A SCOPING REVIEW OF THE USE OF DATA ANALYTICS FOR THE EVALUATION OF MHEALTH APPLICATIONS

ANÉ VAN SCHALKWYK
Department of Industrial Engineering, Stellenbosch University, South Africa
19800444@sun.ac.za

SARA GROBBELAAR
Department of Industrial Engineering, Stellenbosch University, AND DST-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy (SciSTIP), Stellenbosch University, South Africa
ssgrobbelaar@sun.ac.za

MARLIEN HERSELMAN
Next Generation Enterprises and Institutions, CSIR, PO Box 395, Pretoria, 0001 AND Department of Industrial Engineering, Stellenbosch University, South Africa
mherselman@csir.co.za

ABSTRACT

Background: The development and use of mobile health applications (mHealth Apps) have increased drastically over the last few years. There is an evident lack of the needed evidence to support the scalability and sustainability of these mHealth Apps. The vast number of mHealth Apps commercially available on App Stores makes it difficult to select the most effective and trustworthy Apps. Continuous evaluation of these Apps needs to be implemented or conducted to establish the needed evidence-base. Traditional methods for evaluating health interventions are not suitable for the quick and agile approach needed to develop and maintain Apps. In search of alternative and consistent standards and guidelines new evaluation methods are proposed. The aim of this article is to investigate the state of the literature in this regard. A scoping review is conducted to identify how mHealth Apps are currently evaluated, with a focus on data analytics.

Method: Five stages of conducting a scoping review are implemented: 1) identifying the research questions 2) identifying the relevant studies 3) study selection 4) data charting and 5) collecting, summarising, and reporting the results. Electronic databases used for the literature search included Scopus, Google Scholar, PubMed, and JMIR. Using the identified search terms for the scope-specific contents and two levels of screening 43 documents were included for the final data extraction.

Results: There is an increasing trend within the published works in our database, specifically peaking in 2018 and 2019. Usability and Effectiveness Evaluations are the most common reported evaluations for mHealth Apps. User surveys are the most frequently used traditional evaluation method, together with statistical analysis. Newly proposed evaluation methods include log or usage data analytics or implementing new frameworks.

Conclusion: This scoping review highlighted the value of conducting further research for the standardization of the evaluation of mHealth Apps. A growing trend and implementation of data analytics are observed within the health domain. We conclude that a framework that incorporates

data analytic tools and techniques would be of great value to mHealth App developers, evaluators, and users.

Key words: mHealth Apps; Data analytics; Evaluation; Scoping Review

INTRODUCTION

The development, availability, and use of mobile health applications (mHealth Apps) have expanded rapidly. The US Food and Drug Administration estimates that in 2018 the total downloaded mHealth Apps are close to 1.7 billion downloads, with 325 000 different Apps available across various health domains (Henson *et al.*, 2019)). mHealth Apps are enabling technologies and provide various benefits which include for example: increased access to healthcare services, education, and awareness; decreased costs associated with healthcare delivery; improved management and prevention of chronic diseases or symptoms; and data collection and distribution (Osunyomi and Grobbelaar, 2015; Silva *et al.*, 2015).

Although some progress has been made in terms of better understanding the innovation systems and ecosystems through which mHealth Apps are developed as well as the development and cocreation on the platform level, guidance on the evaluation of the technology remains lacking. (Herman, Grobbelaar and Pistorius, 2018; Ngongoni, Grobbelaar and Schutte, 2018; van der Merwe and Grobbelaar, 2018; van der Merwe, Grobbelaar and Bam, 2019). There remains a lack of standards, policies, and evidence-based results and support for the sustainability of mHealth Apps (Henson *et al.*, 2019).

To build the required evidence-base, appropriate evaluations are required throughout the entire lifecycle from development to implementation and operation of the Apps (White *et al.*, 2016). However, the quick, agile, and iterative characteristics used for developing and maintaining Apps propose a challenge for the traditional research and evaluation methods to date (Pham *et al.*, 2018). Randomized Controlled Trials (RCTs) used to be the "golden standard" for evaluating health interventions (Kumar *et al.*, 2013). However, RCTs are characterized by long time lags (takes an average of 5.5 years to complete), high costs, and the need for random assignment and rigid protocols (Pham, Wiljer and Cafazzo, 2016). These characteristics pose a serious problem for technological developments like mHealth Apps that could be obsolete by the time the RCT is completed, have increased costs due to trial expenses, or in need of critical updates from the initial App that was evaluated (Kumar *et al.*,2013).

Kumar *et al.* (2013) explains that to be able to keep up with the fast pace development and implementation of Apps, some of the rigid protocols and time-consuming steps of RCTs are omitted which results in a lack of outcome evaluations (going from pilot to implementation phase) and a resulting lack of evidence of the long-term value of these mHealth Apps. To avoid this scenario and to be able to build a strong evidence-base for the scalability and sustainability of mHealth Apps new methods of evaluation that suits the unique characteristics and requirements of developing and operating mHealth Apps should be considered. The problem with deviating from RCTs for evaluating mHealth Apps is that unlike the RCT that has been tried and tested for years, other methodologies for evaluating mHealth Apps are not yet standardized or consistently used (Stragier *et al.*, 2019).

Research studies have started investigating the use of the vast amount of data generated by the mHealth Apps for continuous evaluation purposes. Evidence suggest that these types of usage or log data analyses could be of considerable value above and beyond the traditional evaluation approaches, but that the most promising aspect is mainly focused on a conceptual level that is not yet exploited to its full extent (Sieverink *et al.*, 2017). Liang *et al.* (2018) reports that with the support of appropriate data analytics and data mining platforms data could be used and analysed, thereby providing valuable outcomes back to the users in an understandable and uncomplicated manner. Data quality, collection, and lengthy cleaning processes used to be the biggest barriers or concerns of using data for evaluation or optimisation purposes. However, with readily available data generated from mHealth Apps, and the efficient data capture protocols that could be implemented from the initial development stages, the potential of continuous or semi-automated evaluation is within reach (Miller *et al.*, 2019).

With the promise and potential of improved evaluation methodologies specifically for mHealth Apps, it is prudent to investigate what has been done and reported on in the current literature. A scoping review is ideal for setting research agendas and can be used to determine the value of conducting a full systematic review (Tricco *et al.*, 2016). Scoping reviews highlight gaps within the literature by identifying key concepts, theories, and sources of evidence, and thereby systematically maps the literature available on a specific topic – especially within emerging fields that have not been comprehensively reviewed (Tricco *et al.*, 2016). Thus, with the broad spectrum of mHealth, Data analytics, and Monitoring & Evaluation (M&E) the identified research topic would benefit greatly from the approach and findings of a scoping review.

