A Methodologic Systematic Review of Mobile Health Behavior Change Randomized Trials

Theodora Oikonomidi, Alexandre Vivot, Viet-Thi Tran, Carolina Riveros, Elisabeth Robin, Philippe Ravaud

A Methodologic Systematic Review of Mobile Health Behavior Change Randomized Trials

Theodora Oikonomidi, MSc,[1,2] Alexandre Vivot, PhD,[1,2] Viet-Thi Tran, PhD,[1,2] Carolina Riveros, MPH,[1] Elisabeth Robin, MSc,[2] Philippe Ravaud, PhD[1,2,3]

From the [1]Clinical Epidemiology Unit, Hôtel-Dieu Hospital, Assistance Publique-Hôpitaux de Paris, Paris, France; [2]Université de Paris, CRESS, INSERM, INRA, F-75004 Paris France; and [3]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York

Address correspondence to: Alexandre Vivot, PhD, Centre d'Epidémiologie Clinique, INSERM U1153, Hôpital Hôtel-Dieu, 1 place du Parvis Notre Dame, 75004 Paris, France. E-mail: alexandre.vivot@aphp.fr.

Context: Mobile health (mHealth) helps providers offer accessible, affordable, tailored behavior change interventions. However, research assessing mHealth interventions may feature methodologic shortcomings and poor reporting. This review aims to summarize the characteristics, methods, and intervention reporting of RCTs evaluating mHealth behavior change interventions.

Evidence acquisition: This was a methodologic systematic review of RCTs assessing mHealth behavior change interventions, published in PubMed from January 1, 2014 to January 1, 2018, in journals with the upper half of Impact Factors (Clarivate Analytics). Three reviewers independently extracted sample characteristics. Primary outcomes were classified as patient-important or not using definitions from the literature. Any non–patient important outcomes were then re-classified by a panel of three patients. Intervention reporting was assessed by the mHealth Evidence Reporting and Assessment checklist. Data were analyzed in December 2018.

Evidence synthesis: Most of the 231 included RCTs assessed text messaging (51%) or smartphone app (28%) interventions aiming to change nutrition and physical activity (36%) or treatment adherence (25%). Only 8% of RCTs had a patient-important primary outcome, follow-up of ≥6 months, and intent-to-treat analysis. Most primary outcomes were behavioral measures (60%). Follow-up was <3 months in 29% of RCTs. Regarding reporting, 12 of the 16 checklist items were reported in less than half of RCTs (e.g., usability/content testing, 32%; data security, 13%).

Conclusions: Reports of RCTs assessing mHealth behavior change interventions lack information that would be useful for providers, including reporting of long-term intervention impact on patient-important primary outcomes and information needed for intervention replicability.

CONTEXT

The rapidly growing field of mobile health (mHealth), or the use of mobile phones and other wireless technology in medical care, could help providers offer behavior change interventions to their patients in a remotely delivered, affordable, and scalable manner.[1–4]

Systematic reviews have shown that mHealth interventions can promote change in a range of behaviors, in healthy and chronically ill populations, at least in the short term.[2,5] Traditional behavior change interventions are resource-intensive and difficult to implement at scale, but their mHealth counterparts can be delivered via devices already used widely in the population and require fewer economic and human resources.[6–8] Additionally, contrary to static interventions, mHealth interventions can adapt over time based on users' characteristics, prior response to the intervention, and contextual parameters, to select and deliver the right intervention components at the right time.[6,9] The increased interest in mHealth is mirrored by a surge in publications detailing the development and comparative effectiveness assessment of mHealth interventions.[10,11]

However, RCTs assessing behavioral interventions may have several shortcomings, including short follow-up, lack of patient-important primary outcomes, and incomplete intervention reporting.[12–14] Incomplete intervention reporting can make it impossible to replicate and implement interventions, or to group them together for evidence synthesis.[15,16] Describing mHealth behavior change interventions is particularly challenging because they are complex, with many technical aspects (e.g., software, data safety), and "active ingredients" in the form of behavioral techniques.[17–19] The mHealth Evidence Reporting and Assessment (mERA) checklist,

published in 2016, aims to improve reporting by helping authors describe key intervention characteristics specific to mHealth. However, the reporting completeness of mHealth interventions has not been assessed systematically.[19]

This study aimed to provide an overview of the characteristics and methods of recently published RCTs of mHealth behavior change interventions and to assess the completeness of intervention reporting. To this end, all RCTs published in high–Impact Factor journals in the past 4 years were reviewed.

EVIDENCE ACQUISITION

This was a prospectively registered methodologic review of published RCTs evaluating mHealth behavior change interventions (PROSPERO CRD42017065826). The protocol is available on the institutional website www.clinicalepidemio.fr/protocols/. This report follows PRISMA guidelines.[20] PRISMA consists of a 27-item checklist and a flow diagram that can guide authors in reporting systematic reviews and meta-analyses completely and transparently.

