

REVIEW ARTICLE OPEN



Development features and study characteristics of mobile health apps in the management of chronic conditions: a systematic review of randomised trials

Maria Cucciniello^{1,2}, Francesco Petracca², Oriana Ciani^{2,3} and Rosanna Tarricone^{2,4}✉

COVID-19 pandemic challenges have accelerated the reliance on digital health fuelling the expanded incorporation of mobile apps into healthcare services, particularly for the management of long-term conditions such as chronic diseases (CDs). However, the impact of health apps on outcomes for CD remains unclear, potentially owing to both the poor adoption of formal development standards in the design process and the methodological quality of studies. A systematic search of randomised trials was performed on Medline, ScienceDirect, the Cochrane Library and Scopus to provide a comprehensive outlook and review the impact of health apps on CD. We identified 69 studies on diabetes ($n = 29$), cardiovascular diseases ($n = 13$), chronic respiratory diseases ($n = 13$), cancer ($n = 10$) or their combinations ($n = 4$). The apps rarely adopted developmental factors in the design stage, with only around one-third of studies reporting user or healthcare professional engagement. Apps differed significantly in content, with a median of eight behaviour change techniques adopted, most frequently pertaining to the 'Feedback and monitoring' (91%) and 'Shaping knowledge' (72%) categories. As for the study methodologies, all studies adopted a traditional randomised control trial (RCT) design, with relatively short follow-ups and limited sample sizes. Findings were not significant for the majority of studies across all CD, with most RCTs revealing a high risk of bias. To support the adoption of apps for CD management, this review reinforces the need for more robust development and appropriate study characteristics to sustain evidence generation and elucidate whether study results reflect the true benefits of apps or a biased estimate due to unsuitable designs.

npj Digital Medicine (2021)4:144; <https://doi.org/10.1038/s41746-021-00517-1>

INTRODUCTION

Smartphone users in the world have steadily increased, surpassing the 3.5 billion mark in 2020 with a further expected growth of several hundred million in the coming years¹, fuelling the expanding interest in mHealth apps². The coronavirus disease 2019 (COVID-19) pandemic has contributed to further accelerate the reliance on digital health³⁻⁵.

Therefore, apps are being increasingly incorporated into healthcare (HC) services owing to their portability, instantaneous access and direct communication, inspiring new models of remote HC delivery and cost-effective solutions for chronic diseases, whose long-term nature and need for continuous monitoring can be positively impacted^{2,6,7}. mHealth apps may be particularly effective in self-management, one of the components of the eHealth Enhanced Chronic Care Model⁸, intended as the 'individual's ability to manage the symptoms, treatment, physical and psychosocial consequences and lifestyle changes inherent in living with a chronic disease'⁹. Apps could as well improve patient empowerment, the process of gaining knowledge of one's health and ability and motivation to influence it^{10,11}.

Despite the ever-growing interest and the increasing number of apps on common platforms, their impact on outcomes for chronic diseases remains unclear. Several systematic reviews and meta-analyses reported mixed impact of apps on clinical outcomes, self-management and behaviour change¹²⁻²¹,

questioning the ability of science to keep up with the continuous technological advances²².

The inconclusive evidence base of apps could be attributed both to the poor adoption of formal development standards^{23,24}, notwithstanding the reported benefits of adopting participatory approaches and behaviour change theories^{25,26}, and to the methodological quality of study designs²⁷.

Previous reviews had summarised the impact of apps on health outcomes across different conditions, concluding that high-quality research is needed to transform the promise of mHealth technology into improved HC delivery and outcomes^{7,28}. More recently, Iribarren et al. comprehensively reviewed app-delivered behaviour change interventions targeting health outcomes and thoroughly analysed the corresponding app features²⁹. However, no study on chronic diseases has comprehensively looked at both the developmental features of apps, namely the strategies and considerations adopted throughout the development stage of mobile apps, thereby including the behavioural change features adopted, and the characteristics of the evaluation study in determining the effect on a wide array of domains.

We, therefore, conducted a systematic review of randomised studies on smartphone apps for the management of high prevalence diseases (diabetes, cardiovascular diseases, chronic respiratory diseases and cancer) to fill this gap.

¹University of Edinburgh Business School - Centre for Service Excellence (CenSE), Scotland, UK. ²Centre for Research in Health and Social Care Management (CERGAS), Government, Health and Non Profit Division, SDA Bocconi, Via Sarfatti 10, 20136 Milan, Italy. ³Institute of Health Research, University of Exeter Medical School, South Cloisters, St. Luke's Campus, Heavitree Road, EX1 2LU Exeter, UK. ⁴Department of Social and Political Sciences, Bocconi University, Via Roentgen 1, 20136 Milan, Italy. ✉email: rosanna.tarricone@unibocconi.it

RESULTS

Study selection

Following the removal of duplicates, the addition of 16 publications cited in reference lists, and title and abstract screening, 175 records were selected for full-text examination. The Cohen's kappa coefficient expressing the inter-rater agreement was 0.84 during the abstract screening.

Overall, 74 papers based on 69 different trials were included in the review. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart of the study selection process is presented as Fig. 1, including the reported reasons for exclusion.

Characteristics of included studies and participants

The studies were published between 2008 and 2019 in 24 countries, with the United States of America ($n = 13$)^{30–43} and China ($n = 7$)^{44–50} ranked first and second, respectively. Only one study covered multiple countries, involving diabetes clinics in Italy, England and Spain⁵¹.

The chronic disease most commonly addressed was diabetes ($n = 29$, 42.0%), with 16 studies sampling type 2 diabetes (T2DM) population in 19 articles^{30–33,42,48,49,52–63}, 10 focusing on type 1 diabetes (T1DM)^{51,64–72} and three enrolling both T1DM and T2DM patients in four articles^{45,73–75}. Thirteen studies (18.8%) targeted cardiovascular diseases, with five focusing on heart failure^{35,36,76–78} and the remaining on cardiovascular diseases (CVDs)^{34,38,39,79–84}. An equal number addressed chronic respiratory diseases ($n = 13$, 18.8%), with seven studies tackling asthma^{41,85–90} and six focusing

on chronic obstructive pulmonary disease (COPD)^{91–96}. Lastly, ten studies focused on cancer ($n = 10$, 14.5%)^{40,43,44,46,47,50,97–100}. One study separately addressed T2DM and CVDs (either ischaemic heart disease or heart failure)¹⁰¹, while two others included both uncontrolled hypertension and diabetes^{37,102} and one focused on T2DM and/or hypertension¹⁰³. Their sample sizes ranged from 18³⁵ to 519¹⁰¹ and the median size was 94, with 23 studies (33.3%) totalling between 50 and 100 participants and only six studies with more than 200 individuals^{33,39,59,88,90,101}.

The vast majority were either individually randomised parallel-group trials ($n = 52$, 75.4% of total studies) or pilot RCTs ($n = 12$, 17.4%), while the remaining five had a cluster-randomised ($n = 3$, 4.3%)^{31,37,90} or an individual cross-over design ($n = 2$, 2.9%)^{69,72}.

Out of the 69 studies, 60 were two-arm, while eight (11.6%) had three and one (1.4%) had four groups³¹. Among those with multiple intervention groups, several studies, along with the standard app version, included a different arm with an enhanced intervention, on top of app use, with the possibility of using either the teleconsultation option⁶⁴, Health Counselling intervention⁵³, classroom-based programmes³³, decision support³¹, physician review¹⁰⁰ or additional app features⁸⁴. Frias et al., instead, had two intervention arms using the same app, but with different follow-ups³⁷.

In the control group, most participants received care as usual. However, some studies included lighter technological features for the control group: Johnston et al. provided the control with a simplified drug adherence e-diary installed on phones⁸³, while other studies either included basic versions of the app⁶⁸ or administered education/information programmes over mobile

