K., Finnerty C.C., Herndon D.N., Tuca A., Rappl T., Smolle C., Kamolz L.P.	Smartphone applications in burns.	Burns
van Mechelen D.M., van Mechelen W., Verhagen E.A.	Sports injury prevention in your pocket?! Prevention apps assessed against the available scientific evidence: a review.	British journal of sports medicine
Ouhbi S., Fernandez-Aleman J.L., Toval A., Idri A., Pozo J.R.	Free Blood Donation Mobile Applications.	Journal of Medical Systems

Authors	Title	Journal
Kumar N., Khunger M., Gupta A., Garg N.	A content analysis of smartphone-based applications for hypertension management.	Journal of the American Society of Hypertension
Schnall R., Iribarren S.J.	Review and analysis of existing mobile phone applications for health care-associated infection prevention.	American Journal of Infection Control
Shelmerdine S.C., Lynch J.O.	Smartphone applications in paediatric radiology: availability and authority.	Pediatric Radiology
Huckvale K., Morrison C., Ouyang J., Ghaghda A., Car J.	The evolution of mobile apps for asthma: An updated systematic assessment of content and tools.	BMC Medicine
Wong M.C., Fung K.	Mobile applications in otolaryngology-head and neck surgery.	Otolaryngology - Head and Neck Surgery (United States)
Lupton D., Jutel A.	'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps.	Social Science and Medicine
Jutel A., Lupton D.	Digitizing diagnosis: A review of mobile applications in the diagnostic process.	Diagnosis
Al Badi M.H., Chitme H.R.	Smartphone applications for improved pharmaceutical care.	International Journal of Pharmaceutical Sciences and Research
Kassianos A.P., Emery J.D., Murchie P., Walter F.M.	Smartphone applications for melanoma detection by community, patient and generalist clinician users: A review.	British Journal of Dermatology
Kalz M., Lenssen N., Felzen M., Rossaint R., Tabuenca B., Specht M., Skorning M.	Smartphone apps for cardiopulmonary resuscitation training and real incident support: a mixed-methods evaluation study.	Journal of medical Internet research
Huckvale K., Adomaviciute S., Prieto J.T., Leow M.K.-S., Car J.	Smartphone apps for calculating insulin dose: A systematic assessment.	BMC Medicine
Pereira-Azevedo N., Carrasquinho E., Cardoso De Oliveira E., Cavadas V., Osorio L., Fraga A., Castelo-Branco M., Roobol M.J.	mHealth in Urology: A review of experts' involvement in app development.	PLoS ONE
Cheng N.M., Chakrabarti R., Kam J.K.	iPhone applications for eye care professionals: a review of current capabilities and concerns.	Telemedicine journal and e-health : the official journal of the American Telemedicine Association

Authors	Title	Journal
Fijacko N., Brzan P.P., Stiglic G.	Mobile Applications for Type 2 Diabetes Risk Estimation: a Systematic Review.	Journal of Medical Systems
Zhang M.W.B., Ho R.C.M., Hawa R., Sockalingam S.	Analysis of the Information Quality of Bariatric Surgery Smartphone Applications Using the Silberg Scale.	Obesity Surgery
Morrissey E.C., Corbett T.K., Walsh J.C., Molloy G.J.	Behavior change techniques in apps for medication adherence: A content analysis.	American Journal of Preventive Medicine
Thomas G.M., Lupton D.	Threats and thrills: pregnancy apps, risk and consumption.	Health, Risk and Society
Kim M.S., Aro M.R., Lage K.J., Ingalls K.L., Sindhwani V., Markey M.K.	Exploring the Usability of Mobile Apps Supporting Radiologists' Training in Diagnostic Decision Making.	Journal of the American College of Radiology
Lee H., Sullivan S.J., Schneiders A.G., Ahmed O.H., Balasundaram A.P., Williams D., Meeuwisse W.H., McCrory P.	Smartphone and tablet apps for concussion road warriors (team clinicians): a systematic review for practical users.	British journal of sports medicine
Perry R., Lunde B., Chen K.T.	An evaluation of contraception mobile applications for providers of family planning services.	Contraception
Coulon S.M., Monroe C.M., West D.S.	A systematic, multi-domain review of mobile smartphone apps for evidence-based stress management.	American Journal of Preventive Medicine
Darby A., Strum M.W., Holmes E., Gatwood J.	A Review of Nutritional Tracking Mobile Applications for Diabetes Patient Use.	Diabetes Technology and Therapeutics
Crane D., Garnett C., Brown J., West R., Michie S.	Behavior change techniques in popular alcohol reduction apps: content analysis.	Journal of medical Internet research
Hoeppner B.B., Hoeppner S.S., Seaboyer L., Schick M.R., Wu G.W.Y., Bergman B.G., Kelly J.F.	How Smart are Smartphone Apps for Smoking Cessation? A Content Analysis.	Nicotine and Tobacco Research
Ong A.A., Gillespie M.B.	Overview of smartphone applications for sleep analysis.	World Journal of Otorhinolaryngology - Head and Neck Surgery
Robustillo Cortes Md.eL., Cantudo Cuenca M.R., Morillo Verdugo R., Calvo Cidoncha E.	High quantity but limited quality in healthcare applications intended for HIV-infected patients.	Telemedicine journal and e-health : the official journal of the American Telemedicine Association