On January 5, 2018, the authors searched MEDLINE via PubMed for eligible studies using Medical Subject Headings terms and free-text words (Appendix Table 1). English-language reports of RCTs published between January 1, 2014 and January 1, 2018 were included if they assessed interventions with an mHealth component (standalone or part of a larger intervention) according to the definition of The Global Observatory for eHealth and aiming to change health behaviors.[12] To avoid restricting the search to specific behaviors and to obtain a broad overview of the literature, broad search terms (e.g., *behav**) were used. Only journal articles published in

the highest half of the average Journal Impact Factor (Clarivate Analytics) were included, because of the impact these journals have on shaping clinician and researcher views.[21] Secondary reports were included. When more than one publication referred to the same trial, the most recent publication was included. Reports of ongoing trials were excluded. Titles and abstracts and then full-text reports were screened independently by two reviewers (TO and AV) and disagreements were resolved by consensus. One researcher (TO) extracted data from all included studies by using a pre-specified, standardized form (Appendix Text 1). Two other researchers (ER and CR) independently extracted 25% of the included records in total. Information was extracted from the following sources: the included publication, supplementary files, and any sources cited by the authors (protocol, website, previous publication by the authors detailing intervention testing or development). Digital forms of presentation (e.g., screenshots) were also used. The extraction was duplicated in full for any items that did not have substantial inter-rater agreement ($\kappa<0.60$ or percentage agreement <80%). Disagreements were resolved by consensus, which was reached after consulting the pre-specified definitions for each extraction item and the RCT reports.

Extraction of Study and Intervention Characteristics

The following RCT and intervention characteristics were extracted: target behavior, participants' health condition, device used to deliver the intervention, intervention frequency, reporting of information on sensor accuracy, presence of behavior change techniques (e.g., self-monitoring, feedback),[12,22–24] the behavior change theory used to develop the intervention as reported by the authors, and the degree of intervention personalization (presence, timing, and input of personalization). Because careful reporting of information relevant to mHealth scale-up at the population level is needed,[13,25] the scalability barriers mentioned by the original authors were

extracted and information on the setting was collected (income level of the country where the RCT was conducted according to The World Bank classification).[26]

The following methodologic components were extracted: primary outcome, follow-up duration in months, use of intent-to-treat (ITT) analysis, and adverse events related to the intervention (Appendix Text 2 provides details on data extraction). All primary outcomes were classified as patient-important or not by two researchers (TO and VTT), following definitions previously used in the literature (including measures of mortality, clinical events, adverse events, function, pain, quality of life, and therapeutic decisions).[27,28] Outcomes not classified as patient-important at this stage were additionally examined by three patients (a 67-year-old man with Type 2 diabetes, a 31-year-old woman with chronic migraine, and a 23-year-old man with asthma, allergies, and a history of depression). In individual teleconference meetings, the researchers presented each primary outcome of each RCT to the participants, summarized by its target population and behavior. Participants were asked to assess if the outcome was in their opinion sufficiently important for patients. Outcomes were considered be patient-important if there was agreement between at least two patients.

Assessment of Intervention Reporting

The completeness of intervention reporting was assessed using the recently published mERA checklist.[19] The mERA consists of 16 items that represent a minimum set of information that should be reported to ensure intervention replicability. The mERA items cover the intervention content, technical features, implementation context, and participants' involvement and experience with the intervention. Checklist items were classified as "reported" if they were adequately described in the included paper and its supplementary files. If the item was missing

from the included paper but was found in a source cited by the authors (e.g., protocol), the item was classified as "reported in cited source." If the item was missing or unclear, it was classified as "unreported." Examples of good reporting are shown in Appendix Table 2.

Descriptive statistics (frequencies and percentages or medians with IQRs) were calculated using R, version 3.3.0.

EVIDENCE SYNTHESIS

In total, 1,934 citations were retrieved from PubMed (Figure 1). After screening, 231 publications were included (full list is shown in Appendix Text 3).

Study and Intervention Characteristics

Key study components are summarized in Table 1 and Figure 2, which shows the components found in each RCT. Most studies were parallel design (*n*=196, 85%) or cluster RCTs (*n*=26, 11%). Fifty-two RCTs (23%) were characterized as pilot studies. They assessed interventions for a wide range of behaviors, mostly nutrition, physical activity, or both (*n*=82, 36%) and adherence to treatment (*n*=57, 25%). Trial populations were most commonly healthy individuals (*n*=97, 42%), individuals with overweight/obesity (*n*=25, 11%), or patients with cardiovascular disease/diabetes (*n*=35, 15%). Overall, 23 RCTs targeted children/adolescents exclusively (10%) and three targeted older adults (1%). In 49% (*n*=114) of RCTs, the mHealth interventions were delivered as standalone.

Half of all interventions were delivered by text message (*n*=118, 51%), 28% (*n*=65) were delivered primarily by smartphone app, and 24% (*n*=55) included a wearable device or other sensor. The 55 RCTs describing interventions with sensors/devices included nine interventions using a smart pill box or smart inhaler, one using digital medicine, and one using a wearable bite-tracking device. The remaining trials concerned wearable activity monitors, accelerometers, or pedometers. Only 22 RCT reports included information on sensor accuracy (representing 40% of RCTs assessing sensor-based interventions). Intervention frequency was mostly daily or multiple days per week (*n*=142, 62%), whereas for several interventions, the frequency varied (e.g., participants could access the app according to their needs; *n*=70, 30%).

Regarding intervention content, goal setting and self-monitoring were delivered together frequently (*n*=62, 30%), often in combination with feedback (*n*=52, 23%). Social features and reminders were described in 15% (*n*=34) and 51% (*n*=117) of RCT reports, respectively, whereas communication with a healthcare professional was described as possible in 12% (*n*=28). Almost half of the interventions (*n*=110, 48%) were not based on behavioral theory. Of the remaining interventions, most were based on the Social-Cognitive Theory (*n*=46, 20%), the Health Belief Model (*n*=13, 6%), or the Transtheoretical Model (*n*=12, 5%).