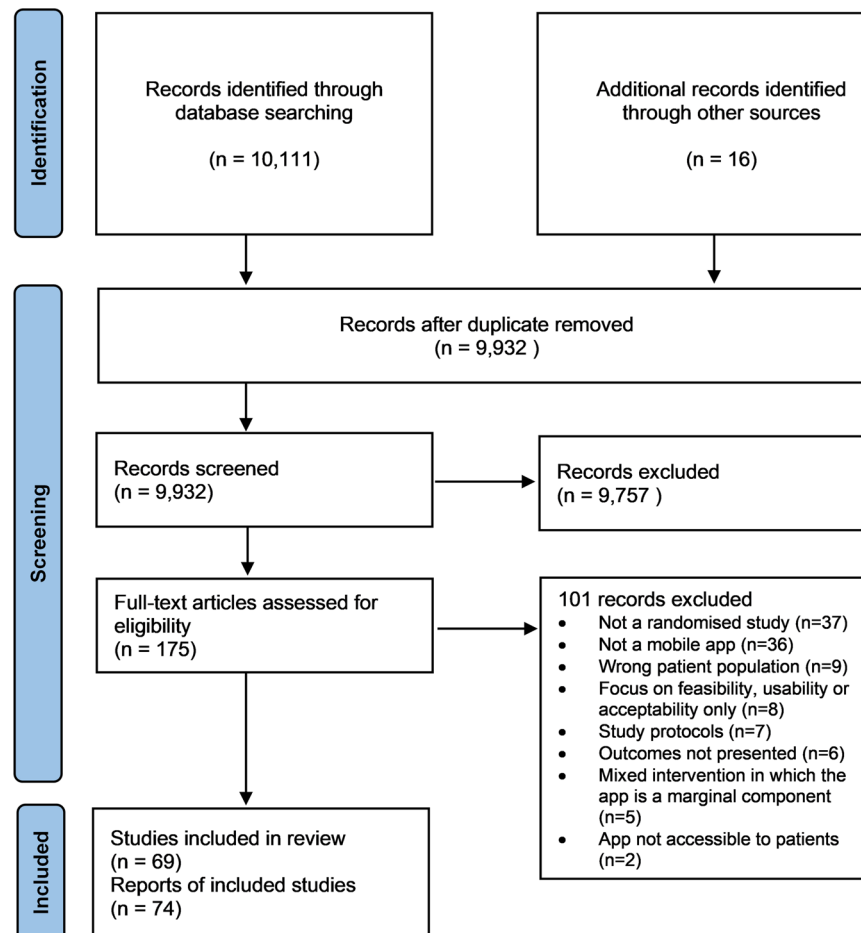


Fig. 1 Study selection flow diagram. Modified Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the article selection process.

phones^{35,43}. Ryan et al.⁸⁸ provided the control with enhanced clinical care to exclude attribution of potential benefits to intense interventions in the app group. Franc et al.'s study⁶⁰, besides the Diabeo app arm, included a simpler telemonitoring system via an interactive voice response system, while Kwon et al.⁹³ adopted two different exercise regimens.

The follow-ups varied: for 26 studies, they were between 4 and 6 months (37.7%); for 24 studies (34.8%), 1–3 months; in 13 cases (18.8%), >6 months and for the remaining six (8.7%) only 4 weeks.

The majority of the studies ($n = 44$, 63.8%) included internet literacy as an explicit eligibility criterion, as either generic familiarity with mobiles or, more explicitly, ownership of smartphones with specific operating systems. A notable exception is Baron et al.'s study⁷³ excluding those with previous mobile telehealth services experience.

Recruitment procedures were mostly traditional, with offline methods for 64 studies (92.7%). Only Morawski et al.³⁹ adopted a completely online recruitment strategy, through patient communities, social media, mobile apps and advertisement, while four studies pursued both offline and online strategies, through online advertisement and community forums^{40,61,65,82}.

Finally, there were two major approaches regarding the mobile phones in the trial: 36 studies (52.2%) provided participants with study devices after randomisation, while 31 (44.9%) adopted a bring your own device approach, downloading the app on the participant's smartphone. Only two studies adopted a mixed strategy, opting to install the app on a loaned phone if the participant was incompatible with the software version^{35,87}. Key study characteristics are summarised in Table 1, while a per-study overview of selected study characteristics can be found in Supplementary Table 1.

App design and development considerations. Five development factors were analysed comparatively across all studies. Based on the study conducted by Adu and colleagues²⁴, we considered the following factors: (i) supporting behavioural theory, (ii) user involvement in the design, (iii) healthcare professional (HCP) involvement in the design, (iv) data security and privacy considerations, (v) pilot testing, which includes all forms of interaction with target users to enhance the usability, acceptability and reliability of the app before its complete testing (full details in Supplementary Table 2).

Regarding health behavioural theories, only a minority of the studies ($n = 11$, 15.9%) reported that the interventions were based on theories and models of behaviour change and were beneficial in developing apps^{35,42,43,47,52,57–59,61,73,90}.

Concerning user involvement in app design, 21 studies (30.4%) explicitly mentioned patient engagement strategies, incorporating patient inputs in app design. While most simply cite an interactive approach in close collaboration with intended users, a few specify the process in greater detail. Ryan et al. based their formative work on qualitative interviews of ten asthma patients and two research staff to identify technological adjustments for improved solutions⁸⁸, while other publications included user requirements through survey results^{93,99}, iterative piloting⁹⁵ or a consecutive series of user studies during a 26-month development phase⁶¹. A similar number also engaged HCPs in the app design phase ($n = 18$, 26.1%). While many studies adopted engagement strategies in design involving both users and HCPs, there were a few exceptions. Quinn et al.'s studies included only endocrinologists and Certified Diabetes Educators in design^{30,31}, while Yang et al. adopted a Delphi Method with two consultation rounds involving 30 clinical and nursing experts⁵⁰ and Greer et al. reported extensive user testing by a team of clinical researchers⁴³. On the contrary, the Few Touch app was developed using focus groups, interviews, feasibility testing and questionnaires over a 3-year period with T2DM patients only⁵³.

Data security and privacy information were documented by 33

Table 1. Key characteristics of included studies ($n = 69$).

Study characteristics	n (%)
Disease area	
Diabetes	29 (42.0%)
Cardiovascular diseases	13 (18.8%)
Respiratory diseases	13 (18.8%)
Cancer	10 (14.5%)
Multiple	4 (5.8%)
Publication year^a	
2008–2013	13 (17.6%)
2014–2016	25 (33.8%)
2017–2019	36 (48.6%)
Country of study	
United States of America	13 (18.8%)
China	7 (10.1%)
Australia	5 (7.2%)
Canada	5 (7.2%)
The Netherlands	5 (7.2%)
United Kingdom	5 (7.2%)
Finland	3 (4.3%)
South Korea	3 (4.3%)
Switzerland	3 (4.3%)
Others	20 (29.0%)
Design	
RCT	52 (75.4%)
Pilot RCT	12 (17.4%)
Cluster RCT	3 (4.3%)
Cross-over RCT	2 (2.9%)
Study arms	
Two	60 (87.0%)
Three	8 (11.6%)
Four	1 (1.4%)
Sample size	
Median (range)	94 (18–519)
<50	14 (20.3%)
51–100	23 (33.3%)
101–150	15 (21.7%)
151–200	11 (15.9%)
>200	6 (8.7%)
Follow-up interval	
Up to 1 month	6 (8.7%)
1–3 months	24 (34.8%)
4–6 months	26 (37.7%)
More than 6 months	13 (18.8%)
Recruitment strategy	
Offline	64 (92.8%)
Online	1 (1.4%)
Mixed	4 (5.8%)
Internet literacy as an eligibility criterion	
Yes	44 (63.8%)
No	25 (36.2%)
Device strategy	
Study device approach	36 (52.2%)
Bring your own device approach	31 (44.9%)
Mixed approach	2 (2.9%)

^aThe total number of included publications is equal to 74.

of the studies (47.8%), mostly pertaining to secure data transfer and storage from app to study servers, the minimisation of medical/personal information stored on the mobile and obtained through the app, and the password protection and encryption of core data. A couple of studies mentioned compliance with the

Health Insurance Portability and Accountability Act^{30,31,38}, while Wayne et al. mentioned that the Connected Wellness Platform to collect data exceeded Canadian privacy standards for software carrying health information⁵⁵. Only one study performed a comprehensive risk analysis before the trial to ensure data privacy and security⁵³.

The incorporation of pilot testing into app development was cited by 28 of the studies (40.6%), with eight additional trials (11.6%) characterised as pilot RCTs themselves. Of those 28 studies, the great majority reported observational single-arm studies instrumental in assessing study feasibility and usability. Some studies instead documented multiphase pilot testing: Vorrink et al. reported three rounds of pilot testing initially starting with healthy volunteers and later including COPD patients⁹⁵, while Kearney et al. developed and validated the advanced symptom management system through a two-arm pilot study in Scotland with ten patients receiving chemotherapy and a feasibility study to evaluate the system's acceptability on a convenience sample of 18 patients and nine HCPs⁹⁸. Boer et al.'s Adaptive Computerised COPD Exacerbation Self-management Support tool was validated through a 3-month prospective observational study, but was further optimised during the trial with a 2-week run-in period to familiarise participants with the technology⁹⁶. Finally, a few studies that did not test apps specifically designed for research purposes, but rather selected commercially available ones adopted different forms of piloting, by identifying software with the highest usability ratings by users or researchers adopting the Mobile App Rating Scale^{39,40,84}. The developmental factors adopted are summarised in Table 2.

Features of the mHealth interventions. The interventions differed significantly in content, level of HCP involvement and degree of automation in decision support (Supplementary Table 3).

Regarding professional participation, 26 studies (37.7%) excluded any in excess of that guaranteed to control patients during routine follow-up visits. Among these, seven also embedded a low technology-automation level in decision support: three studies aimed at sharing information and increasing disease awareness through the app^{58,82,97}, while others either worked as digital diaries allowing parameter recording under clinical supervision without direct intervention^{67,69}, simulated e-therapy through multimedia contents⁴³ or delivered mindfulness training⁴⁰. The remaining 19 studies adopted more intensive automated technology: four recorded participants' medication adherence through reminders^{39,63,84} or an artificial intelligence platform with dosing instructions³⁸; four others fostered improved physical activity automatically selecting the appropriate intensity progression for exercise regimens^{51,93,99} or providing feedbacks on trunk control⁷⁹, while the remaining 11 delivered automated advice or feedback about the data entered in the app based on decision-support systems or automated algorithms, with minimised involvement of HCPs^{35,41,59,68,71,78,83,87,94,96,103}.

