The European Journal of Public Health, Vol. 26, No. 4, 602–610

doi:10.1093/eurpub/ckw090 Advance Access published on 21 June 2016

Mobile and Web 2.0 interventions for weight management: an overview of review evidence and its methodological quality

Marco Bardus[1,2], Jane R. Smith[2], Laya Samaha[3], Charles Abraham[2]

1 Department of Health Promotion and Community Health, American University of Beirut, Beirut, Lebanon
2 Psychology Applied to Health group, University of Exeter Medical School, Exeter, UK
3 Università della Svizzera italiana, Institute of Health Communication, Lugano, Switzerland

Correspondence: Marco Bardus, American University of Beirut, Beirut, Lebanon, Department of Health Promotion and Community Health, Faculty of Health Sciences, PO Box 11-0236, Riad El-Sohl, Beirut 1107 2020, Lebanon; Tel: +961 1 350 000 (ext. 4694), e-mail: marco.bardus@gmail.com

Background: The use of Internet and related technologies for promoting weight management (WM), physical activity (PA), or dietary-related behaviours has been examined in many articles and systematic reviews. This overview aims to summarize and assess the quality of the review evidence specifically focusing on mobile and Web 2.0 technologies, which are the most utilized, currently available technologies. **Methods:** Following a registered protocol (CRD42014010323), we searched 16 databases for articles published in English until 31 December 2014 discussing the use of either mobile or Web 2.0 technologies to promote WM or related behaviors, i.e. diet and physical activity (PA). Two reviewers independently selected reviews and assessed their methodological quality using the AMSTAR checklist. Citation matrices were used to determine the overlap among reviews. **Results:** Forty-four eligible reviews were identified, 39 of which evaluated the effects of interventions using mobile or Web 2.0 technologies. Methodological quality was generally low with only 7 reviews (16%) meeting the highest standards. Suggestive evidence exists for positive effects of mobile technologies on weight-related outcomes and, to a lesser extent, PA. Evidence is inconclusive regarding Web 2.0 technologies. **Conclusions**: Reviews on mobile and Web 2.0 interventions for WM and related behaviors suggest that these technologies can, under certain circumstances, be effective, but conclusions are limited by poor review quality based on a heterogeneous evidence base.

Introduction

Many noncommunicable diseases (NCDs) can be prevented by addressing modifiable behavior patterns associated with raised blood glucose, elevated blood lipids, and obesity,[1] in particular, weight management (WM) through healthy eating and physical activity (PA).[2] Behavioral WM interventions have had limited impact[3] and information and services delivered through the Internet and related technologies ('eHealth')[4] may enhance users capacity to change their behavior and manage their weight.[5] Internet technologies now incorporate Web 2.0 tools (also referred to as 'social media'), such as social networking sites, blogs and content communities,[6] which are actively used by 2.8 billion people worldwide.[7] Social media profiles are increasingly accessed through mobile phones or smartphones,[7] which represent 38% of the estimated 7.1 billion global mobile connections.[8]

Several reviews of Internet- and computer-based interventions report generally positive effects for smoking cessation,[9] diet and PA,[10] and WM[11,12] interventions. However, despite mobile phones and Web 2.0 now being the most utilized technologies, it remains unclear how these have been effectively employed in interventions. A recent scoping review[13] identified a large number of reviews including interventions utilizing mobile and Web 2.0 technologies to facilitate WM. It is unclear, however whether these reviews provide methodologically sound evidence on effectiveness. To date, only a few overviews have systematically assessed the methodological quality of eHealth reviews. These focused on online interventions promoting lifestyle behaviors[14] for somatic diseases,[15] and self-directed interventions for weight loss.[16] None focused on Web 2.0 or mobile technologies or used a citation matrix to estimate overlap in the evidence base of the included studies. Although infrequently used,[17] citation matrices can usefully assess the extent to which similar reviews are comparable and draw conclusions from a coherent evidence base.[17,18]

This overview aimed to: (1) assess the methodological quality of reviews of mobile and Web 2.0 technologies to support WM and related behaviors (i.e. PA and diet), and (2) summarize review evidence on the use and effectiveness these technologies, taking into account the quality and breadth of the evidence base of primary studies included in the reviews. Reviews are the main unit of analysis (henceforth referred to as 'reviews'). The studies included in those reviews are henceforth referred to as 'research studies' or just 'studies'.