Nearly half of all RCTs assessed non-personalized, “one-size-fits-all” interventions (*n*=101, 44%). One quarter (*n*=53, 23%) assessed interventions with static personalization: The adaptation took place only once, at baseline (e.g., by incorporating the participant’s name or appointment date in the message or allowing participants to set their preferred time to receive notifications). The remaining 33% (*n*=77) assessed interventions with dynamic personalization,

with the adaptation occurring multiple times during the intervention (periodically or in real time). Among the dynamic interventions, most (*n*=57) were personalized on the basis of participants' behavior or outcomes (e.g., weight change). Six were described by the authors as "just-in-time" adaptive interventions (continuously adapted interventions that offer the right component at the right time) and ten relied on ecological momentary assessment (the real-time capture of behavior cues).

Comments on intervention scalability were identified in 72 (31%) RCTs, 24 of which described barriers. These concerned intervention costs, technical and privacy issues, limited adoption, and the long-term intervention impact (Appendix Table 3). An additional scalability barrier may be the need to adapt the assessed interventions for middle- and low-income settings, because 81% (*n*=186) of RCTs were conducted in high-income countries.

Only 8% (*n*=18) of all RCTs had a patient-important primary outcome, duration of ≥6 months, and ITT analysis (Figure 2). Overall, 130 RCTs (56%) had a patient-important primary outcome (Appendix Table 4 provides a list of outcomes classified as patient-important). Most primary outcomes were behavioral measures (e.g., number of steps, weekly red meat consumption) (*n*=131). In all, 27 RCTs did not define a primary outcome in the article, trial registry, or protocol. Regarding follow-up duration, more than one third of RCTs had a follow-up of <6 months (*n*=123, 53%), and in one third, the follow-up was very short (<3 months, *n*=67, 29%) (median follow-up: 4 months, IQR=2–6 months). Moreover, more than two thirds of RCTs (*n*=176, 76%) did not include any follow-up measurements after the end of the intervention to demonstrate sustainability of outcomes (e.g., maintenance of weight loss, relapse to unhealthy

behaviors). ITT analysis was reported in 34% (*n*=79) of RCTs. A breakdown of methodologic characteristics of RCTs by pilot status and commercial availability of the interventions showed that a smaller proportion of pilot RCTs had follow-up of ≥6 months compared with non-pilot RCTs (31% vs 51%), whereas a greater proportion of pilot RCTs reported participant feedback (63% vs 44%) (Appendix Table 5). A larger proportion of RCTs of commercially available interventions reported ITT analysis (49% vs 31% in RCTs of unavailable interventions). The presence of these components was similar between RCTs targeting different behaviors (Appendix Table 6).

In the overall sample (N=231), sample size calculation was reported in 52% (*n*=119). The median sample size included in the primary analysis was 119 (IQR=61–281). Finally, adverse events were reported in 18% (*n*=42) of RCTs. The ten events possibly related to the intervention were mostly unintended behavior changes (e.g., decreased condom use by adolescents in an intervention to prevent pregnancy)[29] or psychosocial effects (e.g., sustained depressive symptoms in an intervention for patients with bipolar disorder).[30]

Intervention Reporting

Seventy-two studies (31%) reported information elsewhere than in the primary RCT paper, including intervention pre-testing and development papers. The reporting completeness of selected mERA checklist items is presented in Figure 3 (Appendix Table 7 provides the complete checklist). Only two of the 16 items were frequently reported: Intervention delivery and intervention content were found in more than two thirds of the studies (*n*=224, 97% and *n*=219, 95%). Several key items were reported poorly, including user feedback (*n*=112, 48%), usability/content testing (which describes patient/public involvement in intervention

development; *n*=75, 32%), and fidelity (the receipt and use of the intervention by participants; e.g., number of app openings per week; *n*=124, 54%), and data security (*n*=29, 13%). RCTs of commercially unavailable interventions reported a mean of 6.9 (SD=2.5) items compared to 7.2 (SD=2.2) for RCTs assessing interventions with commercial software/sensors (Appendix Figure 1).

DISCUSSION

This methodologic review shows that recently published RCTs of mHealth behavior change interventions lacked long follow-up duration and ITT analysis, but more than half used patient-important primary outcomes. Reporting of information on technical aspects of the intervention is limited, which may render interventions difficult to replicate. More traditional intervention aspects (e.g., intervention contents) are better reported.

Regarding intervention contents, a small proportion of RCTs assessed interventions that were personalized dynamically by participants' psychosocial characteristics and environmental variables (e.g., beliefs, perceived barriers). Previous work has called for such personalized interventions,[31–34] based on behavioral models that are more dynamic than those used in the included RCTs.[35]

Achieving a sustainable impact on patient-important outcomes with acceptable, low-burden behavioral interventions is an important research gap.[36] The relationship between change in the continuous behavioral outcomes found in this sample (e.g., step count) and important health outcomes such as clinical events is complex.[14] In combination with the short follow-up duration,

the impact of mHealth interventions from a public health viewpoint may be unclear, particularly in light of the difficulty maintaining long-term adherence to behavioral interventions and sustaining behavior change.[14,37] Importantly, usability/content testing, interoperability, participant feedback, and fidelity were often reported poorly, thereby offering limited information on mHealth acceptability and potential uptake. These issues characterize the field overall, because methodology and reporting did not differ between studies targeting different interventions with different commercial availability or target behaviors.

The development of mHealth interventions has several stages. At earlier stages, RCTs may be used to assess intervention feasibility and optimization, rather than its impact. Methodologic components may vary to match the research question. However, as with all RCTs assessing non-pharmacologic interventions, the assessment of mHealth interventions should follow good practice including patient-centric methodology and complete reporting.