The aim of this scoping review is to investigate how mHealth Apps are evaluated, with a specific focus on how data analytics are being applied in this field. This aim is accomplished by (1) identifying the relevant literature relating to mHealth App evaluations, (2) analysing, synthesising and reporting all research findings relating to (a) the environments where mHealth apps are applied, (b) how mHealth apps are evaluated, and (c) the types of data analytic tools and techniques in the mHealth environment. To the knowledge of the author no prior scoping review with this aim has been completed.

These objectives formulated the review questions to be answered by this research paper as follows:

i. RQ1: What trends are observed within the selected literature (with regards to the Authors/Year of publications/Journals)?
ii. RQ2: What aspects of mHealth Apps are evaluated most frequently?
iii. RQ3: When (at what stage) does the evaluation take place?
iv. RQ4: What evaluation tools or techniques are used to evaluate mHealth Apps?
v. RQ5: What data analytic tools or techniques are used for the evaluation of mHealth Apps?

The findings from this scoping review identify the gap in the literature that warrants a further in-depth systematic review. Thereby the existing evidence can be mapped and applied towards a standardization or practical framework for how mHealth Apps could be evaluated using data analytics. This will allow new research to be contributed to the Health, Data Analytics, and M&E domains. It could assist potential and current mHealth App developers for improved development and evaluation methodologies and decision-making, while also increasing the needed evidence-base for selecting, using, and updating mHealth Apps.

This scoping review highlighted the increasing rates of publications and changes in mHealth App evaluations. It shows the complex field that mHealth is developing towards that requires the collaboration of distinct disciplines of computer science, healthcare, and M&E. The literature reviewed identified the need for a standardized framework that can be practically applied to standardize and incorporate data analytics for continuous evaluation practices of developed and developing mHealth Apps.

METHODOLOGY

This section will provide details about the specific approach that was followed for conducting the scoping review.

Phase 1: Develop a Protocol

The review protocol provides the needed background, predefined objectives, methods, and detailed plan for conducting the scoping review. The protocol was developed by following the scoping review guidelines provided by the Joanna Briggs Institute as described by Tricco *et al.*, (2016). Tricco *et al.* developed the scoping review methodological guidelines by conducting a scoping review of all the requirements for a scoping review. Together with these guidelines, the following scoping review stages are conducted (Arksey and O'Malley, 2005):

i. Stage 1: Identifying the research question.
ii. Stage 2: Identifying relevant studies.
iii. Stage 3: Study Selection.
iv. Stage 4: Charting the Data.
v. Stage 5: Collecting, summarising, and reporting the results.

These stages and guidelines are considered comprehensive and ensures that all needed information is included in this review paper. The draft protocol was revised based on feedback as provided by Professors Sara Grobbelaar and Marlien Herselman. If required, the final version can be requested from the corresponding author.

Eligibility Criteria

Clearly defined eligibility criteria is developed and used to narrow down the inclusion of available literature for scope specific content, while ensuring repeatability and reliability of the scoping process. The criteria are developed by using the specific research questions as identified in Section 1. A scoping review is an iterative process. As the review is conducted new criteria are formulated and added accordingly. The final inclusion and resulting exclusion criteria for selecting the applicable literature for review is provided in *Table 1*. Literature is included if it meets all the inclusion criteria or excluded if one of the exclusion criteria is applicable.

Table 1: Inclusion and Exclusion Criteria for Literature Selection

Code	Inclusion Criteria	Code	Exclusion Criteria
IC1	Article focus is on the mHealth domain.	**EC1**	Paper not related to the mHealth domain.
IC2	Mentions/Evaluation of applications.	**EC2**	No evaluation mentioned or conducted.
IC3	Provides information regarding the use of data analytics.	**EC3**	No information or indication of data analytics.
IC4	Paper available in English	**EC4**	Evaluation of specific function not overall application.

Phase 2: Identification of Relevant Data Sources

Data Collection

Relevant literature is identified by using specific search terms and search engines or databases. The electronic databases that were used for this scoping review included: Scopus, Google Scholar, PubMed, and the Journal of Medical Internet Research (JMIR). These specific databases or journals were selected due to the vast range and inclusion of peer-reviewed literature provided from various fields and domains. The two specific journal search engines that were used, namely PubMed and JMIR, is included because of its high publication rate of scope-related articles identified throughout the initial Scopus and Google Scholar searches. Grey literature was also sourced using Google to identify any relevant data that might not be published or peer reviewed. The literature search was conducted from the 1st to the 15th of June 2020.

To avoid bias in selecting literature for inclusion it is important to source data from a broad spectrum of authors, institutions, and publishers. Therefore, the search strategy used had no limit on author or publication, year of publication, or study design, but was limited to 'English Only' literature. Each electronic database made use of a combination of specific search terms (as provided in *Table 2*) to identify scope specific content. The search terms were assigned to three domains namely: mHealth Apps, Data analytics, and Evaluation, and was focused on the title, abstract, and keywords of the publications. Harzing's Publish or Perish software was used for the Scopus and Google Scholar search and selection process. The initial search results yielded: 1000+ results for Google Scholar and JMIR; 55 results for Scopus; 170 results for PubMed; and 7 potential articles in the Grey Literature domain.

Since only one researcher was available for screening it is not feasible to screen all 1000+ records. It was decided that an acceptable general overview would be obtained by only considering the top 20 records from each database that had more than 100 potential papers. The results were ranked according to the relevance to the entered search terms or by using Harzing's h-index (for the Google Scholar search). The h-index indicates the impact of the research by considering the number of citations and authors. This resulted in a total of 122 potential records that should be screened for selection. The data selection process will be explained in the following subsection.

Table 2: Search Terms for Data Collection

Scope	Search Terms
Data Analytics	Data analysis* OR Big data OR machine learning OR data mining OR log data* OR artificial intelligence
mHealth Apps	mHealth App* OR mobile health app* OR digital health app* OR medical app*
Evaluation	Eval* OR framework OR assessment

Data Selection

By applying the eligibility criteria as shown in *Table 1*, all 122 of the search results collectively identified by the electronic databases undergoes two levels of screening to determine whether the paper should be included for the final data extraction. The data selection process is conducted as displayed in *Figure 1*. Level 1 screening involves applying the eligibility criteria to the title and abstracts of each paper. Record is kept of which papers are excluded and which are included and the reason thereof. If any uncertainties existed about whether the paper should be included or not, it was kept for Level 2 screening. As shown in *Figure 1*, of the 122 initial records 53 were excluded during the Level 1 screening. EC4 being the main reason for exclusion (n= 16). Besides the eligibility criteria, papers are also removed if they are duplicates, non-English, or not accessible. Level 1 screening identified 12 duplicate papers. The included papers are used for reference scanning to identify additional potential papers that was not included from the initial databases search. From the reference scanning 13 additional records were included, which resulted in a total of 82 records eligible for Level 2 screening.