Authors	Title	Journal
Nguyen A.D., Baysari M.T., Kannangara D.R.W., Tariq A., Lau A.Y.S., Westbrook J.I., Day R.O.	Mobile applications to enhance self-management of gout.	International Journal of Medical Informatics
Portelli P., Eldred C.	A quality review of smartphone applications for the management of pain.	British Journal of Pain
Brzan P.P., Rotman E., Pajnkihar M., Klanjsek P.	Mobile Applications for Control and Self Management of Diabetes: A Systematic Review.	Journal of Medical Systems
Penzenstadler L., Chatton A., Van Singer M., Khazaal Y.	Quality of Smartphone Apps Related to Alcohol Use Disorder.	European Addiction Research
Gal N., Zite N.B., Wallace L.S.	Evaluation of smartphone oral contraceptive reminder applications.	Research in social & administrative pharmacy : RSAP
Gibbs J., Gkatzidou V., Tickle L., Manning S.R., Tilakkumar T., Hone K., Ashcroft R.E., Sonnenberg P., Sadiq S.T., Estcourt C.S.	'Can you recommend any good STI apps?' A review of content, accuracy and comprehensiveness of current mobile medical applications for STIs and related genital infections.	Sexually Transmitted Infections
Lunde B., Perry R., Sridhar A., Chen K.T.	An Evaluation of Contraception Education and Health Promotion Applications for Patients.	Women's Health Issues
Calero J.J., Oton L.F., Oton C.A.	Apps for radiation oncology. A comprehensive review.	Translational Oncology
Collado-Borrell R., Escudero-Vilaplana V., Ribed-Sanchez A., Ibanez-Garcia S., Herranz-Alonso A., Sanjurjo-Saez M.	Smartphone applications for cancer patients; what we know about them?.	Farmacia hospitalaria : organo oficial de expresion cientifica de la Sociedad Espanola de Farmacia Hospitalaria
Radovic A., Vona P.L., Santostefano A.M., Ciaravino S., Miller E., Stein B.D.	Smartphone Applications for Mental Health.	Cyberpsychology, behavior and social networking
Nieto-Gutierrez W., Aguirre-Tipismana L., Torres-Mallma C., Salazar-Rojas R., Taype-Rondan A.	Features of Zmobile provider education applications for prehospital trauma life support.	Journal of Emergency Medicine, Trauma and Acute Care
Nicholas J., Larsen M.E., Proudfoot J., Christensen H.	Mobile Apps for Bipolar Disorder: A Systematic Review of Features and Content Quality.	Journal of medical Internet research

Authors	Title	Journal
Sucala M., Cuijpers P., Muench F., Cardos R., Soflau R., Dobrean A., Achimas-Cadariu P., David D.	Anxiety: There is an app for that. A systematic review of anxiety apps.	Depression and Anxiety
Schoeppe S., Alley S., Rebar A.L., Hayman M., Bray N.A., Van Lippevelde W., Gnam J.-P., Bachert P., Direito A., Vandelanotte C.	Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: A review of quality, features and behaviour change techniques.	International Journal of Behavioral Nutrition and Physical Activity
Tellez W.A., Nieto-Gutierrez W., Taype-Rondan A.	Sunscreen mobile apps: A content analysis.	European Research in Telemedicine
Hoeppner B.B., Schick M.R., Kelly L.M., Hoeppner S.S., Bergman B., Kelly J.F.	There is an app for that - Or is there? A content analysis of publicly available smartphone apps for managing alcohol use.	Journal of Substance Abuse Treatment
Meghani S.H., MacKenzie M.A., Morgan B., Kang Y., Wasim A., Sayani S.	Clinician-Targeted Mobile Apps in Palliative Care: A Systematic Review.	Journal of Palliative Medicine
Edwards E.A., Lumsden J., Rivas C., Steed L., Edwards L.A., Thiyagarajan A., Sohanpal R., Caton H., Griffiths C.J., Munafo M.R., Taylor S., Walton R.T.	Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps.	BMJ open
Shaia K.L., Farag S., Chyjek K., Knopman J., Chen K.T.	An Evaluation of Mobile Applications for Reproductive Endocrinology and Infertility Providers.	Telemedicine journal and e-health : the official journal of the American Telemedicine Association
Ye Q., Khan U., Boren S.A., Simoes E.J., Kim M.S.	An Analysis of Diabetes Mobile Applications Features Compared to AADE7TM: Addressing Self-Management Behaviors in People With Diabetes.	Journal of Diabetes Science and Technology
Hoppe C.D., Cade J.E., Carter M.	An evaluation of diabetes targeted apps for Android smartphone in relation to behaviour change techniques.	Journal of human nutrition and dietetics : the official journal of the British Dietetic Association
Househ M., Hossain N., Jamal A., Zakaria N., Elmetwally A., Alsalamah M., Khalifa M.	A cross-sectional content analysis of Android applications for asthma.	Health informatics journal
Thornton L., Quinn C., Birrell L., Guillaumier A., Shaw B., Forbes E., Deady M., Kay-Lambkin F.	Free smoking cessation mobile apps available in Australia: a quality review and content analysis.	Australian and New Zealand journal of public health