On the other hand, 43 studies (62.3%) tested app interventions including different forms of HCP involvement, which typically engaged either clinicians or nurses. Among these, 23 did not pair

the HCP intervention with additional technological decision support: one study facilitated therapy optimisation by providers based on adherence overview³⁷; two studies included HCP involvement in classroom-based programmes³³ or group sessions⁴² as co-interventions to a lifestyle changing app with no automated feedbacks; two others supported professional monitoring of participants' physical activity with the possibility of adjusting goals⁹⁵ or reinforcing adherence⁹²; four enabled direct patient-physician communication through forums or real-time consultations^{44,46,47,70}, while the remaining 14 transferred data to dedicated professionals. A further distinction between this latter group of 14 studies can be made between studies where the data input was used by HCPs to provide personalised feedbacks during visits or share recommendations through text messages with a predefined frequency^{48,53,65,72,75,80,100,101}, and those where data transferred through the app continuously triggered study coordinator intervention whenever the responses suggested a deteriorating health^{36,45,55,73,85,90}.

The final 20 studies included both HCP involvement and higher-intensity technology automation. Some of the app interventions calculated and suggested optimal insulin doses under general professional oversight and with facilitated patient-physician communication^{51,56,64,66}, while the remaining provided automatic machine-based feedbacks as a first-level support with the possibility for professionals to intervene if warranted^{30,31,34,49,50,52,57,60,76,77,86,88,89,91,98,102}.

Only five studies provided HCP with a mobile app coupled with the patients' ones. In Cingi et al.'s study, physicians could view all patients' inputs and respond to messages and broadcast multimedia to participants through their app⁸⁵. Similarly, in other studies, a clinician app was associated with patient data^{34,45} or enabled patient-physician communication^{44,55}.

Regarding the behaviour change techniques (BCTs) adopted in the studies¹⁰⁴, the median number of BCT categories identified was eight (mean 7.46, SD 3.02). Twenty interventions (29.0%) included up to five BCTs, 37 had between five and ten BCTs, while the remaining 12 (17.4%) had more than ten BCTs. The maximum identified BCTs was 14^{67,68}, while the most common individual BCTs were: 'Instruction on how to perform a behaviour' ($n = 49$, 71.0%), 'Feedback on outcomes of behaviour' ($n = 41$, 59.4%), 'Social support' ($n = 39$, 56.5%), 'Adding objects to the environment' ($n = 39$, 56.5%), and 'Self-monitoring of outcome(s) of behaviour' ($n = 38$, 55.1%) (Fig. 2). Of the 93 possible BCTs identified in the taxonomy, 40 were used in at least one study and 14 in at least 14 (20%) studies.

Primary outcomes. Primary outcomes were explicitly identified in 56 studies (where a formal power calculation was performed) and were classified based on the taxonomy for outcomes in clinical research¹⁰⁵. The types of primary outcome measures varied in terms of core areas, although they most frequently pertained to physiological and clinical outcomes ($n = 29$, 51.8% of the total primary outcomes identified in the studies), followed by life impact ($n = 23$, 41.1%), resource use ($n = 3$, 5.4%) and adverse events ($n = 1$, 1.8%).

Regarding specific outcome domains, endocrine outcomes were the most commonly reported, haemoglobin A1c levels ($n = 20$), with only seven studies reporting significant differences between groups at follow-up^{31,45,57,60,64,65,70}. Among the remaining studies, seven analysed the impact on physical functioning and seven on delivery of care, both belonging to the life impact domain. Regarding physical functioning, three studies evaluated physical activity in terms of steps/day^{51,94,95}, while two assessed the impact on disease-specific self-care^{76,78} and two analysed the change in respiratory function parameters through the 6-min walking test^{93,99}. Statistically significant improvements were observed only in two studies^{61,78}. With regard to care delivery, six studies assessed the impact on medication adherence, through

Table 2. Developmental factors considered by the included studies ($n = 69$).

Name	Articles
Data security and privacy considerations	33 (47.8%)
Healthcare professional involvement in the design	18 (26.1%)
Pilot testing	28 (40.6%)
Supporting behavioural theory	11 (15.9%)
User involvement in the design	21 (30.4%)

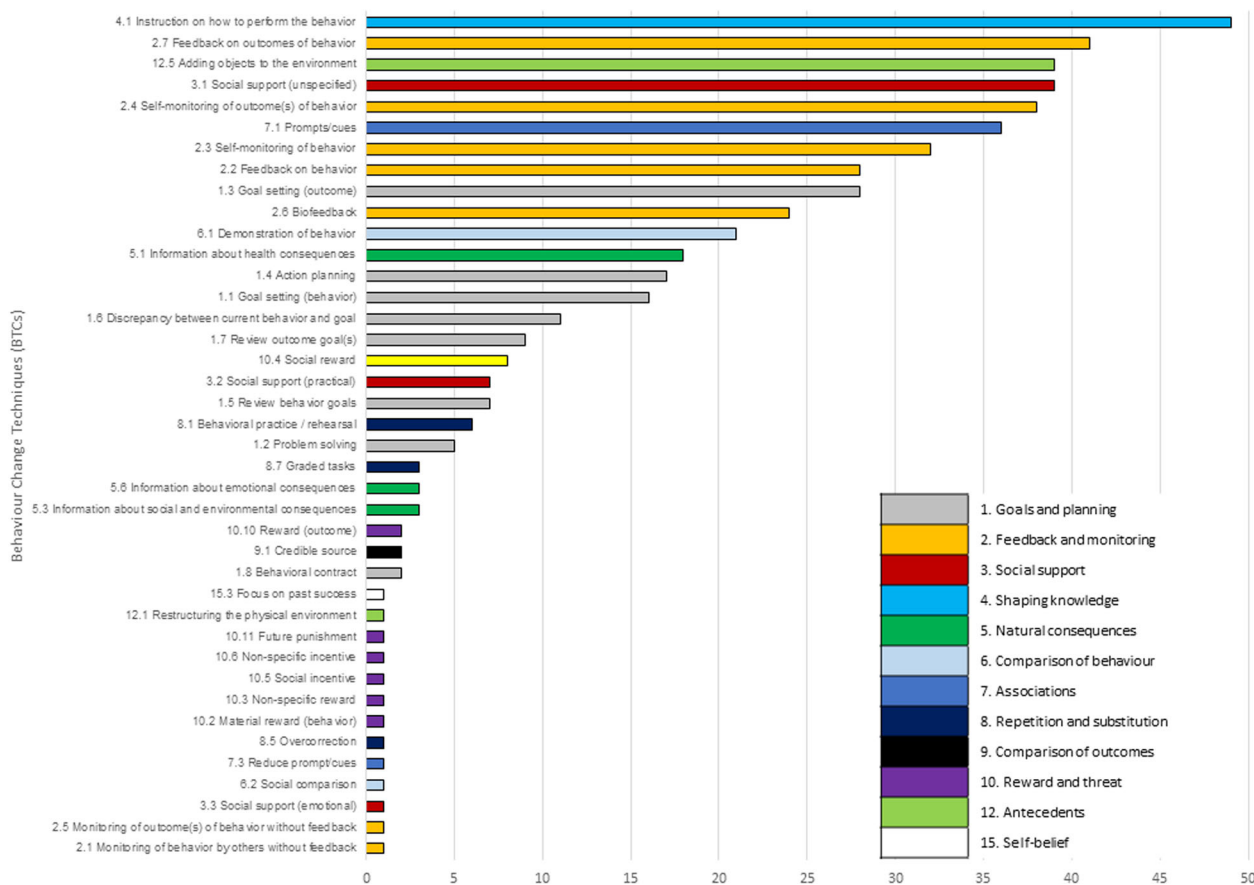


Fig. 2 BCT components adopted. Behaviour Change Techniques components and related clusters included in the selected studies ($n = 69$).

the administration of either a self-reported rating scale^{39,63,84,90}, a composite score⁸³ or pill count³⁶, while the remaining study evaluated a cardiac rehabilitation programme⁸⁰. Of these, five studies demonstrated significant outcome improvements^{39,63,80,83,84}. All the reported primary outcome domains classified by a core area and the statistical significance of the trial results (with p value ≤ 0.05) are shown in Table 3, with additional information in Supplementary Table 4.

Risk-of-bias assessment

Sixty-four (93%) studies presented an overall high risk of bias per the Cochrane RoB 2 tool, whilst only five had a low risk for at least four domains^{53,61,64,70,82}. The main issue was the plausible impossibility to blind study participants to the intervention, together with potential deviations from the intended interventions. The randomisation process and the selection of the reported results were the domains showing the lowest bias risk. The detailed, individual risk-of-bias analysis is provided in the Supplementary information materials (Supplementary Fig. 1).