Level 2 screening involves applying the eligibility criteria (from *Table 1*) to the full articles. From the Level 2 screening 26 records were excluded, with EC3 being the main reason for exclusion (n=10). Only one duplicate record was identified and removed. Finally, a total of 11 records were not accessible despite the authors best attempts, and 2 were non-English. This resulted in a total of 43 records eligible for data extraction process as will be explained in the following subsection. The 43 records are referenced from 1 to 43 (as shown in *Table 3*).

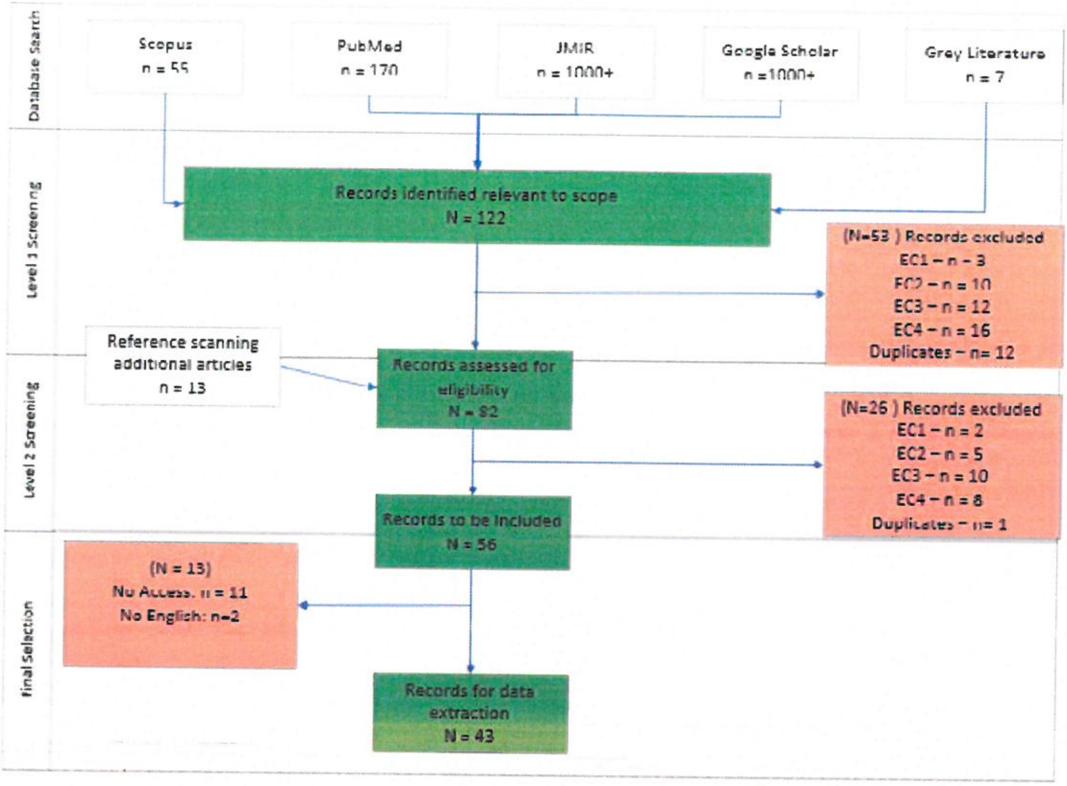

Figure 1: Data Selection Process

Data Charting

Data charting is completed by rereading each paper and identifying and recording specific similarities or trends to be able to obtain a fundamental understanding of the current research field. All the records that were selected for final data extraction were imported into Atlas.ti. Atlas.ti is a software program that is used for the qualitative analysis of research by providing the functionality to code, categorize, and visually display unstructured text – as would be desired to scope the current layout of available literature. The predefined charting form was developed from the research questions and refined throughout the review process as new understanding and information are obtained. Charting focuses on characterizing the literature not analysing the outcomes, and therefore the coding functionality provided by Atlas.ti was applied. The main categories and components that were focused on was the app aspects evaluated, the methods used for the evaluation, and the data analytic tools and techniques that were used or mentioned in each article. These categories provide the needed data to answer the research questions provided in Section 1. The final refined charting form is provided in the results section that follows.

REVIEW ANALYSIS AND RESULTS

This section provides the findings from the data extraction and coding process conducted on the 43 selected records. It aims to achieve the mentioned research objectives by answering the research questions as provided. The selected 43 records consisted of two conference papers, two book chapters, transcripts from a course video (from grey literature) and the remaining majority (38) of journal articles. As mentioned previously, Atlas.ti was used for coding and the identified categories coded from the papers can be observed in *Figure 2*. Each category also highlights what review question(s) it is linked to (shown by the blue triangle in the corner). The charted data findings are explained in detail in this section that follows.

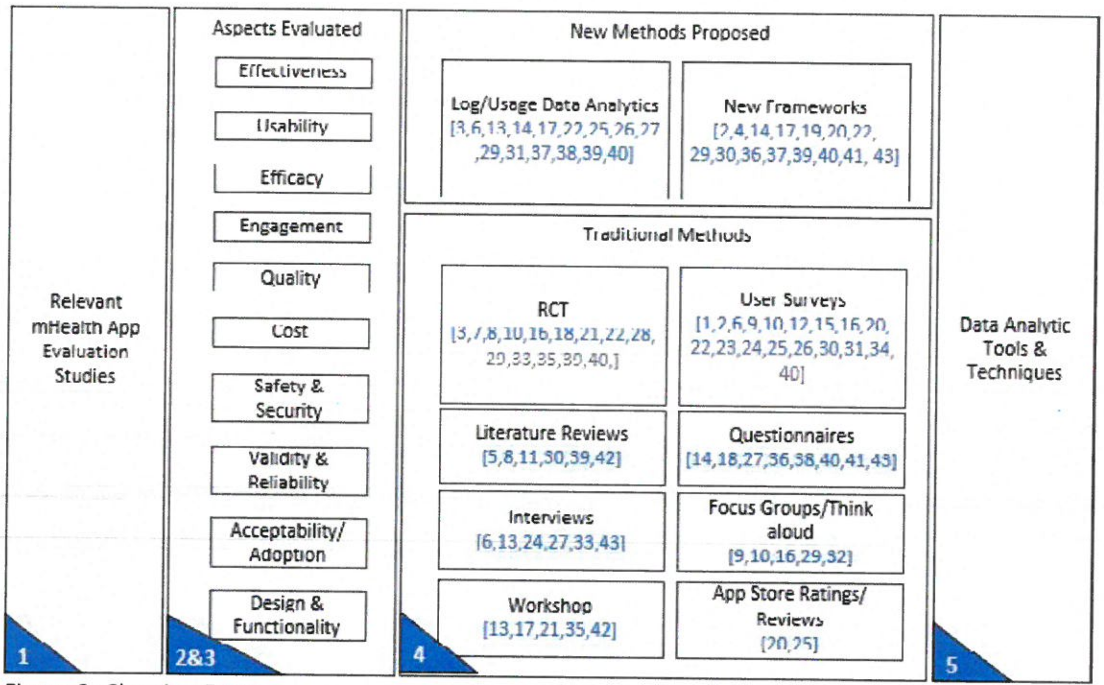

Figure 2: Charting Form

RQ1: What trends are observed within the selected literature?