These findings agree with previous methodologic work on studies assessing commercially available apps.[38] As compared with a previous review assessing the reporting of 35 studies with the mERA, better reporting was found for technology platforms and worse reporting for interoperability, possibly because the previous review assessed several study designs in the domain of reproductive health, whereas this review included RCTs in any domain.[39] The prevalence of ITT analysis and patient-important primary outcomes is similar to that reported in other reviews.[40,41] This review found a higher proportion of patient-important primary outcomes than previous reviews of studies of diabetes and critically ill patients, which could be attributed to differences in outcomes classification or time trends.[28,42]

This is the first methodologic review to examine RCTs assessing mHealth interventions and summarize reported scalability barriers and the first to use the final version of the mERA. This review examined a large number of interventions for diverse behaviors, published recently in journals in the upper half of the Journal Impact Factor and representing widely cited, influential publications, and all primary outcomes were assessed. An important effort was made to source information from previous work relevant to the intervention.

Limitations

The results are not generalizable to publications in journals with a lower Journal Impact Factor. To thoroughly examine the recent literature, the review was restricted to papers published between 2014 and 2018. As the mERA checklist was published in 2016, it would have had limited impact on the reporting of studies in this review. Because this review summarizes papers published up to 2018, the findings may not be generalizable to more recent publications that differ in methods and reporting completeness. However, RCT methods and reporting tend to evolve slowly over time in the absence of regulatory changes (i.e., mandatory use of reporting guidelines).[43] The use of broad search terms (e.g., *behav**) may have led to missing some studies.

There are no standards for what constitutes a patient-important behavioral outcome. For the purpose of this paper, outcomes were classified using a combination of literature definitions and the opinion of a patient panel for outcomes not classified as patient-important based on prior literature. Patients' beliefs on what constitutes a patient-important outcome are diverse and may be affected by their interaction with their own physicians. The inclusion of different patients or

the use of different definitions (e.g., agreement between all patients) may have led to a different proportion of outcomes classified as patient-important.

Recommendations for Future Research

This review provides a broad assessment of the mHealth behavior change literature. Some of the methodologic components examined here may not be achievable for all studies because of practical barriers (i.e., longer duration may be prohibited by budget constraints), whereas others are modifiable.

Several steps could be taken to improve RCTs evaluating mHealth interventions. First, these RCTs should be held to the same standards as all RCTs of non-pharmacologic interventions and follow good practice (e.g., pre-specifying outcomes, ITT analysis).[44,45] Because mHealth interventions are complex with many aspects, complete reporting in the space of the main paper may be challenging. Important information should be reported in the main paper, including intervention fidelity and participant feedback. Protocols and online supplements can be used for complete reporting. As several experts have proposed, future mHealth research should focus on assessing the effectiveness of highly dynamic, personalized interventions and resolving the identified barriers to intervention scale-up.[6,9,46]

CONCLUSIONS

This methodologic review found that mHealth behavioral change interventions assessed in RCT articles recently published in high–Impact Factor journals lack long follow-up and have gaps in intervention reporting, which prohibits the assessment of their long-term clinical impact and

limits their replicability. There is potential to improve the field through consensus for more parsimonious methodologic choices and adoption of good reporting practices.

ACKNOWLEDGMENTS

The authors thank Laura Smales for editing the manuscript, Elise Diard (Centre de Recherche Epidémiologie et Statistique, INSERM U1153, French Cochrane Center) for the design of Figure 2, and the patients who participated in the outcomes classification.

Authors' contributions: TO, AV, and PR conceived and designed the study; TO, AV, ER, and CR performed the data collection and VTT contributed to data interpretation; TO wrote the first draft of the paper; and AV, ER, CR, VTT, and PR critically revised the manuscript for important intellectual content. All authors approved of the final version of this manuscript to be published.

No financial disclosures were reported by the authors of this paper.

REFERENCES

1. van Heerden A, Tomlinson M, Swartz L. Point of care in your pocket: a research agenda for the field of m-health. *Bull World Health Organ*. 2012;90(5):393–394. https://doi.org/10.2471/blt.11.099788.
2. Mbuagbaw L, van der Kop ML, Lester RT, et al. Mobile phone text messages for improving adherence to antiretroviral therapy (ART): an individual patient data meta-analysis of randomised trials. *BMJ Open*. 2013;3(12):e003950. https://doi.org/10.1136/bmjopen-2013-003950.
3. Free C, Phillips G, Galli L, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. *PLoS Med*. 2013;10(1):e1001362. https://doi.org/10.1371/journal.pmed.1001362.
4. Yanovski SZ. Weight management in adults with obesity: what is a primary care clinician to do? *JAMA*. 2018;320(11):1111–1113. https://doi.org/10.1001/jama.2018.11031.
5. Bacigalupo R, Cudd P, Littlewood C, Bissell P, Hawley M, Buckley Woods H. Interventions employing mobile technology for overweight and obesity: an early systematic review of randomized controlled trials. *Obes Rev*. 2013;14(4):279–291. https://doi.org/10.1111/obr.12006.
6. Collins FS, Riley WT. NIH's transformative opportunities for the behavioral and social sciences. *Sci Transl Med*. 2016;8(366):366ed14. https://doi.org/10.1126/scitranslmed.aai9374.

7. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. *PLoS One*. 2017;12(2):e0170581. https://doi.org/10.1371/journal.pone.0170581.
8. WHO. *mHealth: New horizons for health through mobile technologies: second global survey on ehealth. Global Observatory for eHealth series—Volume 3*. Geneva, Switzerland: WHO Press, WHO; 2011.
9. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med*. 2018;52(6):446–462. https://doi.org/10.1007/s12160-016-9830-8.
10. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. *J Med Internet Res*. 2013;15(5):e95. https://doi.org/10.2196/jmir.2430.
11. Statista. Number of Apps Available in Leading App Stores As of March 2017.
12. McCarroll R, Eyles H, Ni Mhurchu C. Effectiveness of mobile health (mHealth) interventions for promoting healthy eating in adults: a systematic review. *Prev Med*. 2017;105:156–168. https://doi.org/10.1016/j.ypmed.2017.08.022.
13. Mbuagbaw L, Mursleen S, Lytvyn L, Smieja M, Dolovich L, Thabane L. Mobile phone text messaging interventions for HIV and other chronic diseases: an overview of systematic reviews and framework for evidence transfer. *BMC Health Serv Res*. 2015;15:33. https://doi.org/10.1186/s12913-014-0654-6.
14. McNellis RJ, Ory MG, Lin JS, O'Connor EA. Standards of evidence for behavioral counseling recommendations. *Am J Prev Med*. 2015;49(3 suppl 2):S150–S157. https://doi.org/10.1016/j.amepre.2015.06.002.

15. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*. 2014;348:g1687. https://doi.org/10.1136/bmj.g1687.

16. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*. 2014;383(9913):267–276. https://doi.org/10.1016/s0140-6736(13)62228-x.

17. Michie S, Abraham C, Eccles MP, Francis JJ, Hardeman W, Johnston M. Strengthening evaluation and implementation by specifying components of behaviour change interventions: a study protocol. *Implement Sci*. 2011;6:10. https://doi.org/10.1186/1748-5908-6-10.

18. Michie S, Johnston M. Theories and techniques of behaviour change: developing a cumulative science of behaviour change. *Health Psychol Rev*. 2012;6(1):1–6. https://doi.org/10.1080/17437199.2012.654964.

19. Agarwal S, LeFevre AE, Lee J, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ*. 2016;352:i1174. https://doi.org/10.1136/bmj.i1174.

20. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535. https://doi.org/10.1136/bmj.b2535.

21. Clarivate Analytics. 2015 Journal Citation Reports®. Published 2017.

22. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol*. 2009;28(6):690–701. https://doi.org/10.1037/a0016136.

23. Epton T, Currie S, Armitage CJ. Unique effects of setting goals on behavior change: systematic review and meta-analysis. *J Consult Clin Psychol*. 2017;85(12):1182–1198. https://doi.org/10.1037/ccp0000260.

24. Thakkar J, Kurup R, Laba TL, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. *JAMA Intern Med*. 2016;176(3):340–349. https://doi.org/10.1001/jamainternmed.2015.7667.

25. Milat AJ, King L, Bauman AE, Redman S. The concept of scalability: increasing the scale and potential adoption of health promotion interventions into policy and practice. *Health Promot Int*. 2012;28(3):285–298. https://doi.org/10.1093/heapro/dar097.

26. The World Bank. The world by income - 2017. http://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html. Accessed July 10, 2019.

27. Yordanov Y, Dechartres A, Ravaud P. Patient-important outcomes in systematic reviews: poor quality of evidence. *PLoS One*. 2018;13(4):e0195460 https://doi.org/10.1371/journal.pone.0195460.

28. Gandhi GY, Murad MH, Fujiyoshi A, et al. Patient-important outcomes in registered diabetes trials. *JAMA*. 2008;299(21):2543–2549. https://doi.org/10.1001/jama.299.21.2543.

29. Rokicki S, Cohen J, Salomon JA, Fink G. Impact of a text-messaging program on adolescent reproductive health: a cluster–randomized trial in Ghana. *Am J Public Health*. 2017;107(2):298–305. https://doi.org/10.2105/ajph.2016.303562.

30. Faurholt-Jepsen M, Frost M, Ritz C, et al. Daily electronic self-monitoring in bipolar disorder using smartphones–the MONARCA I trial: a randomized, placebo-controlled,

single-blind, parallel group trial. *Psychol Med*. 2015;45(13):2691–2704. https://doi.org/10.1017/s0033291715000410.

31. Ma J, Rosas LG, Lv N. Precision lifestyle medicine: a new frontier in the science of behavior change and population health. *Am J Prev Med*. 2016;50(3):395–397. https://doi.org/10.1016/j.amepre.2015.09.035.
32. Dunton GF. Sustaining health-protective behaviors such as physical activity and healthy eating. *JAMA*. 2018;320(7):639–640. https://doi.org/10.1001/jama.2018.6621.
33. Ellis DA, Piwek L. Failing to encourage physical activity with wearable technology: what next? *J R Soc Med*. 2018;111(9):310–313. https://doi.org/10.1177/0141076818788856.
34. Tong HL, Laranjo L. The use of social features in mobile health interventions to promote physical activity: a systematic review. *NPJ Digit Med*. 2018;1:43. https://doi.org/10.1038/s41746-018-0051-3.
35. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl Behav Med*. 2011;1(1):53–71. https://doi.org/10.1007/s13142-011-0021-7.
36. Kurth AE, Miller TL, Woo M, Davidson KW. Understanding research gaps and priorities for improving behavioral counseling interventions: lessons learned from the U.S. Preventive Services Task Force. *Am J Prev Med*. 2015;49(3 suppl 2):S158–S165. https://doi.org/10.1016/j.amepre.2015.06.007.
37. Pfaeffli Dale L, Dobson R, Whittaker R, Maddison R. The effectiveness of mobile-health behaviour change interventions for cardiovascular disease self-management: a systematic review. *Eur J Prev Cardiol*. 2016;23(8):801–817. https://doi.org/10.1177/2047487315613462.

38. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. *NPJ Digit Med*. 2018;1(1):12. https://doi.org/10.1038/s41746-018-0021-9.

39. L'Engle KL, Mangone ER, Parcesepe AM, Agarwal S, Ippoliti NB. Mobile phone interventions for adolescent sexual and reproductive health: a systematic review. *Pediatrics*. 2016;138(3):e20160884. https://doi.org/10.1542/peds.2016-0884.

40. Bell ML, Fiero M, Horton NJ, Hsu C-H. Handling missing data in RCTs: a review of the top medical journals. *BMC Med Res Methodol*. 2014;14:118. https://doi.org/10.1186/1471-2288-14-118.

41. Adie S, Harris IA, Naylor JM, Mittal R. Are outcomes reported in surgical randomized trials patient-important? A systematic review and meta-analysis. *Can J Surg*. 2017;60(2):86–93. https://doi.org/10.1503/cjs.010616.

42. Gaudry S, Messika J, Ricard J-D, et al. Patient-important outcomes in randomized controlled trials in critically ill patients: a systematic review. *Ann Intensive Care*. 2017;7:28. https://doi.org/10.1186/s13613-017-0243-z.

43. Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. *BMJ*. 2012;344:e4178. https://doi.org/10.1136/bmj.e4178.

44. Lewis J, Machin D. Intention to treat--who should use ITT? *Br J Cancer*. 1993;68(4):647–650. https://doi.org/10.1038/bjc.1993.402.

45. Smyth R, Kirkham J, Jacoby A, Altman D, Gamble C, Williamson P. Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. *BMJ*. 2011;342:c7153. https://doi.org/10.1136/bmj.c7153.

46. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. *JAMA*. 2015;313(5):459–460. https://doi.org/10.1001/jama.2014.14781.

LIST OF FIGURES

and light-grey interventions for which information was reported in an additional source cited in the main publication. Technical aspects of the intervention (e.g., Technology platform, Data security) were frequently poorly reported as compared with more “traditional” aspects of intervention reporting (i.e., Intervention contents and Intervention delivery). Information on patient/participant engagement in the development and pre-testing of the intervention (Usablity/content testing) and training offered to participants for the use of the mHealth intervention (Adoption inputs/program entry) was reported in less than half of the sample.

Table 1. Key Sample Characteristics of 231 RCTs Assessing mHealth Behavior Change Interventions[a]

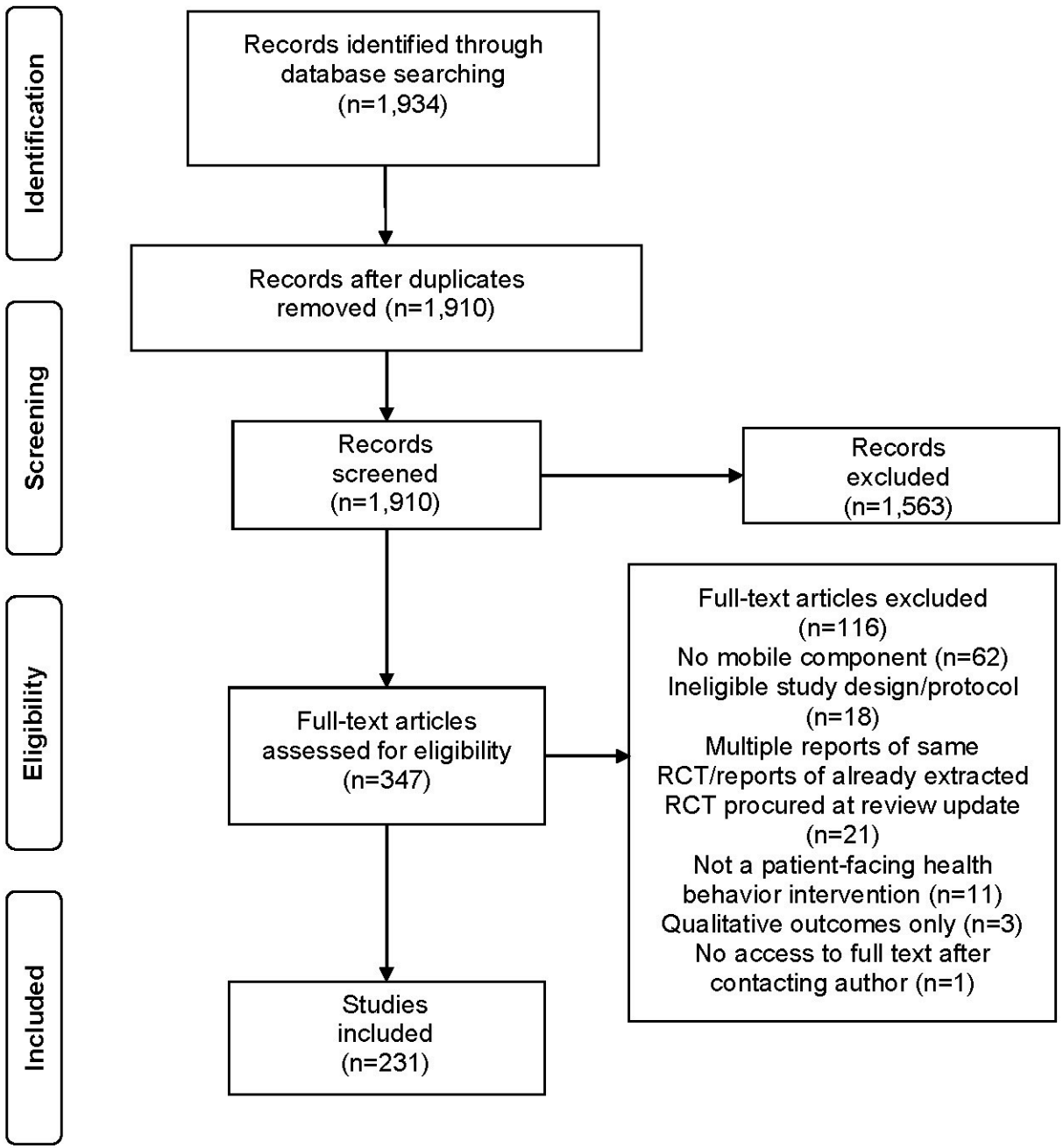
Sample characteristics	Studies (n=231)
Sample size included in primary analysis, median (IQR)[b]	119 (61–281)
Patient condition, n (%)	
Healthy individuals	97 (42)
Overweight/Obese	25 (11)
HIV infection	19 (8)
Cardiovascular disease/Diabetes	35 (15)
Pregnancy	14 (6)
Other[c]	41 (18)
Target behavior, n (%)	
Nutrition and/or physical activity	82 (36)
Adherence to treatment	57 (25)
Self-management	29 (13)
Smoking	19 (8)
Sexual/reproductive	9 (4)
Vaccination	7 (3)
Sun protection	6 (3)
Other[d]	22 (10)
Intervention technology, n (%)	
Text messaging only	118 (51)
Smartphone app	65 (28)
Wearable device plus text messaging or app	39(17)
Other	9 (4)
Use of sensor technology, n (%)	55 (24)
Use of behavioral techniques, n (%)	
Self-monitoring	124 (54)
Goal setting	82 (36)
Feedback	95 (41)
Intervention frequency, n (%)	
Weekly to daily	142 (62)
Monthly or less	5 (1)
One-time intervention	12 (5)
Variable frequency	70 (30)
Unreported	2 (1)
Follow-up duration, n (%)	
<3 months	68 (29)
3 to 5 months	56 (24)
≥6 months	107 (46)