DISCUSSION

We systematically reviewed the randomised studies on the impact of mHealth apps on four NCDs. We identified 69 studies (74 papers), published since 2008 from 24 countries, focusing on diabetes ($n = 29$), CVDs ($n = 13$), chronic respiratory diseases ($n = 13$), cancer ($n = 10$) or combinations of these.

The mHealth app impact was assessed on a wide range of primary outcomes. Per Dodd et al.'s taxonomy¹⁰⁵, they most frequently pertained to endocrine, cardiac or respiratory clinical outcomes ($n = 27$, e.g. glycated haemoglobin and systolic blood

pressure), but also physical, emotional or cognitive functioning ($n = 11$), care delivery ($n = 7$, e.g. medication adherence) and global quality of life ($n = 5$). Although a quantitative synthesis was not possible in this review due to the broad study aims and the subsequent heterogeneity of the studies, we noted inconclusive significance for many of them across chronic diseases. Statistically significant, improved primary outcomes were reported in 26 studies (eight on endocrine outcomes, five on delivery of care outcomes and three on cardiac outcomes, among others). These results confirm the widespread concerns in the scientific literature about the inability to couple the abundant production of mHealth apps with adequate vetting processes to generate evidence for their adoption^{106–108}.

This review focused on two main cornerstones that help explain the inadequacy of current peer-reviewed randomised studies of mHealth apps and could be leveraged to improve future ones.

On one side, there are important considerations related to app development, before the actual use in clinical studies. Our review evaluated several developmental factors for possible adoption in the design stage, all of which were inadequately recorded in the majority of the studies.

Despite the criticality of data security and privacy issues, often unaddressed with health apps^{109,110}, only 33 studies documented these aspects, even if preliminarily, in the trial report. Similarly, only about a third of them reported user and HCP engagement during the app design, through basic or more articulated (e.g. Delphi process) strategies. Although our findings show some improvements, vis à vis a previous review on diabetes self-management apps²⁴, more effort is needed to assign end-users a pivotal role in app development through the incorporation of their needs, expectations and experiences¹¹¹. Furthermore, only 28 (45.9%) of the non-pilot RCT studies mentioned previous pilot

Table 3. Primary outcome domains and statistical significance ($\alpha \leq 0.05$) of the included studies ($n = 56$).

Core area	Outcome domain	Positive results	Neutral results	Negative results
II. Clinical outcomes	3. Cardiac outcomes	3	0	0
	5. Endocrine outcomes	7	13	0
	9. General outcomes	1	0	0
	14. Metabolism and nutrition outcomes	1	0	0
	22. Respiratory, thoracic and mediastinal outcomes	0	4	0
III. Life impact	25. Physical functioning	2	5	0
	28. Emotional functioning/wellbeing	2	1	0
	29. Cognitive functioning	1	0	0
	30. Global quality of life	2	3	0
	32. Delivery of care	5	2	0
IV. Resource use	34. Economic	0	1	0
	35. Hospital	0	1	0
	27. Societal/care burden	1	0	0
V. Adverse events	38. Adverse events/effects	0	1	0

testing of the intervention, which seems a missed opportunity to prove a new intervention's viability, manage its risk and identify any deficiencies before substantial resource commitment. Finally, few interventions (11, 15.9%) were grounded in behavioural change theory. This shortcoming is currently a key missing element for digital health tools to achieve sustained behaviour change¹¹² and could be related to software developers overlooking important components for robust development and evaluation¹¹³.

These critical steps in the app development process gain further relevance considering that few mHealth solutions, among those investigated in RCTs, progress to wide availability in routine practice and major stores¹¹⁴, thus emphasising the paradox of wide adoption of apps untested in clinical research, while clinically investigated apps rarely scale up to real-world adoption.

As for HCP involvement, this review found 43 studies (62.3%) with additional human-led components. The support or encouragement provided to the patient by someone, whether it is directed at praising or rewarding behaviour or supporting self-management could decisively impact the intervention. A previous meta-analysis identified that, compared with diabetes apps with low-frequency HCP feedback, those with high-frequency feedback had a significant effect¹⁴. In the present review, studies involving additional human-led components showed no higher likelihood of a positive effect on outcomes for individuals in the intervention arm.

In addition, app-based interventions can differ greatly in their content and in the way they induce behaviour change. A median of eight BCTs was tapped by each mHealth intervention, with significant variability among them (SD 3.03). In apps containing gamification strategies, which share significant overlap with health BCTs and draw upon leader boards, prizes and rewards to motivate individuals¹¹⁵, the median rises to 14 BCTs with consistent inclusion of the 'reward and threat' category (81% apps)¹¹⁶. Gamification is indeed a significant factor to acknowledge in the development of mHealth apps for its promise to drive user behaviour and increase engagement through game elements, and ad hoc models have been proposed to support designers of gamified, condition-oriented solutions¹¹⁷. However, the adoption of gaming components was infrequent in the apps for chronic NCDs in this review, with only 12 (17.4%) studies implementing 'reward and threat' techniques. The most common behaviour change categories were instead 'Feedback and monitoring' (91%) and 'Shaping knowledge' (72%). High prevalence of self-regulatory

techniques and, to a lower extent, prompts such as alerts and reminders is coherent with previous studies^{118,119}, although these adopted an earlier 26-category taxonomy version¹²⁰ and specifically aimed at promoting physical activity. A striking difference from earlier reviews is the frequent incorporation of the BCT 'Adding objects to the environment', highlighting the ambivalence between the adoption of personal and study devices in intervention delivery. When an ad hoc device was provided after randomisation, an extra object was added in the environment to support the desired behaviour, and the related BCT was hence recorded. Lack of awareness and research on specific combinations of BCTs additionally emphasises the interventions' poor grounding in behaviour change theories.

Our findings highlight further methodological issues concerning app testing.

A recent review of study protocols available via clinicaltrials.gov examined the clinical evidence underlying digital health interventions, highlighting the studies' relatively low quality, small size and limited likelihood of being significantly powered to demonstrate treatment effects compared to drugs and traditional medical devices, which mandate stricter regulatory guidelines on safety and efficacy¹²¹. Although we considered published peer-reviewed reports of interventional studies and not study protocols, our results agree with the previous contribution: relatively short follow-ups—only 13 studies followed up >6 months—and median sample size of 94 patients, with several studies adopting convenience sampling and others not reaching the expected enrolment level or inadequately accounting for drop-out rates in their sample size calculations. Furthermore, 28 were pilot RCTs in need of subsequent investigation¹²². Also, formal risk-of-bias assessment performed according to standard tools revealed a high risk of bias for the majority of RCTs, highlighting the inability of both currently published studies and available assessment metrics to keep up with the specific methodological challenges of mobile apps.

These results should be assessed considering the ongoing literature debate on the evaluation of digital health interventions and mHealth apps. While some reject digital exceptionalism and accept RCTs as the golden standard¹²³, others think that digital technologies cannot be held to the same standards as new drugs/devices¹²⁴ and support the adoption of more agile, yet equally robust, methodologies¹²⁵. Given the distinguishing characteristics of mHealth apps¹²⁶, designs other than parallel-group RCTs have been proposed. These include the multiphase optimisation

strategy (MOST), the sequential multiple assignment randomised trial (SMART) and the micro-randomised trial^{127–130}. This also seems to be the direction embraced by some regulatory bodies such as—for instance—the Federal Institute for Drugs and Medical Devices in Germany (Bundesinstitut für Arzneimittel und Medizinprodukte, BfArM) that recognises the relevance of alternative study designs and methods such as Pragmatic Clinical Trials, SMART or MOST¹³¹.

Nonetheless, none of these was implemented in the peer-reviewed studies identified, emphasising the preference for parallel-group RCTs or checklists of quality criteria that show significant variability and no clear method for their development¹³².

The review results additionally light up the current debate: with 55% of the randomised studies showing neutral findings—i.e. non-significant difference—it is legitimate to discuss whether results could have differed with more robust development and design of the underlying apps, implementation of agile study designs or appropriate study characteristics of the RCT studies.

Limitations

There are several limitations of the current review. A meta-analysis was impossible due to the heterogeneity of studies explored. Non-English papers were excluded. We reviewed randomised studies only, which allowed focusing on the theoretically more robust evidence to assess the impact of mHealth apps in chronic disease while excluding other potentially relevant evidence sources. Furthermore, the taxonomy adopted for the BCT analysis was not specifically developed for mHealth apps, inevitably entailing authors' judgement. Finally, the coding was based on the information in the study report and related sources. Some of the features may have been untracked, possibly underestimating the number of BCTs tapped.

CONCLUSIONS

This systematic review sheds light on how design and development processes, choice of outcomes and behavioural change theory underpinning mHealth apps are fundamental in determining the value of digital interventions.