As mentioned, the majority of the selected literature are journal articles (88%). The articles are sourced from 18 different journals, with the most popular being the JMIR mHealth and uHealth as it published 35% of the selected articles (n=11), followed by JMIR that published 13% (n=4). Dimensions online platform was used together with Microsoft Excel to analyse the specifics about the research publications and their current impact. Firstly, by visualising the 'year of publications' of the selected research and comparing it to the overall publication rates as provided by Dimensions it highlights the current trends of publication rates within the selected scope. *Figure 3* provides the year of publications of the selected literature, and *Figure 4* provides the year of publications of the overall published works when entering the search terms (refer to *Table 2*) into Dimensions. Both *Figures 3* and *4* shows an increasing trend observed within the publications of this research field. Clear visible peaks are observed between 2018 and 2019, and it should be noted that the 2020 published data could still increase as the scoping review was conducted before the yearend of 2020.

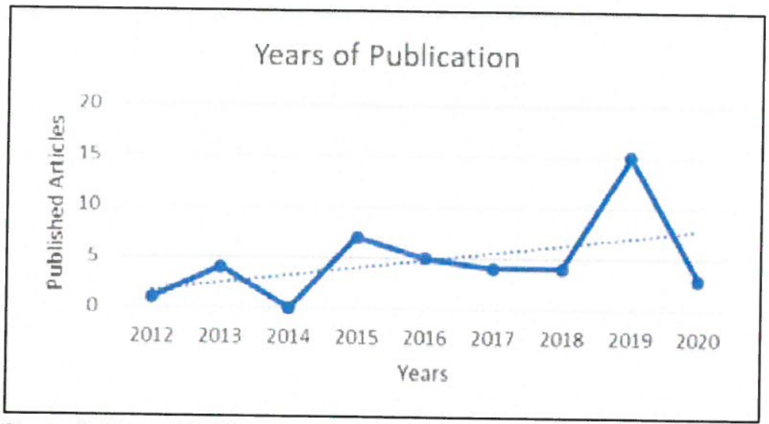

Figure 3: Year of Publication of Selected Literature

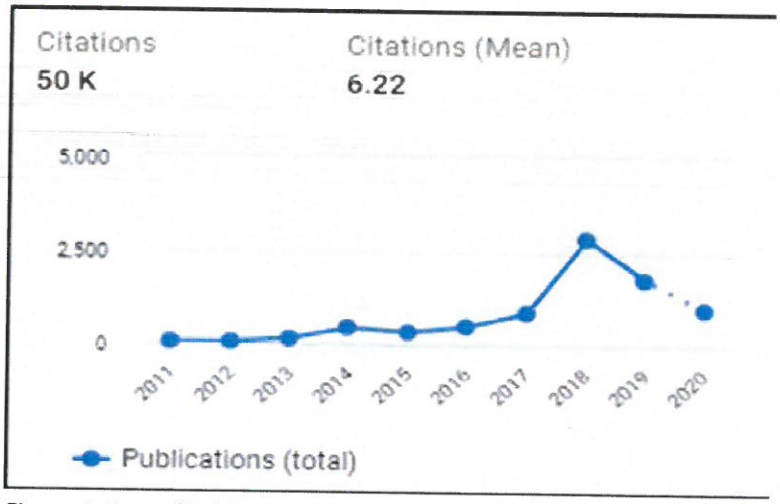

Figure 4: Year of Publication of Total Publications (sourced from Dimensions (2020))

Furthermore, *Dimensions* (2020) also provides the Research Categories and Content Types of the overall publications (when entering the search terms of *Table 2*). It shows that the most popular category is the Medical and Health Sciences (40%) followed by Public Health and Health Services (32%) and Information and Computing Sciences (12%). The content breakdown is given as: 8023 Publications, 11 Patents, 4 Clinical Trials, and 130 Policy Documents. This information is useful to identify where the focus area within this scope lies, and what documents to consider when doing additional research in this domain. Lastly, the most common authors identified through the selected literature appears to be Joseph A. Cafazzo, Quynh Pham, and Paul Kreb. Upon inspection the identified authors have 31, 8, and 6 published articles in JMIR respectively. Knowing the authors that are popular within the research scope allows for more focused literature searches to be conducted, and expansion possibilities in the inclusion criteria.

RQ2 & RQ3: What aspects of mHealth Apps are evaluated most frequently, and at what stage does the evaluation take place?

The coding process and categorization of the literature identified recurring similarities of aspects that researchers focus on when evaluating mHealth Apps as shown in *Figure 2*. *Table 3* provides a breakdown of what aspect(s) of the mHealth App are evaluated by each individual article. This breakdown shows how the aspects are not mutually exclusive, and highlights what the evaluations are focused on most in the current literature and provides the percentage of occurrences for each aspect evaluation. As shown in *Table 3*, 76.74% of the selected literature evaluated or mentioned the importance of evaluating the Usability of the mHealth Apps.

Usability is defined differently amongst the different sources but refers to for example the usefulness (i.e. using the App for its perceived or stated purpose), ease of use, frequency or duration of use, or the specific use of the technology (i.e. bugs in software). Usability is also considered as one facet of engagement and encompasses the effectiveness, efficiency, and satisfaction aspects of the technology or evaluation. Therefore, it makes sense that usability would be the most popular aspect to evaluate since it correlates with or includes several of the other identified aspects. The least mentioned aspects of evaluation are Validity and Reliability (9.30%).