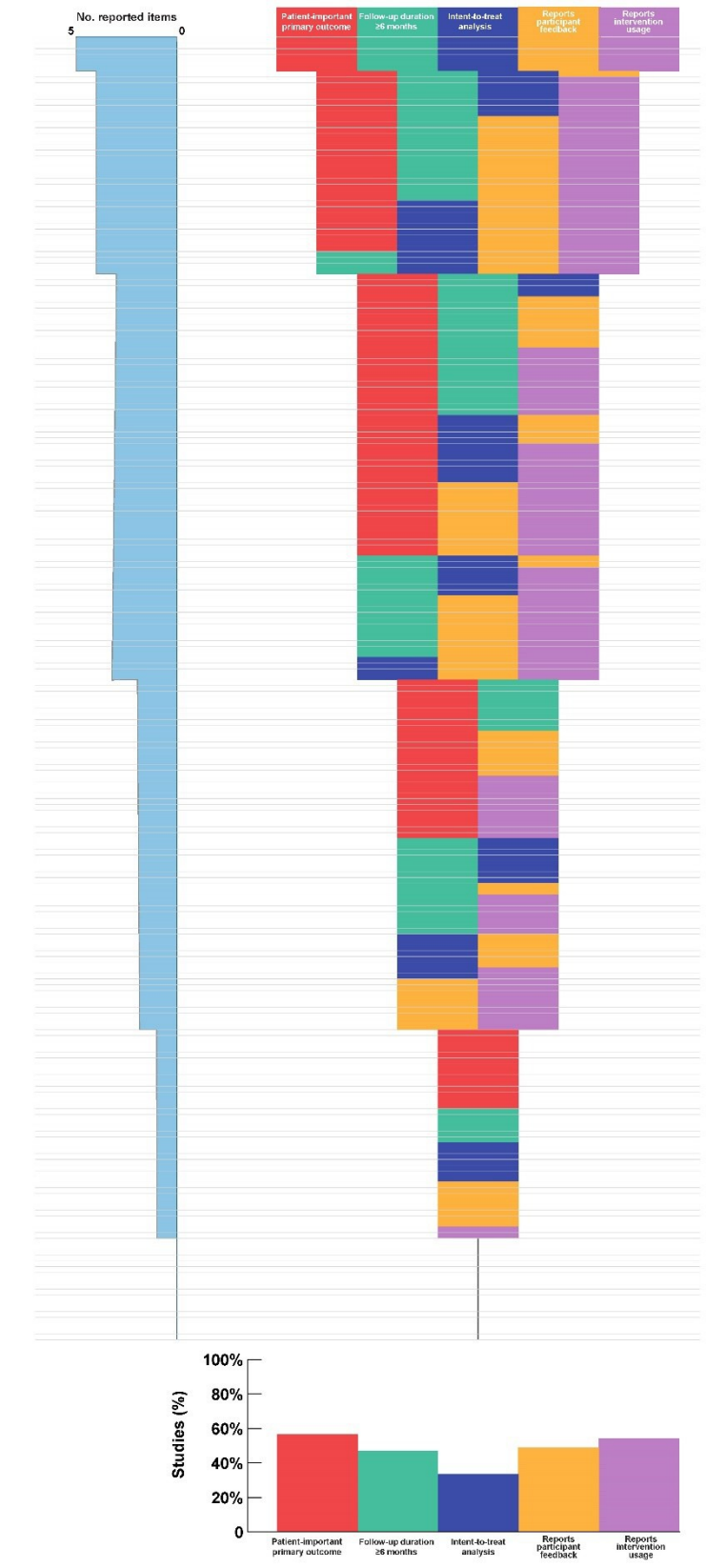
[a]Percentages may not add up to 100 because of rounding.