Authors	Title	Journal
Lambert K., Mullan J., Mansfield K., Owen P.	Should We Recommend Renal Diet-Related Apps to Our Patients? An Evaluation of the Quality and Health Literacy Demand of Renal Diet-Related Mobile Applications.	Journal of renal nutrition : the official journal of the Council on Renal Nutrition of the National Kidney Foundation
Leigh S., Ouyang J., Mimnagh C.	Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder.	Evidence-based mental health
Zhang M.W.B., Ho R.C.M., Loh A., Wing T., Wynne O., Chan S.W.C., Car J., Fung D.S.S.	Current status of postnatal depression smartphone applications available on application stores: An information quality analysis.	BMJ Open
Zhou A.H., Patel V.R., Baredes S., Eloy J.A., Hsueh W.D.	Mobile Applications for Allergic Rhinitis.	Annals of Otology, Rhinology and Laryngology
Sharpe J.D., Kamara M.T.	A systematic evaluation of mobile apps to improve the uptake of and adherence to HIV pre-exposure prophylaxis.	Sexual Health
Casale M., Costantino A., Rinaldi V., Forte A., Grimaldi M., Sabatino L., Oliveto G., Aloise F., Pontari D., Salvinelli F.	Mobile applications in otolaryngology for patients: An update.	Laryngoscope Investigative Otolaryngology
Jovicic S., Siodmiak J., Watson I.D.	Quality evaluation of smartphone applications for laboratory medicine.	Clinical Chemistry and Laboratory Medicine
Fougerouse P.-A., Yasini M., Marchand G., Aalami O.O.	A Cross-Sectional Study of Prominent US Mobile Health Applications: Evaluating the Current Landscape.	AMIA ... Annual Symposium proceedings. AMIA Symposium
Yu J.S., Kuhn E., Miller K.E., Taylor K.	Smartphone apps for insomnia: examining existing apps' usability and adherence to evidence-based principles for insomnia management.	Translational behavioral medicine
Khander A., Farag S., Chen K.T.	Identification and Evaluation of Medical Translator Mobile Applications Using an Adapted APPLICATIONS Scoring System.	Telemedicine journal and e-health : the official journal of the American Telemedicine Association
Ben-Mussa A., Paget A.M.	Popular apps on the medical category targeting patients and the general public in the United Kingdom: Do they conform to the Health On the Net Foundation principles?.	Health informatics journal
Wisniewski H., Liu G., Henson P., Vaidyam A., Hajratalli N.K., Onnela J.-P., Torous J.	Understanding the quality, effectiveness and attributes of top-rated smartphone health apps.	Evidence-based mental health

Authors	Title	Journal
Brown H.M., Bucher T., Collins C.E., Rollo M.E.	A review of pregnancy iPhone apps assessing their quality, inclusion of behaviour change techniques, and nutrition information.	Maternal and Child Nutrition
Talwar D., Yeh Y.-L., Chen W.-J., Chen L.-S.	Characteristics and quality of genetics and genomics mobile apps: a systematic review.	European Journal of Human Genetics
Owens O.L., Beer J.M., Reyes L.I., Thomas T.L.	Systematic Review of Commercially Available Mobile Phone Applications for Prostate Cancer Education.	American journal of men's health
Kwan V., Bihelek N., Anderson V., Yeates K.	A Review of Smartphone Applications for Persons With Traumatic Brain Injury: What Is Available and What Is the Evidence?.	The Journal of head trauma rehabilitation
Carmody J.K., Denson L.A., Hommel K.A.	Content and Usability Evaluation of Medication Adherence Mobile Applications for Use in Pediatrics.	Journal of pediatric psychology
Grundy Q., Chiu K., Held F., Continella A., Bero L., Holz R.	Data sharing practices of medicines related apps and the mobile ecosystem: Traffic, content, and network analysis.	BMJ (Online)
Singh K., Diamantidis C.J., Ramani S., Bhavsar N.A., Mara P., Warner J., Rodriguez J., Wang T., Wright-Nunes J.	Patients' and Nephrologists' Evaluation of Patient-Facing Smartphone Apps for CKD.	Clinical journal of the American Society of Nephrology : CJASN
Luo D., Wang P., Lu F., Elias J., Sparks J.A., Lee Y.C.	Mobile Apps for Individuals with Rheumatoid Arthritis: A Systematic Review.	Journal of Clinical Rheumatology

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	N/A
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	N/A
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	N/A
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	9
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	9
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	S1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	N/A
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	N/A
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	12
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	13
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	17

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.