Table 3: Specific Aspects Evaluated as Coded in Selected Sources

Ref #	Source	Effectiveness	Usability	Efficacy	Engagement	Quality	Acceptance & Adoption	Safety & Security	Valid & Reliable	Design & Functionality & Aesthetics	Cost
1	(Al-Azzam, Alazzam and Al-Manasra, 2019)	-	✓	-	-	-	✓	✓	-	✓	✓
2	(Bot *et al.*, 2016)	-	✓	-	✓	-	-	-	-	-	-
3	(Chen *et al.*, 2012)	✓	-	✓	✓	-	-	-	-	-	-
4	(Cui *et al.*, 2016)	✓	-	-	-	-	-	-	-	-	-
5	(Davies *et al.*, 2017)	-	✓	-	✓	✓	-	-	-	-	-
6	(Ding *et al.*, 2019)	✓	✓	-	-	-	-	✓	✓	-	✓
7	(Free *et al.*, 2013)	✓	-	-	-	-	-	-	-	-	-
8	(Friesen, Hamel and McLeod, 2013)	-	✓	-	-	-	-	✓	-	✓	✓
9	(Hamine *et al.*, 2015)	✓	✓	✓	-	-	✓	-	-	-	-

10	(Jeon *et al.*, 2016)	✓	✓	✓	-	-	-	-	-	✓	-
11	(Kernebeck *et al.*, 2019)	✓	✓	-	-	-	✓	-	-	-	-
12	(Kocsis *et al.*, 2019)	✓	-	-	-	-	-	-	-	-	-
13	(Krebs and Duncan, 2015)	✓	✓	-	-	-	-	✓	-	-	✓
14	(Krebs *et al.*, 2019)	✓	-	-	-	-	✓	-	-	-	-
15	(Frisbee, 2016)	✓	✓	-	-	-	-	✓	-	-	-
16	(Kumar *et al.*, 2013)	✓	-	✓	-	-	-	✓	✓	-	-
17	(Lee *et al.*, 2015)	✓	✓	-	-	-	-	✓	-	-	-
18	(Liang *et al.*, 2018)	✓	✓	-	-	✓	-	✓	-	✓	-
19	(Messner *et al.*, 2020)	-	✓	✓	✓	✓	-	✓	✓	✓	-
20	(Michie *et al.*, 2017)	✓	✓	-	✓	-	-	-	-	✓	✓
21	(Miller *et al.*, 2019)	-	✓	-	✓	-	-	-	-	-	-
22	(Murnane, Huffaker and Kossinets, 2015)	✓	✓	✓	-	-	-	-	-	✓	-
23	(O'reilly *et al.*, 2018)	✓	✓	-	-	✓	-	-	-	✓	-
24	(Owen *et al.*, 2015)	-	✓	✓	✓	-	-	-	-	-	-
25	(Park *et al.*, 2018)	-	-	-	✓	-	-	-	-	-	-
26	(Peiris *et al.*, 2019)	✓	✓	-	-	-	-	-	-	-	-
27	(Pham *et al.*, 2019)	✓	-	✓	-	-	-	-	-	-	✓
28	(Pham *et al.*, 2018)	✓	✓	-	✓	-	✓	✓	-	-	-
29	(Pham, Wiljer and Cafazzo, 2016)	✓	✓	✓	-	✓	-	-	-	-	-
30	(Quan *et al.*, 2020)	✓	✓	-	✓	-	✓	-	-	-	-
31	(Ramukumba and Hagglund, 2019)	-	✓	-	-	-	-	✓	-	✓	-
32	(Rodrigues *et al.*, 2015)	-	✓	✓	-	-	-	✓	-	-	-
33	(Roosan *et al.*, 2019)	✓	✓	-	-	-	-	-	-	-	-
34	(Ryu *et al.*, 2017)	✓	-	✓	-	-	-	-	-	-	-

35	(Safi, Danzer and Schmailzl, 2019)	-	✓	-	-	-	✓	-	-	-	-
36	(Sieverink *et al.*, 2017)	-	✓	-	✓	✓	-	-	-	-	-
37	(Stragier *et al.*, 2019)	✓	✓	-	✓	-	-	-	-	-	-
38	(Tomlinson *et al.*, 2013)	✓	-	✓	-	-	-	-	-	-	✓
39	(Coursera, 2020)	-	✓	-	✓	✓	-	-	✓	✓	-
40	(Vriend, Coehoorn and Verhagen, 2015)	✓	✓	-	-	-	✓	-	-	-	-
41	(White *et al.*, 2019)	✓	✓	-	✓	✓	-	-	-	✓	-
42	(Wisniewski *et al.*, 2019)	✓	✓	-	✓	-	✓	-	-	-	✓
43	(Woods *et al.*, 2019)	-	✓	-	✓	-	✓	-	-	✓	-
	% of Sources	67.44	**76.74**	27.91	37.21	18.60	23.26	27.91	**9.30**	27.91	18.60

Depending on what aspect of the mHealth App is being evaluated (e.g. engagement, effectiveness, usability) or what evaluation tools or techniques are applied the stage or period when the evaluation will be conducted differs. The selected literature conducted the mHealth App evaluations at one of five identified stages namely: Early in the development; Later in development; During the pilot phase; On a launched or existing App; or Throughout which means during and after the App's development, or alternatively the researchers did not indicate when (*i.e.* Not Specified). The reviewed literature (refer to *Figure 5*) indicated that most mHealth Apps (48%) were evaluated after being launched such as the Apps found on the existing App Stores. *Figure 5* also highlights that in the current literature the

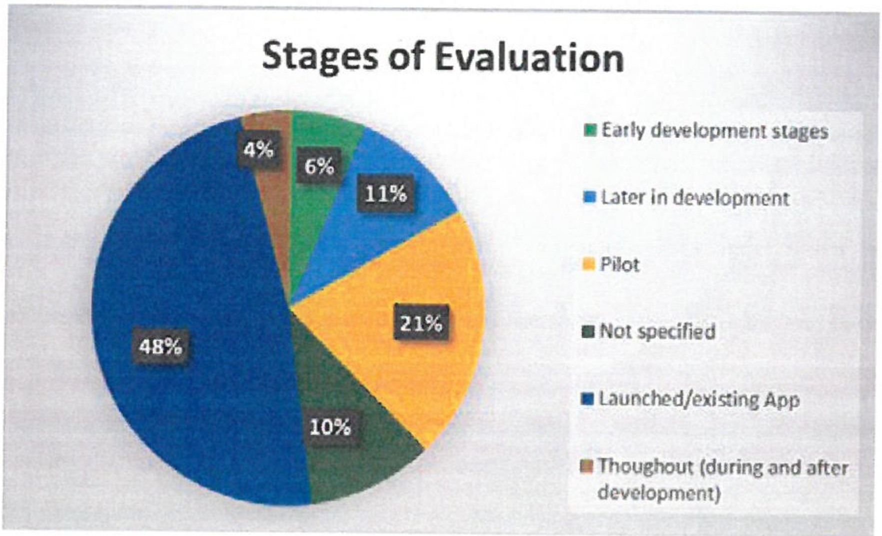

Figure 5: Breakdown for Stages of Evaluation

least common finding was continuous evaluations (4%) that evaluate from development to operational phases of the Apps.

RQ4: What evaluation tools or techniques are used to evaluate mHealth Apps?