Furthermore, based on exploratory quantitative analyses of the association between study characteristics and a positive statistically significant result, our study seems to suggest a relatively lower success among apps for respiratory conditions (asthma, COPD) compared to CVDs, oncology or diabetes. This finding must be confirmed in future studies, but certainly highlights how the evidence on effective or cost-effective apps is not generalisable *tout court* across different health conditions. The widely adopted concept of 'equivalence' used in medical devices regulation systems must be carefully considered in the case of mHealth apps^{133,134}. Another interesting finding is the lower proportion of positive studies among those with longer (≥ 6 months) follow-up. This may signal issues with sustained adherence and engagement with the intervention and calls for research intended to identify strategies that maintain higher levels of engagement of end-users and successful retention rate in the long run.

Finally, among four different outcome areas, life impact outcomes seem to report more positive results compared to clinical outcomes, resource use or adverse events. This finding draws attention to the importance of the most appropriate outcome selection in the evaluation of mobile app interventions.

However, especially in view of the recently adopted EU Regulation of Medical Devices and the upcoming EU Regulation on Health Technology Assessment^{135,136}, follow-up research is urgently needed to further explore these elements that deserve deeper attention, confirm these initial hypotheses and to better

investigate the impact of BCTs on the effectiveness of mHealth apps.

Developers, methodologists and trialists are encouraged to carefully consider these additional elements when designing or running clinical evaluations of apps, to elucidate whether trial results reflect their true benefits or a biased estimate due to unsuitable designs or suboptimal implementation of the intervention.

Acknowledging the specific features of health apps (e.g. security issues, continuous content and software updates, and the impossibility of blinding participants and investigators)¹²⁶ is necessary to support policy-makers, providers and developers in their large-scale evaluation and consistent and effective deployment for chronic disease management.

The COVID-19 pandemic may have favoured large-scale adoption of digital tools¹³⁷, but only through sustained and appropriate evidence generation they will shape successful care models and significantly contribute to CD management.

METHODS

Information sources

This systematic review was guided by the PRISMA statement and checklist¹³⁸.

A literature search strategy was developed using keywords and MeSH (Medical Subject Headings) related to mHealth apps and pertinent chronic diseases. The full search strategy was first defined on Medline (Supplementary Note 1) and then adapted to the other engines. The search was performed on Medline (OVID interface), ScienceDirect, the Cochrane Library and Scopus, and last updated on 23 November 2019. To ensure literature saturation, we scanned the reference lists of the studies and systematic reviews and meta-analyses initially retrieved by the search or recently completed on PROSPERO and added additional publications of interest to the review. The review was not registered. A detailed protocol was prepared and circulated among the project team.

Eligibility criteria

The search strategy targeted trials with a randomised design involving individuals with one of the four main NCDs as identified by the World Health Organisation—cardiovascular diseases, cancers, chronic respiratory diseases and diabetes¹³⁹—and the use of a mobile health app, a small, self-contained software coded for a specific purpose and usually optimised to be downloaded and run on mobile phones and tablets¹⁴⁰. In terms of study design, we selected any type of randomised trials (parallel-group trials, cluster-randomised trials and cross-over trials, and other recent study designs specifically developed for digital technology). Concerning the target population, no further restriction criteria based on patient characteristics, such as age, gender, ethnicity or employment status were applied. All possible comparators were deemed relevant, including current standards of medical care as well as other eHealth interventions, such as telephone follow-up, text messaging or a simplified version of the app intervention. All quantitatively measured outcomes were included and classified based on Dodd et al.'s taxonomy¹⁰⁵. In terms of publication status, only peer-reviewed journal articles published in English after 2008 (when the App Store and PlayStore were first opened) were eligible. We excluded studies where the app was only used by HCPs or that focused on feasibility only. Further details about all inclusion and exclusion criteria are provided in Supplementary Table 5.

Study selection

Studies identified via database searches were first screened by title and abstract by two independent reviewers (F.P. and M.C.). Duplicates were excluded and, in case of disagreement, a third independent reviewer was consulted (R.T.). For studies meeting the inclusion criteria, or when unclear, the full article was retrieved. The same reviewers assessed the full-text papers to determine inclusion based on the stated eligibility criteria. For studies deemed ineligible for inclusion at the full-text analysis stage, the reasons for exclusion were recorded. Cohen's kappa was calculated to determine the inter-rater agreement between the reviewers.

Data collection and extraction

We used a pre-piloted data extraction form elaborating the CONSORT-EHEALTH checklist extension¹⁴¹ and the Cochrane Consumers and Communication Review Group's data extraction template¹⁴². Information was extracted from each study on: (i) study design and participant characteristics; (ii) app design and development; (iii) features of the mHealth intervention, including embedded BCTs¹⁰⁴; (iv) types of outcome measures and reported results.

Risk-of-bias assessment

Since no official methodological standards are available for empirical studies on mobile apps, we used the Revised Cochrane risk-of-bias tool for randomised trials (RoB 2)¹⁴³.

Assessment of risk of bias is regarded as an essential component of systematic reviews on the effects of an intervention. An evaluation of the risk of bias in each study included in the systematic review documents potential flaws in the evidence summarised and contributes to the certainty in the overall evidence. The RoB 2 tool provides a framework for assessing the risk of bias in a single estimate of an intervention effect reported from a randomised trial for a specific outcome or endpoint. This revised Cochrane tool was published in July 2019 and is structured into five different bias domains (i.e. risk of bias arising from the randomisation process, risk of bias due to deviations from the intended interventions (effect of assignment to intervention/effect of adhering to intervention), missing outcome data, risk of bias in the measurement of the outcome, risk of bias in the selection of the reported result). The assessment of each domain eventually leads to one of the following three judgements: low risk, some concerns or high risk. An overall risk of bias judgement, based on the five dimensions introduced before, is given to judge the overall confidence in the result. We assessed the treatment effect on the primary outcome or, when the primary outcome was not mentioned explicitly or could not be identified indirectly through sample size calculations, on the first outcome reported in the results¹⁴⁴. One co-author performed the quality assessment of studies (F.P.), and a second co-author independently double-checked the assessment (O.C.).

Data synthesis

Given the broad scope of the review, considerably high heterogeneity in terms of the type of NCDs, study characteristics, developmental factors, intervention structure and outcomes observed were anticipated. Therefore, the preferred approach to summarising the data was through a narrative synthesis, tabulation and descriptive analysis of all items extracted from the studies.

DATA AVAILABILITY

All data included in this study are available within the paper and its Supplementary information. Source data are available from the corresponding author upon request.

Received: 27 April 2021; Accepted: 13 September 2021;
Published online: 05 October 2021

REFERENCES

- Newzoo. Global mobile market report. <https://newzoo.com/insights/articles/newzoos-global-mobile-market-report-insights-into-the-worlds-3-2-billion-smartphone-users-the-devices-they-use-the-mobile-games-they-play> (2019).
- WHO. mHealth: New horizons for health through mobile technologies: second global survey on eHealth. <https://apps.who.int/iris/handle/10665/44607> (2011).
- Fagherazzi, G., Goetzinger, C., Rashid, M. A., Aguayo, G. A. & Huiart, L. Digital health strategies to fight COVID-19 worldwide: challenges, recommendations, and a call for papers. *J. Med. Internet Res.* **22**, e19284 (2020).
- Mahmood, S., Hasan, K., Colder Carras, M. & Labrique, A. Global preparedness against COVID-19: we must leverage the power of digital health. *JMIR Public Health Surveill.* **6**, e18980 (2020).
- Petracca, F., Ciani, O., Cucciniello, M. & Tarricone, R. Harnessing digital health technologies during and after the COVID-19 pandemic: context matters. *J. Med. Internet Res.* **22**, e21815 (2020).
- Meskó, B., Drobni, Z., Bényei, É., Gergely, B. & Gyórfy, Z. Digital health is a cultural transformation of traditional healthcare. *mHealth* **3**, 38 (2017).
- Hamine, S., Gerth-Guyette, E., Faulx, D., Green, B. B. & Ginsburg, A. S. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. *J. Med. Internet Res.* **17**, e52 (2015).
- Gee, P. M., Greenwood, D. A., Paterniti, D. A., Ward, D. & Miller, L. M. The eHealth Enhanced Chronic Care Model: a theory derivation approach. *J. Med. Internet Res.* **17**, e86 (2015).
- Barlow, J., Wright, C., Sheasby, J., Turner, A. & Hainsworth, J. Self-management approaches for people with chronic conditions: a review. *Patient Educ. Couns.* **48**, 177–187 (2002).
- Aujoulat, I., d'Hoore, W. & Deccache, A. Patient empowerment in theory and practice: polysemy or cacophony? *Patient Educ. Couns.* **66**, 13–20 (2007).
- Tomes, N. Patient empowerment and the dilemmas of late-modern medicalisation. *Lancet* **369**, 698–700 (2007).
- Cui, M., Wu, X., Mao, J., Wang, X. & Nie, M. T2DM self-management via smartphone applications: a systematic review and meta-analysis. *PLoS ONE* **11**, e0166718 (2016).
- Sun, C., Malcolm, J. C., Wong, B., Shorr, R. & Doyle, M. A. Improving glycemic control in adults and children with type 1 diabetes with the use of smartphone-based mobile applications: a systematic review. *Can. J. Diabetes* **43**, 51–58.e3 (2018).
- Martínez-García, M., Ruiz-Cárdenas, J. D. & Rabinovich, R. A. Effectiveness of smartphone devices in promoting physical activity and exercise in patients with chronic obstructive pulmonary disease: a systematic review. *COPD* **14**, 543–551 (2017).
- Hou, C., Carter, B., Hewitt, J., Francis, T. & Mayor, S. Do mobile phone applications improve glycemic control (HbA1c) in the self-management of diabetes? A systematic review, meta-analysis, and GRADE of 14 randomized trials. *Diabetes Care* **39**, 2089–2095 (2016).
- Bonoto, B. C. et al. Efficacy of mobile apps to support the care of patients with diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. *JMIR mHealth uHealth* **5**, e4 (2017).
- Wu, Y. et al. Mobile app-based interventions to support diabetes self-management: a systematic review of randomized controlled trials to identify functions associated with glycemic efficacy. *JMIR mHealth uHealth* **5**, e35 (2017).
- Alessa, T., Abdi, S., Hawley, M. S. & de Witte, L. Mobile apps to support the self-management of hypertension: systematic review of effectiveness, usability, and user satisfaction. *JMIR mHealth uHealth* **6**, e10723 (2018).
- Holmen, H., Wahl, A. K., Cvancarova Småstuen, M. & Ribul, L. Tailored communication within mobile apps for diabetes self-management: a systematic review. *J. Med. Internet Res.* **19**, e227 (2017).
- Armitage, L. C., Kassavou, A. & Sutton, S. Do mobile device apps designed to support medication adherence demonstrate efficacy? A systematic review of randomised controlled trials, with meta-analysis. *BMJ Open* **10**, e032045 (2020).
- Wu, X., Guo, X. & Zhang, Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. *JMIR mHealth uHealth* **7**, e12297 (2019).
- Nilsen, W. et al. Advancing the science of mHealth. *J. Health Commun.* **17**, 5–10 (2012).
- Schnall, R. et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J. Biomed. Inform.* **60**, 243–251 (2012).
- Adu, M. D., Malabu, U. H., Callander, E. J., Malau-Aduli, A. E. & Malau-Aduli, B. S. Considerations for the development of mobile phone apps to support diabetes self-management: systematic review. *JMIR mHealth uHealth* **6**, e10115 (2018).
- Riley, W. T. et al. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl. Behav. Med.* **1**, 53–71 (2011).
- Webb, T. L., Joseph, J., Yardley, L. & Michie, S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of

- theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J. Med. Internet Res.* **12**, e4 (2010).
27. Rowland, S. P., Fitzgerald, J. E., Holme, T., Powell, J. & McGregor, A. What is the clinical value of mHealth for patients? *NPJ Digit. Med.* **3**, 4 (2020).
 28. 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.* **18**, e97 (2016).
 29. Iribarren, S. J. et al. Effectiveness of mobile apps to promote health and manage disease: systematic review and meta-analysis of randomized controlled trials. *JMIR mHealth uHealth* **9**, e21563 (2021).
 30. Quinn, C. C. et al. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. *Diabetes Technol. Ther.* **10**, 160–168 (2008).
 31. Quinn, C. C. et al. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. *Diabetes Care* **34**, 1934–1942 (2011).
 32. Quinn, C. C. et al. The impact of a mobile diabetes health intervention on diabetes distress and depression among adults: secondary analysis of a cluster randomized controlled trial. *JMIR mHealth uHealth* **5**, e183 (2017).
 33. Forjuoh, S. N. et al. Behavioral and technological interventions targeting glycaemic control in a racially/ethnically diverse population: a randomized controlled trial. *BMC Public Health* **14**, 71 (2014).
 34. Moore, J. et al. Technology-supported apprenticeship in the management of hypertension: a randomized controlled trial. *J. Clin. Outcomes Manag.* **21**, 1 (2014).
 35. Athilingam, P., Jenkins, B., Johansson, M. & Labrador, M. A mobile health intervention to improve self-care in patients with heart failure: pilot randomized control trial. *JMIR Cardio* **1**, e3 (2017).
 36. Dang, S., Karanam, C. & Gómez-Marín, O. Outcomes of a mobile phone intervention for heart failure in a minority county hospital population. *Telem. J. E Health* **23**, 473–484 (2017).
 37. Frias, J. et al. Effectiveness of digital medicines to improve clinical outcomes in patients with uncontrolled hypertension and type 2 diabetes: prospective, open-label, cluster-randomized pilot clinical Trial. *J. Med. Internet Res.* **19**, e246 (2017).
 38. Labovitz, D. L., Shafner, L., Reyes Gil, M., Virmani, D. & Hanina, A. Using artificial intelligence to reduce the risk of nonadherence in patients on anticoagulation therapy. *Stroke* **48**, 1416–1419 (2017).
 39. Morawski, K. et al. Association of a smartphone application with medication adherence and blood pressure control: the MedSAFE-BP Randomized Clinical Trial. *JAMA Intern. Med.* **178**, 802–809 (2018).
 40. Rosen, K. D., Paniagua, S. M., Kazanis, W., Jones, S. & Potter, J. S. Quality of life among women diagnosed with breast Cancer: a randomized waitlist controlled trial of commercially available mobile app-delivered mindfulness training. *Psychonology* **27**, 2023–2030 (2018).
 41. Stukus, D. R. et al. Real-world evaluation of a mobile health application in children with asthma. *Ann. Allergy Asthma Immunol.* **120**, 395–400.e1 (2018).
 42. Wang, J., Cai, C., Padhye, N., Orlander, P. & Zare, M. A behavioral lifestyle intervention enhanced with multiple-behavior self-monitoring using mobile and connected tools for underserved individuals with type 2 diabetes and comorbid overweight or obesity: pilot comparative effectiveness trial. *JMIR mHealth uHealth* **6**, e92 (2018).
 43. Greer, J. A. et al. Randomized trial of a tailored cognitive-behavioral therapy mobile application for anxiety in patients with incurable cancer. *Oncologist* **24**, 1111–1120 (2019).
 44. Lyu, K. X. et al. Smartphone application wechat for clinical follow-up of discharged patients with head and neck tumors: a randomized controlled trial. *Chin. Med. J.* **129**, 2816–2823 (2016).
 45. Zhou, W., Chen, M., Yuan, J. & Sun, Y. Welltang - a smart phone-based diabetes management application - improves blood glucose control in Chinese people with diabetes. *Diabetes Res. Clin. Pract.* **116**, 105–110 (2016).
 46. Sun, Y. et al. Development and testing of an intelligent pain management system (ipms) on mobile phones through a randomized trial among chinese cancer patients: a new approach in cancer pain management. *JMIR mHealth uHealth* **5**, e108 (2017).
 47. Zhu, J., Ebert, L., Liu, X., Wei, D. & Chan, S. W. Mobile breast cancer e-support program for chinese women with breast cancer undergoing chemotherapy (part 2): multicenter randomized controlled trial. *JMIR mHealth uHealth* **6**, e104 (2018).
 48. Sun, C. et al. Mobile phone-based telemedicine practice in older chinese patients with type 2 diabetes mellitus: randomized controlled trial. *JMIR mHealth uHealth* **7**, e10664 (2019).
 49. Wang, Y. et al. Effects of continuous care for patients with type 2 diabetes using mobile health application: a randomised controlled trial. *Int. J. Health Plann. Manag.* **34**, 1025–1035 (2019).
 50. Yang, J. et al. Development and testing of a mobile app for pain management among cancer patients discharged from hospital treatment: randomized controlled trial. *JMIR mHealth uHealth* **7**, e12542 (2019).
 51. Rossi, M. C. et al. Diabetes Interactive Diary: a new telemedicine system enabling flexible diet and insulin therapy while improving quality of life: an open-label, international, multicenter, randomized study. *Diabetes Care* **33**, 109–115 (2010).
 52. Orsma, A. L. et al. Active assistance technology reduces glycosylated hemoglobin and weight in individuals with type 2 diabetes: results of a theory-based randomized trial. *Diabetes Technol. Ther.* **15**, 662–669 (2013).
 53. Torbjørnsen, A. et al. A low-intensity mobile health intervention with and without health counseling for persons with type 2 diabetes, part 1: baseline and short-term results from a randomized controlled trial in the norwegian part of RENEWING HEALTH. *JMIR mHealth uHealth* **2**, e52 (2014).
 54. Holmen, H. et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-year results from the norwegian randomized controlled trial RENEWING HEALTH. *JMIR mHealth uHealth* **2**, e57 (2014).
 55. Wayne, N., Perez, D. F., Kaplan, D. M. & Ritvo, P. Health coaching reduces hba1c in type 2 diabetic patients from a lower-socioeconomic status community: a randomized controlled trial. *J. Med. Internet Res.* **17**, e224 (2015).
 56. Bee, Y. M. et al. A smartphone application to deliver a treat-to-target insulin titration algorithm in insulin-naive patients with type 2 diabetes: a pilot Randomized Controlled Trial. *Diabetes Care* **39**, e174–e176 (2016).
 57. Kleinman, N. J., Shah, A., Shah, S., Phatak, S. & Viswanathan, V. Improved medication adherence and frequency of blood glucose self-testing using an m-health platform versus usual care in a multisite randomized clinical trial among people with type 2 diabetes in india. *Telem. J. E Health* **23**, 733–740 (2017).
 58. Alanzi, T., Bah, S., Alzahrani, S., Alshammari, S. & Almunsef, F. Evaluation of a mobile social networking application for improving diabetes Type 2 knowledge: an intervention study using WhatsApp. *J. Comp. Eff. Res.* **7**, 891–899 (2018).
 59. Agarwal, P. et al. Mobile app for improved self-management of type 2 diabetes: multicenter pragmatic Randomized Controlled Trial. *JMIR mHealth uHealth* **7**, e10321 (2019).
 60. Franc, S. et al. Efficacy of two telemonitoring systems to improve glycaemic control during basal insulin initiation in patients with type 2 diabetes: The TeleDiab-2 randomized controlled trial. *Diabetes Obes. Metab.* **21**, 2327–2332 (2019).
 61. Höchsmann, C. et al. Effectiveness of a behavior change technique-based smartphone game to improve intrinsic motivation and physical activity adherence in patients with type 2 diabetes: Randomized Controlled Trial. *JMIR Serious Games* **7**, e11444 (2019).
 62. Höchsmann, C. et al. Novel smartphone game improves physical activity behavior in type 2 diabetes. *Am. J. Prev. Med.* **57**, 41–50 (2019).
 63. Huang, Z. et al. A smartphone app to improve medication adherence in patients with type 2 diabetes in asia: feasibility Randomized Controlled Trial. *JMIR mHealth uHealth* **7**, e14914 (2019).
 64. Charpentier, G. et al. The Diabeo software enabling individualized insulin dose adjustments combined with telemedicine support improves HbA1c in poorly controlled type 1 diabetic patients: a 6-month, randomized, open-label, parallel-group, multicenter trial (TeleDiab 1 Study). *Diabetes Care* **34**, 533–539 (2011).
 65. Kirwan, M., Vandelanotte, C., Fenning, A. & Duncan, M. J. Diabetes self-management smartphone application for adults with type 1 diabetes: randomized controlled trial. *J. Med. Internet Res.* **15**, e235 (2013).
 66. Rossi, M. C. et al. Impact of the 'Diabetes Interactive Diary' telemedicine system on metabolic control, risk of hypoglycemia, and quality of life: a randomized clinical trial in type 1 diabetes. *Diabetes Technol. Ther.* **15**, 670–679 (2013).
 67. Drion, I. et al. The effects of a mobile phone application on quality of life in patients with type 1 diabetes mellitus: a Randomized Controlled Trial. *J. Diabetes Sci. Technol.* **9**, 1086–1091 (2015).
 68. Skrvøseth, S. O., Årsand, E., Godtlielsen, F. & Joakimsen, R. M. Data-Driven personalized feedback to patients with type 1 diabetes: a Randomized Trial. *Diabetes Technol. Ther.* **17**, 482–489 (2015).
 69. Burckhardt, M. A. et al. The use of continuous glucose monitoring with remote monitoring improves psychosocial measures in parents of children with type 1 diabetes: a Randomized Crossover Trial. *Diabetes Care* **41**, 2641–2643 (2018).
 70. Castensøe-Seidenfaden, P. et al. Testing a smartphone app (Young with Diabetes) to improve self-management of diabetes over 12 months: Randomized Controlled Trial. *JMIR mHealth uHealth* **6**, e141 (2018).
 71. Goyal, S. et al. A mobile app for the self-management of type 1 diabetes among adolescents: a Randomized Controlled Trial. *JMIR mHealth uHealth* **5**, e82 (2017).
 72. Klee, P. et al. An intervention by a patient-designed do-it-yourself mobile device app reduces hba1c in children and adolescents with type 1 diabetes: a randomized double-crossover study. *Diabetes Technol. Ther.* **20**, 797–805 (2018).