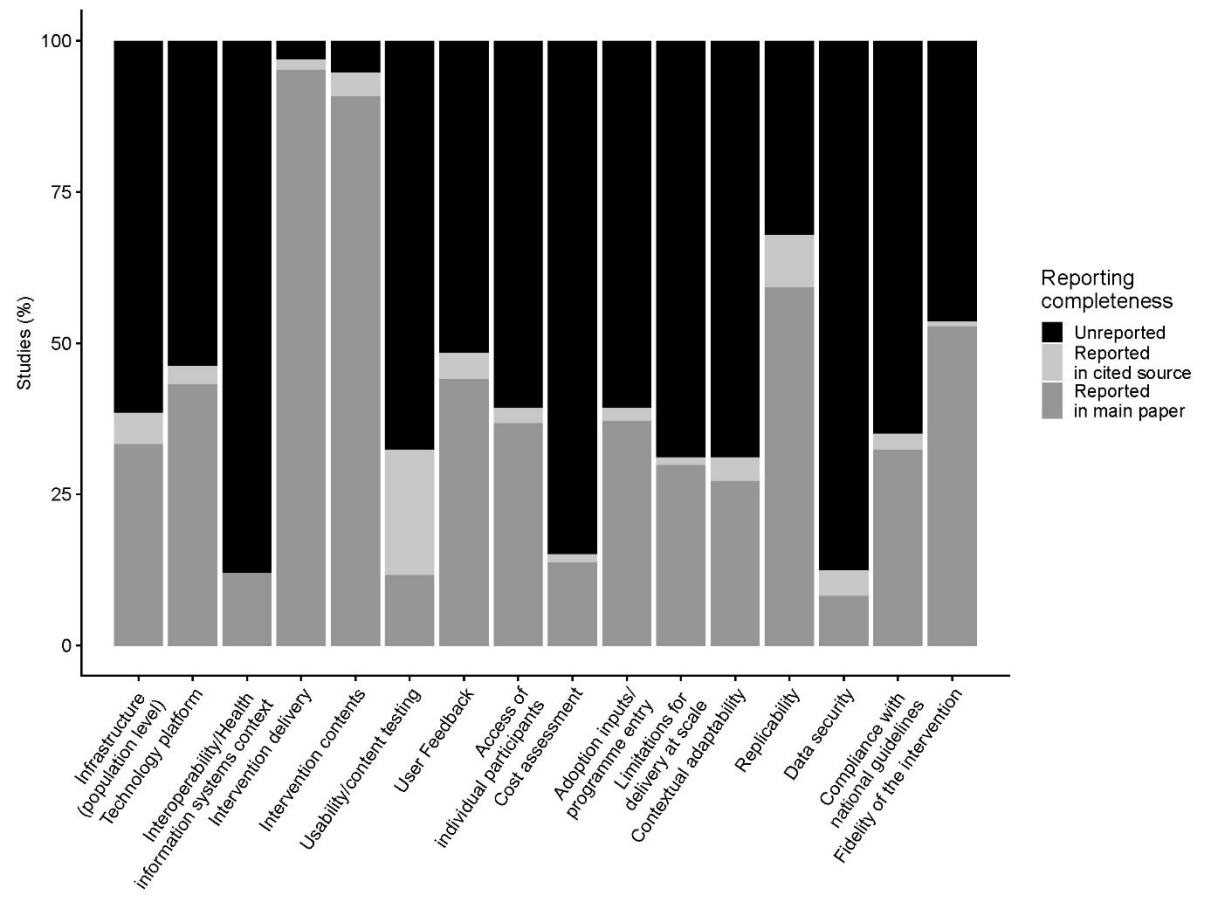
[b]Excluding 26 cluster trials.

[c]Other conditions include bladder pain syndrome, lower back pain, transplant recipients, post-circumcision care, esophageal cancer, at-home misoprostol abortion, menopause, orthodontic treatment, being recently discharged from the emergency department, and taking multiple medications due to multimorbidity.

[d]Other behaviors include rabies prevention, dengue prevention, promotion of breastfeeding, parental coaching, cravings management, rehabilitation after cardiac event, self-injurious behavior.


Identification
Screening
Eligibility
Included
Records identified through database searching (n=1,934)
Records after duplicates removed (n=1,910)
Records screened (n=1,910)
Records excluded (n=1,563)
Full-text articles assessed for eligibility (n=347)
Full-text articles excluded (n=116)
No mobile component (n=62)
Ineligible study design/protocol (n=18)
Multiple reports of same RCT/reports of already extracted RCT procured at review update (n=21)
Not a patient-facing health behavior intervention (n=11)
Qualitative outcomes only (n=3)
No access to full text after contacting author (n=1)
Studies included (n=231)


No. reported items
5
0
Patient-important primary outcome
Follow-up duration ≥6 months
Intent-to-treat analysis
Reports participant feedback
Reports intervention usage
Studies (%)
100%
80%
60%
40%
20%
0
Patient-important primary outcome
Follow-up duration ≥6 months
Intent-to-treat analysis
Reports participant feedback
Reports intervention usage


100
75
50
25
0
Studies (%)
Reporting completeness
Unreported
Reported in cited source
Reported in main paper
Infrastructure (population level)
Technology platform
Interoperability/Health information systems context
Intervention delivery
Intervention contents
Usability/content testing
User Feedback
Access of individual participants
Cost assessment
Adoption inputs/ programme entry
Limitations for delivery at scale
Contextual adaptability
Replicability
Data security
Compliance with national guidelines
Fidelity of the intervention