The evaluation tools and techniques referred to in the articles can be categorized into traditional methods or newly proposed methods. As shown in *Figure 2* some references mention or apply more than one evaluation technique. Due to the specifics of the search terms traditional methods like the Random Controlled Trials (RCT) are mentioned or used less than would be expected as it does not necessarily involve data analytics (which would exclude it from the selection criteria). *Figure 6* visualises the different tools and techniques and the frequency that it was mentioned or used to evaluate mHealth Apps throughout the data charting process.

Again, the techniques are not mutually exclusive and are often used in parallel such as applying statistical analysis to evaluate the user survey responses or making use of a questionnaire while interviewing a user. This combination of techniques explains why statistical analysis is reported as the most frequent and why workshops, interviews, focus groups, and think aloud techniques have similar numbers. Based on the selected literature, from the traditional methods user surveys is the most popular, followed by RCTs. While the new methods report equal amounts of Log/Usage Data analyses compared with applying New Frameworks that can be used. The New Frameworks category can then be further analysed into specific frameworks that were mentioned or used. A total of 16 different frameworks are mentioned with only three of these occurring in more than one article specifically: Mobile App Rating Scale [29,36,38]; Multiphase Optimisation Strategy or MOST [29,36, 38]; and N-of-1 study [3,20,29].

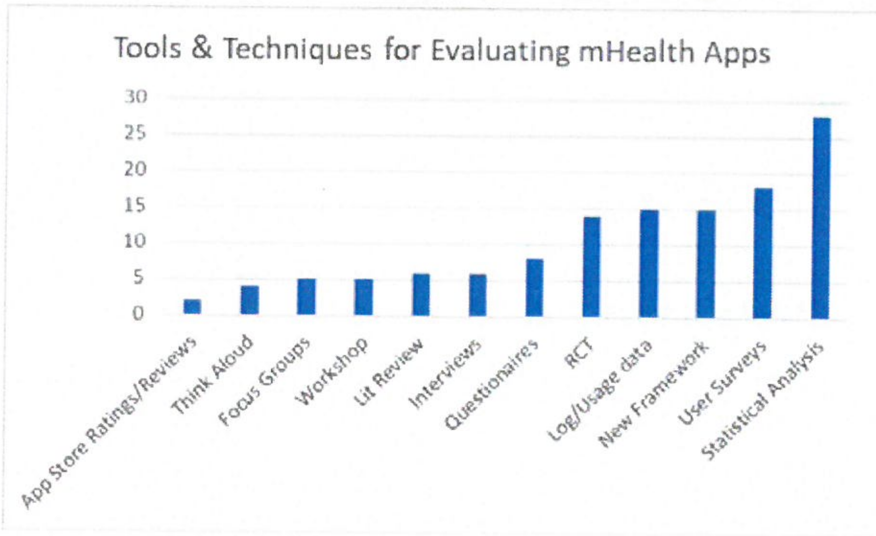

Figure 6: Tools and Techniques for Evaluating mHealth Apps

RQ5: What data analytic tools or techniques are used for the evaluation of mHealth Apps?

The final data charting category as seen in *Figure 2* is used to highlight what data analytics tools and techniques are specifically applied for the evaluation of mHealth Apps as highlighted by the selected literature. Eleven different data analytic techniques were identified as shown in *Figure 7*. As expected, Statistical Analysis is the most frequently used or mentioned technique (45%) as it is compatible with most evaluation techniques as mentioned previously. The second most used or mentioned technique is Usage Analysis (13%) which is similar to Log Analysis (9%) and consistent with the previously mentioned frequency of Usage Evaluations. The wide range of available or mentioned techniques shows the growing potential for applying data analytics for mHealth App evaluation purposes.

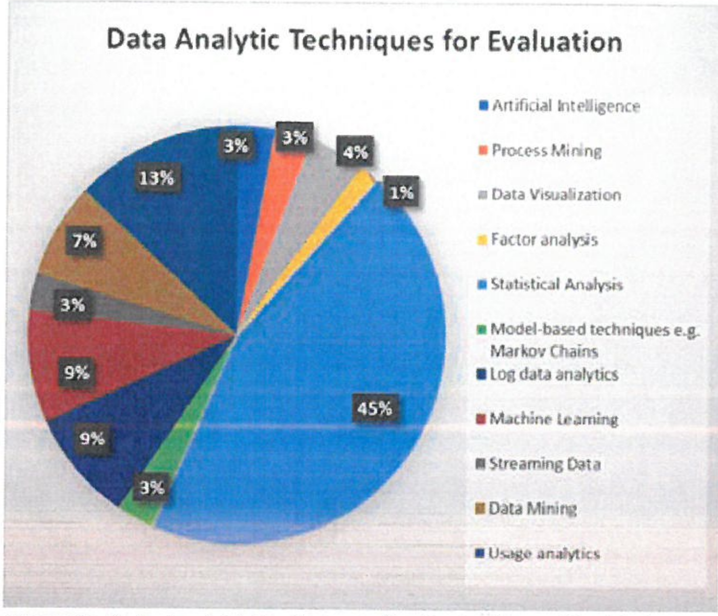

Figure 3: Data Analytic Techniques for Evaluating mHealth Apps

Finally, the tools for applying these data analytic techniques differ between articles with limited reoccurrences found. The tools mentioned include: SPSS; LifeGuide Visualisation Tool; InfoVis; Kibana; Weka Tool; Excel; R Software; Synapse; Google Analytics (GA); Flurry Analytics; and Perl. With only R Software mentioned 4 times, GA twice, and Flurry Analytics twice, while the remaining tools were only mentioned once. Most of the literature reviewed omitted to mention the specific tools used, and some mentioned a combination of the identified tools for applying specific data techniques.

CONCLUSION AND RECOMMENDATIONS

This scoping review was conducted using the stages as provided by Arksey & O'Malley (2005) together with the guidelines of the Johanna Briggs Institute (Tricco *et al.*, 2016). This ensured proper documentation and repeatability of the scoping review process. The scoping review highlighted the lack of evidence and standardization for the continuous evaluation of mHealth Apps. The literature gap identified showed a need for a standardized approach that specifies which tools and techniques to apply when evaluating mHealth Apps. This article showed the increasing rate of publications within a specific scope. Furthermore, it also highlighted the increased use of data analytics for the mHealth domain, and the expansion of available tools and techniques. Current literature is focused on usability evaluations of existing mHealth Apps, with user surveys being the most frequently applied method. Statistical analysis used to be the most common implementation of data analytics as it is compatible with most methods of evaluation, but the review also highlighted the increasing popularity of usage data analysis.