73. Baron, J. S., Hirani, S. & Newman, S. P. A randomised, controlled trial of the effects of a mobile telehealth intervention on clinical and patient-reported outcomes in people with poorly controlled diabetes. *J. Telemed. Telecare* **23**, 207–216 (2017).
74. Baron, J. S., Hirani, S. P. & Newman, S. P. Investigating the behavioural effects of a mobile-phone based home telehealth intervention in people with insulin-requiring diabetes: results of a randomized controlled trial with patient interviews. *J. Telemed. Telecare* **23**, 503–512 (2017).
75. Grady, M., Katz, L. B., Cameron, H. & Levy, B. L. Diabetes app-related text messages from health care professionals in conjunction with a new wireless glucose meter with a color range indicator improves glycemic control in patients with type 1 and type 2 diabetes: Randomized Controlled Trial. *JMIR Diabetes* **2**, e19 (2017).
76. Seto, E. et al. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. *J. Med. Internet Res.* **14**, e31 (2012).
77. Vuorinen, A. L. et al. Use of home telemonitoring to support multidisciplinary care of heart failure patients in Finland: randomized controlled trial. *J. Med. Internet Res.* **16**, e282 (2014).
78. Hägglund, E. et al. Patient-centred home-based management of heart failure. Findings from a randomised clinical trial evaluating a tablet computer for self-care, quality of life and effects on knowledge. *Scand. Cardiovasc. J.* **49**, 193–199 (2015).
79. Shin, D. C. & Song, C. H. Smartphone-based visual feedback trunk control training using a gyroscope and mirroring technology for stroke patients: single-blinded, randomized clinical trial of efficacy and feasibility. *Am. J. Phys. Med. Rehabil.* **95**, 319–329 (2016).
80. Varnfield, M. et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. *Heart* **100**, 1770–1779 (2014).
81. Whittaker, F. & Wade, V. The costs and benefits of technology-enabled, home-based cardiac rehabilitation measured in a randomised controlled trial. *J. Telemed. Telecare* **20**, 419–422 (2014).
82. Eyles, H. et al. A salt-reduction smartphone app supports lower-salt food purchases for people with cardiovascular disease: findings from the SaltSwitch randomised controlled trial. *Eur. J. Prev. Cardiol.* **24**, 1435–1444 (2017).
83. Johnston, N. et al. Effects of interactive patient smartphone support app on drug adherence and lifestyle changes in myocardial infarction patients: a randomized study. *Am. Heart J.* **178**, 85–94 (2016).
84. Santo, K. et al. MEDication reminder APPs to improve medication adherence in Coronary Heart Disease (MedApp-CHD) Study: a randomised controlled trial protocol. *BMJ Open* **7**, e017540 (2017).
85. Cingi, C. et al. The "physician on call patient engagement trial" (POPET): measuring the impact of a mobile patient engagement application on health outcomes and quality of life in allergic rhinitis and asthma patients. *Int. Forum Allergy Rhinol.* **5**, 487–497 (2015).
86. Zairina, E. et al. Telehealth to improve asthma control in pregnancy: a randomized controlled trial. *Respirology* **21**, 867–874 (2016).
87. Liu, W. T. et al. A mobile telephone-based interactive self-care system improves asthma control. *Eur. Respir. J.* **37**, 310–317 (2011).
88. Ryan, D. et al. Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. *BMJ* **344**, e1756 (2012).
89. Kim, M. Y. et al. Feasibility of a smartphone application based action plan and monitoring in asthma. *Asia Pac. Allergy* **6**, 174–180 (2016).
90. Kosse, R. C. et al. Effect of a mHealth intervention on adherence in adolescents with asthma: a randomized controlled trial. *Respir. Med.* **149**, 45–51 (2019).
91. Farmer, A. et al. Self-management support using a digital health system compared with usual care for chronic obstructive pulmonary disease: Randomized Controlled Trial. *J. Med. Internet Res.* **19**, e144 (2017).
92. Wang, C. H. et al. Mobile-phone-based home exercise training program decreases systemic inflammation in COPD: a pilot study. *BMC Pulm. Med.* **14**, 142 (2014).
93. Kwon, H. et al. An mhealth management platform for patients with chronic obstructive pulmonary disease (efil breath): Randomized Controlled Trial. *JMIR mHealth uHealth* **6**, e10502 (2018).
94. Tabak, M., Vollenbroek-Hutten, M. M., van der Valk, P. D., van der Palen, J. & Hermens, H. J. A telerehabilitation intervention for patients with Chronic Obstructive Pulmonary Disease: a randomized controlled pilot trial. *Clin. Rehabil.* **28**, 582–591 (2014).
95. Vorrink, S. N., Kort, H. S., Troosters, T., Zanen, P. & Lammers, J. J. Efficacy of an mHealth intervention to stimulate physical activity in COPD patients after pulmonary rehabilitation. *Eur. Respir. J.* **48**, 1019–1029 (2016).
96. Boer, L. et al. A smart mobile health tool versus a paper action plan to support self-management of chronic obstructive pulmonary disease exacerbations: Randomized Controlled Trial. *JMIR mHealth uHealth* **7**, e14408 (2019).
97. Foley, N. M. et al. PATI: patient accessed tailored information: a pilot study to evaluate the effect on preoperative breast cancer patients of information delivered via a mobile application. *Breast* **30**, 54–58 (2016).
98. Kearney, N. et al. Evaluation of a mobile phone-based, advanced symptom management system (ASyMS) in the management of chemotherapy-related toxicity. *Support Care Cancer* **17**, 437–444 (2009).
99. Ji, W. et al. Mobile health management platform-based pulmonary rehabilitation for patients with non-small cell lung cancer: prospective clinical trial. *JMIR mHealth uHealth* **7**, e12645 (2019).
100. Egbring, M. et al. A mobile app to stabilize daily functional activity of breast cancer patients in collaboration with the physician: a Randomized Controlled Clinical Trial. *J. Med. Internet Res.* **18**, e238 (2016).
101. Karhula, T. et al. Telemonitoring and mobile phone-based health coaching among finnish diabetic and heart disease patients: Randomized Controlled Trial. *J. Med. Internet Res.* **17**, e153 (2015).
102. Logan, A. G. et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. *Hypertension* **60**, 51–57 (2012).
103. Or, C. & Tao, D. A 3-month randomized controlled pilot trial of a patient-centered, computer-based self-monitoring system for the care of type 2 diabetes mellitus and hypertension. *J. Med. Syst.* **40**, 81 (2016).
104. Michie, S. et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann. Behav. Med.* **46**, 81–95 (2013).
105. Dodd, S. et al. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J. Clin. Epidemiol.* **96**, 84–92 (2018).
106. Eapen, Z. J. & Peterson, E. D. Can mobile health applications facilitate meaningful behavior change?: time for answers. *JAMA* **314**, 1236–1237 (2015).
107. Jimenez, G., Lum, E. & Car, J. Examining diabetes management apps recommended from a Google search: content analysis. *JMIR Mhealth Uhealth* **7**, e11848 (2019).
108. Lum, E. et al. Decision support and alerts of apps for self-management of blood glucose for type 2 diabetes. *JAMA* **321**, 1530–1532 (2019).
109. Huckvale, K., Prieto, J. T., Tilney, M., Benghozi, P. J. & Car, J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Med.* **13**, 214 (2015).
110. Grundy, Q. et al. Data sharing practices of medicines related apps and the mobile ecosystem: traffic, content, and network analysis. *BMJ* **364**, 1920 (2019).
111. McCurdie, T. et al. mHealth consumer apps: the case for user-centered design. *Biomed. Instrum. Technol.* **46**, 49–56 (2012).
112. Klonoff, D. C. Behavioral theory: the missing ingredient for digital health tools to change behavior and increase adherence. *J. Diabetes Sci. Technol.* **13**, 276–281 (2019).
113. Cowan, L. T. et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. *Health Educ. Behav.* **40**, 133–139 (2013).
114. Veazie, S. et al. Rapid evidence review of mobile applications for self-management of diabetes. *J. Gen. Intern. Med.* **33**, 1167–1176 (2018).
115. Cugelman, B. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games* **1**, e3 (2013).
116. Edwards, E. A. et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. *BMJ Open* **6**, e012447 (2016).
117. Giunti, G. 3MD for chronic conditions, a model for motivational mHealth design: embedded case study. *JMIR Serious Games* **6**, e11631 (2018).
118. de Korte, E., Wiezer, N., Bakhuys Roozeboom, M., Vink, P. & Kraaij, W. Behavior change techniques in mhealth apps for the mental and physical health of employees: systematic assessment. *JMIR mHealth uHealth* **6**, e167 (2018).
119. Direito, A. et al. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques?. *BMC Public Health* **14**, 646 (2014).
120. Abraham, C. & Michie, S. A taxonomy of behavior change techniques used in interventions. *Health Psychol.* **27**, 379–387 (2008).
121. Chen, C. E., Harrington, R. A., Desai, S. A., Mahaffey, K. W. & Turakhia, M. P. Characteristics of digital health studies registered in ClinicalTrials.gov. *JAMA Intern. Med.* **179**, 838–840 (2019).
122. Byambasuren, O., Sanders, S., Beller, E. & Glasziou, P. Prescribable mHealth apps identified from an overview of systematic reviews. *NPJ Digit. Med.* **1**, 12 (2018).
123. McCartney, M. Margaret McCartney: innovation without sufficient evidence is a disservice to all. *BMJ* **358**, j3980 (2017).
124. The Lancet. Is digital medicine different? *Lancet* **392**, 95 (2018).
125. Greaves, F. et al. What is an appropriate level of evidence for a digital health intervention? *Lancet* **392**, 2665–2667 (2019).