Review Limitations

The limitations of this scoping review include severe time constraints together with the fact that only one researcher conducted the screening and coding process. This might result in bias that exists where only the top 20 articles were considered in some search engines or where the screening and coding process is completed by the same researcher. However, the rigid guidelines followed with the proper documentation of the process is considered sufficient measures to reduce the bias. Since the scoping review only gives a general overview of what has been done (not in-depth like a systematized review) considering the top 20 search results are deemed adequate. Lastly, because of the specific search terms and eligibility criteria the resulting data might be skewed to show higher percentages of new evaluation methods (e.g. log data analyses) compared to traditional methods (e.g. RCTs) then if publications with evaluation without data analytics were also included.

Future Recommendations

Based on the findings of this scoping review it is prudent to next conduct a systematized review that focuses on the evaluation of mHealth Apps using log or usage data. The various evaluation frameworks found in literature lack proper guidance on how to set it up properly and would benefit from a practical framework that incorporates data analytic tools and techniques. Future work could also include researching the difference between Apps and mHealth Apps and whether their evaluations would differ. The current trend as highlighted by this review focuses on usability or effectiveness evaluations of mHealth Apps. Future work could investigate how these specific evaluations are being conducted, thereby identifying what should be implemented during the design phase and assist with building the needed evidence-base for the long-term value of mHealth Apps.

REFERENCES

Al-Azzam, M. K., Alazzam, M. B. and Al-Manasra, M. K. (2019) 'MHealth for decision making support: A case study of EHealth in the public sector', *International Journal of Advanced Computer Science and Applications*, 10(5). doi: 10.14569/ijacsa.2019.0100547.

Arksey, H. and O'Malley, L. (2005) 'Scoping studies: Towards a methodological framework', *International Journal of Social Research Methodology: Theory and Practice*. doi: 10.1080/1364557032000119616.

Bot, B. M. *et al.* (2016) 'The mPower study, Parkinson disease mobile data collected using ResearchKit', *Scientific Data*, 3. doi: 10.1038/sdata.2016.11.

Chen, C. *et al.* (2012) 'Making sense of mobile health data: An open architecture to improve individual- and population-level health', *Journal of Medical Internet Research*, 14(4). doi: 10.2196/jmir.2152.

Coursera (2020) 'eHealth: more than just an electronic record'. University of Sydney.

Cui, M. *et al.* (2016) 'T2DM self-management via smartphone applications: A systematic review and meta-analysis', *PLoS ONE*, 11(11), pp. 1–15. doi: 10.1371/journal.pone.0166718.

Davies, E. B. *et al.* (2017) 'Proportionate methods for evaluating a simple digital mental health tool', *Evidence-Based Mental Health*, 20(4). doi: 10.1136/eb-2017-102755.

Dimensions (2020). Available at: https://app.dimensions.ai/discover/publication?search_text=Data analytics AND evaluation AND mHealth Apps&search_type=kws&search_field=full_search (Accessed: 15 June 2020).

Ding, H. *et al.* (2019) 'Digital health for COPD care: The current state of play', *Journal of Thoracic*

Disease, pp. 2210–2220. doi: 10.21037/jtd.2019.10.17.

Free, C. *et al.* (2013) 'The Effectiveness of Mobile-Health Technologies to Improve Health Care Service Delivery Processes: A Systematic Review and Meta-Analysis', *PLoS Medicine*, 10(1). doi: 10.1371/journal.pmed.1001363.

Friesen, M. R., Hamel, C. and McLeod, R. D. (2013) 'A mHealth application for chronic wound care: Findings of a user trial', *International Journal of Environmental Research and Public Health*, 10(1). doi: 10.3390/ijerph10116199.

Frisbee, K. L. (2016) 'Variations in the Use of mHealth Tools: The VA Mobile Health Study', *JMIR mHealth and uHealth*, 4(3). doi: 10.2196/mhealth.3726.

Hamine, S. *et al.* (2015) 'Impact of mHealth chronic disease management on treatment adherence and patient outcomes: A systematic review', *Journal of Medical Internet Research*. doi: 10.2196/jmir.3951.

Henson, P. *et al.* (2019) 'Deriving a practical framework for the evaluation of health apps', *The Lancet Digital Health*. doi: 10.1016/S2589-7500(19)30013-5.

Herman, H., Grobbelaar, S. S. and Pistorius, C. W. I. (2018) 'Towards a framework for technology platform design, development and implementation in South African health: Preliminary validation', in *2018 3rd Biennial South African Biomedical Engineering Conference (SAIBMEC)*. IEEE, pp. 1–5. doi: 10.1109/SAIBMEC.2018.8363194.

Jeon, E. *et al.* (2016) 'Mobile apps providing tailored nursing interventions for patients with metabolic syndrome', in *Studies in Health Technology and Informatics*. doi: 10.3233/978-1-61499-658-3-510.

Kernebeck, S. *et al.* (2019) 'A tablet app– and sensor-based assistive technology intervention for informal caregivers to manage the challenging behavior of people with dementia (the InsidedeM study): Protocol for a feasibility study', *Journal of Medical Internet Research*, 8(2). doi: 10.2196/11630.

Kocsis, O. *et al.* (2019) 'Multi-model Short-term Prediction Schema for mHealth Empowering Asthma Self-management', in *Electronic Notes in Theoretical Computer Science*. doi: 10.1016/j.entcs.2019.04.007.

Krebs, P. *et al.* (2019) 'The Quitit coping skills game for promoting tobacco cessation among smokers diagnosed with cancer: Pilot randomized controlled trial', *Journal of Medical Internet Research*. doi: 10.2196/10071.

Krebs, P. and Duncan, D. T. (2015) 'Health App Use Among US Mobile Phone Owners: A National Survey', *JMIR mHealth and uHealth*. doi: 10.2196/mhealth.4924.

Kumar, S. *et al.* (2013) 'Mobile health technology evaluation: The mHealth evidence workshop', *American Journal of Preventive Medicine*. doi: 10.1016/j.amepre.2013.03.017.

Lee, Y. *et al.* (2015) 'Evaluation of mobile health applications developed by a tertiary hospital as a tool for quality improvement breakthrough', *Healthcare Informatics Research*. doi: 10.4258/hir.2015.21.4.299.

Liang, J. *et al.* (2018) 'Chinese mobile health APPs for hypertension management: A systematic evaluation of usefulness', *Journal of Healthcare Engineering*. doi: 10.1155/2018/7328274.

van der Merwe, E., Grobbelaar, S. and Bam, W. (2019) 'Exploring the functional dynamics of innovation for inclusive development innovation systems: a case study of a large scale maternal mHealth project in South Africa', *Innovation and Development*. Routledge, pp. 1–22. doi: 10.1080/2157930X.2019.1567884.

van der Merwe, E. and Grobbelaar, S. S. (Saartjie. (2018) 'Systemic policy instruments for inclusive innovation systems: Case study of a maternal mHealth project in South Africa', *African Journal of*

Science, Technology, Innovation and Development. doi: 10.1080/20421338.2018.1491678.

Messner, E. M. *et al.* (2020) 'The german version of the mobile app rating scale (MARS-G): Development and validation study', *Journal of Medical Internet Research*. doi: 10.2196/14479.

Michie, S. *et al.* (2017) 'Developing and evaluating digital interventions to promote behavior change in health and health care: Recommendations resulting from an international workshop', *Journal of Medical Internet Research*. doi: 10.2196/jmir.7126.

Miller, S. *et al.* (2019) 'A framework for analyzing and measuring usage and engagement data (AMUSED) in digital interventions: Viewpoint', *Journal of Medical Internet Research*. doi: 10.2196/10966.

Murnane, E. L., Huffaker, D. and Kossinets, G. (2015) 'Mobile health apps: Adoption, adherence, and abandonment', in *UbiComp and ISWC 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the Proceedings of the 2015 ACM International Symposium on Wearable Computers*. doi: 10.1145/2800835.2800943.

Ngongoni, C. N., Grobbelaar, S. S. and Schutte, C. S. L. (2018) 'Platforms in healthcare innovation ecosystems: The lens of an innovation intermediary', in *2018 3rd Biennial South African Biomedical Engineering Conference, SAIBMEC 2018*. doi: 10.1109/SAIBMEC.2018.8363191.

O'reilly, M. A. *et al.* (2018) 'A wearable sensor-based exercise biofeedback system: Mixed methods evaluation of formulift', *JMIR mHealth and uHealth*. doi: 10.2196/mhealth.8115.

Osunyomi, B. D. and Grobbelaar, S. S. (2015) 'Integrating eHealth in HIV/AIDS intervention programmes in South Africa', *SA Journal of Information Management*. doi: 10.4102/sajim.v17i1.623.

Owen, J. E. *et al.* (2015) 'mHealth in the Wild: Using Novel Data to Examine the Reach, Use, and Impact of PTSD Coach', *JMIR Mental Health*. doi: 10.2196/mental.3935.

Park, Y. R. *et al.* (2018) 'Managing patient-generated health data through mobile personal health records: Analysis of usage data', *JMIR mHealth and uHealth*. doi: 10.2196/mhealth.9620.

Peiris, D. *et al.* (2019) 'A smartphone app to assist smoking cessation among aboriginal australians: Findings from a pilot randomized controlled trial', *Journal of Medical Internet Research*. doi: 10.2196/12745.

Pham, Q. *et al.* (2018) 'An analytics platform to evaluate effective engagement with pediatric mobile health apps: Design, development, and formative evaluation', *JMIR mHealth and uHealth*. doi: 10.2196/11447.

Pham, Q. *et al.* (2019) 'A library of analytic indicators to evaluate effective engagement with consumer mHealth apps for chronic conditions: Scoping review', *JMIR mHealth and uHealth*. doi: 10.2196/11941.

Pham, Q., Wiljer, D. and Cafazzo, J. A. (2016) 'Beyond the Randomized Controlled Trial: A Review of Alternatives in mHealth Clinical Trial Methods', *JMIR mHealth and uHealth*. doi: 10.2196/mhealth.5720.

Quan, A. M. L. *et al.* (2020) 'Mobile clinical decision tools among emergency department clinicians:web-based survey and analytic data for evaluation of the ottawa rules app', *JMIR mHealth and uHealth*. doi: 10.2196/15503.

Ramukumba, M. M. and Hagglund, M. (2019) '"i feel like a nurse and my clients learn more": MHealth, Capacity Building and Empowerment in Community Based Care', in *Studies in Health Technology and Informatics*. doi: 10.3233/SHTI190163.

Rodrigues, R. *et al.* (2015) '"The phone reminder is important, but will others get to know about my illness?" Patient perceptions of an mHealth antiretroviral treatment support intervention in the HIVIND trial in South India', *BMJ Open*. doi: 10.1136/bmjopen-2015-007574.

Roosan, D. *et al.* (2019) 'Improving medication information presentation through interactive visualization in mobile apps: Human factors design', *Journal of Medical Internet Research*. doi: 10.2196/15940.

Ryu, B. *et al.* (2017) 'Impact of an electronic health record-integrated personal health record on patient participation in health care: Development and randomized controlled trial of myhealthkeeper', *Journal of Medical Internet Research*. doi: 10.2196/jmir.8867.

Safi, S., Danzer, G. and Schmailzl, K. J. G. (2019) 'Empirical research on acceptance of digital technologies in medicine among patients and healthy users: Questionnaire study', *Journal of Medical Internet Research*. doi: 10.2196/13472.

Sieverink, F. *et al.* (2017) 'Opening the Black Box of Electronic Health: Collecting, Analyzing, and Interpreting Log Data', *JMIR Research Protocols*. doi: 10.2196/resprot.6452.

Silva, B. M. C. *et al.* (2015) 'Mobile-health: A review of current state in 2015', *Journal of Biomedical Informatics*. doi: 10.1016/j.jbi.2015.06.003.

Stragier, J. *et al.* (2019) 'Data mining in the development of mobile health apps: Assessing in-app navigation through markov chain analysis', *Journal of Medical Internet Research*. doi: 10.2196/11934.

Tomlinson, M. *et al.* (2013) 'Scaling Up mHealth: Where Is the Evidence?', *PLoS Medicine*. doi: 10.1371/journal.pmed.1001382.

Tricco, A. C. *et al.* (2016) 'A scoping review on the conduct and reporting of scoping reviews', *BMC Medical Research Methodology*. doi: 10.1186/s12874-016-0116-4.

Vriend, I., Coehoorn, I. and Verhagen, E. (2015) 'Implementation of an App-based neuromuscular training programme to prevent ankle sprains: A process evaluation using the RE-AIM framework', *British Journal of Sports Medicine*. doi: 10.1136/bjsports-2013-092896.

White, B. *et al.* (2019) 'Gamifying breastfeeding for fathers: Process evaluation of the milk man mobile app', *Journal of Medical Internet Research*. doi: 10.2196/12157.

White, B. K. *et al.* (2016) 'Designing evaluation plans for health promotion mHealth interventions: A case study of the Milk Man mobile app', *Health Promotion Journal of Australia*. doi: 10.1071/HE16041.

Wisniewski, H. *et al.* (2019) 'Understanding the quality, effectiveness and attributes of top-rated smartphone health apps', *Evidence-Based Mental Health*. doi: 10.1136/ebmental-2018-300069.

Woods, L. S. *et al.* (2019) 'Patients' experiences of using a consumer mhealth app for self-management of heart failure: Mixed-methods study', *Journal of Medical Internet Research*. doi: 10.2196/13009.

back to paper index >> back to author index >>