126. Tarricone, R., Petracca, F., Ciani, O. & Cucciniello, M. Distinguishing features in the assessment of mHealth apps. *Expert Rev. Pharmacoecon. Outcomes Res.* **21**, 521–526 (2021).
127. Pham, Q., Wiljer, D. & Cafazzo, J. A. Beyond the randomized controlled trial: a review of alternatives in mhealth clinical trial methods. *JMIR mHealth uHealth* **4**, e107 (2016).
128. Mohr, D. C., Cheung, K., Schueller, S. M., Hendricks Brown, C. & Duan, N. Continuous evaluation of evolving behavioral intervention technologies. *Am. J. Prev. Med.* **45**, 517–523 (2013).
129. Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D. & Murphy, S. A. A "SMART" design for building individualized treatment sequences. *Annu. Rev. Clin. Psychol.* **8**, 21–48 (2012).
130. Klasnja, P. et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. *Health Psychol.* **34S**, 1220–1228 (2015).
131. Federal Institute for Drugs and Medical Devices (BfArM). The fast-track process for digital health applications (DiGA) according to Section 139e SGB V. A Guide for Manufacturers, Service Providers and Users. https://www.bfarm.de/SharedDocs/Downloads/EN/MedicalDevices/DiGA_Guide.pdf?session%20id=CF CB8E00B0826085FAEDB0A35F406F49.2_cid3547__blob=publicationFile&v=2 (2020).
132. Agarwal, P. et al. Assessing the quality of mobile applications in chronic disease management: a scoping review. *NPJ Digit. Med.* **4**, 46 (2021).
133. Medical Device Coordination Group (MDCG). Guidance on clinical evaluation (MDR)/performance evaluation (IVDR) of medical device software. https://ec.europa.eu/health/sites/default/files/md_sector/docs/md_mdcg_2020_1_guidance_clin_eva_md_software_en.pdf (2020).
134. Medical Device Coordination Group (MDCG). Guidance on qualification and classification of software in regulation (EU) 2017/745 – MDR and regulation (EU) 2017/746 – IVDR. https://ec.europa.eu/health/sites/default/files/md_topics-interest/docs/md_mdcg_2019_11_guidance_en.pdf (2019).
135. European Union. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No. 178/2002 and Regulation (EC) No. 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. *Off. J. Eur. Union* **117**, 1–175 (2017).
136. European Commission. *Proposal for a Regulation of the European Parliament and of the Council on Health Technology Assessment and Amending Directive 2011/24/EU* (European Commission, 2018).
137. Golinelli, D. et al. Adoption of digital technologies in health care during the covid-19 pandemic: systematic review of early scientific literature. *J. Med Internet Res* **22**, e22280 (2020).
138. Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G., PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J. Clin. Epidemiol.* **62**, 1006–1012 (2009).
139. WHO. 10 facts on noncommunicable diseases. http://www.who.int/features/factfiles/noncommunicable_diseases/en/ (2013).
140. Boulos, M. N., Brewer, A. C., Karimkhani, C., Buller, D. B. & Dellavalle, R. P. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online J. Public Health Inf.* **5**, 229 (2014).
141. Eysenbach, G., Consort-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J. Med. Internet Res.* **13**, e1923 (2011).
142. Ryan, R., Synnot, A., Prictor, M., & Hill, S. Cochrane consumers and communication group data extraction template for included studies. <http://cccr.org/cochrane.org/author-resources> (2016).
143. Sterne, J. et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, l4898 (2019).

144. Gates, A. et al. Evaluation of the reliability, usability, and applicability of AMSTAR, AMSTAR 2, and ROBIS: protocol for a descriptive analytic study. *Syst. Rev.* **7**, 85 (2018).

ACKNOWLEDGEMENTS

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement #779306 (COMED—Pushing the Boundaries of Cost and Outcome Analysis of Medical Technologies).

AUTHOR CONTRIBUTIONS

M.C., F.P., O.C. and R.T. designed the study, M.C. developed the search strategy, F.P. conducted the original literature searches, M.C. and F.P. were involved in data screening and study selection. F.P. extracted the data, while M.C. checked extracted data for consistency and R.T. mediated where there was disagreement or uncertainty over inclusion. F.P. and O.C. appraised the risk of bias using the Cochrane RoB 2 tool. F.P. synthesised findings. All authors contributed to and approved the final manuscript.

COMPETING INTERESTS

All authors are involved in an RCT to evaluate a mobile supportive care app for patients with metastatic lung cancer. The authors declare no other relevant competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-021-00517-1>.

Correspondence and requests for materials should be addressed to Rosanna Tarricone.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021