Methods

Data sources

Following a registered protocol (PROSPERO: CRD42014010323)[19] articles were identified through a comprehensive search of 16 electronic databases, and 'grey literature' resources (WorldCAT Dissertations and OpenGrey). Reference lists of included articles were also checked.

Search strategy

Applying the PICO framework, keywords and MeSH terms were used to describe the Population (e.g. obesity, overweight),[20] Interventions/Comparators (mobile[21,22] and Web 2.0 technologies[23,24]), and Outcomes (e.g. weight loss, BMI, diet and PA).[25] The Medline search strategy is provided in Supplementary table 1. Preliminary searches were pilot-tested in June-August 2014

and refined with input from an information specialist and expert systematic reviewer. Finalized searches were performed in August 2014 and updated in February 2015 to identify peer-reviewed articles published in English between 1 January 2004 and 31 December 2014, to ensure relevance to current mobile and Web 2.0 technologies.

Inclusion/exclusion criteria

Reviews included in this overview had to evaluate studies involving either mobile and/or Web 2.0 technologies for WM and/or related behaviors (i.e. PA and diet). We defined mobile devices as: 'mobile phones, personal digital assistants (PDA), handheld and ultra-portable computers such as tablets, and smartphones'[26], which encompassed devices delivering interventions through text messaging or standalone apps. Web 2.0 technologies were: 'social networking sites, collaborative projects, micro-blogging and blogging tools, content communities, virtual worlds'.[6] We included only 'reliable systematic reviews', defined in the *NICE NHS Evidence Process and Methods Manual*[27] as those published in a journal conforming to the PRISMA standards[28] or presenting inclusion/exclusion criteria, and confirming in the abstract that a synthesis of studies from two or more information sources had been undertaken.

We excluded reviews that: did not focus on influencing WM, PA, or diet (e.g. disease management, dietary or PA assessment using mobile devices unless they also discussed their use for behavior change); analysed only computer or Internet-based interventions that did not include any studies using mobile or Web 2.0 technologies; or were not deemed reliable.[27]

Review selection

Reviews were selected using a three-step process involving two reviewers who independently: (1) screened titles, abstracts and full texts where necessary and applied the inclusion/exclusion criteria as per the study protocol;[19] (2) read full-texts and categorized selected articles according to their focus (i.e. behavioral assessment or behavior change) and research type (i.e. design and development, feasibility, evaluations) as detailed elsewhere;[13] (3) selected only 'reliable'[27] systematic reviews, that focused on behavior change. Reliability of selection was assessed (see below) and consensus achieved through discussion.

Study quality assessment and data extraction

Two reviewers independently applied the 'A Measurement Tool to Assess Systematic Reviews' (AMSTAR)[29] checklist, which consists of 11 items and is considered a valid and reliable instrument for assessing review reporting and methodological quality.[30,31] Supplementary materials (e.g. protocols), where available, were considered when applying the checklist. As done in other overviews using AMSTAR,[32] an overall score (range: 0–11) was used to categorize reviews according to their percentile rank (e.g. $>90^{th}$ percentile = A; $<50^{th}$ percentile = F): A-B corresponded to 'high', C-D to 'medium or moderate', and E-F to 'low' quality.

The following information was extracted for each review: target population, interventions/comparators considered, outcomes reported, any study design criteria for inclusion; number of studies reviewed in relation to WM, PA or diet; databases searched and years covered; summary of findings related to WM, PA or diet, including reported measures of effect sizes (e.g. Cohen's *d* values, standardized mean difference SMD, etc.), sample sizes and details of the design and quality of studies where available. The first author extracted all information. The second author checked data extraction and independently extracted data from all meta-analyses.

Inter-rater reliability was assessed for each step of study selection, quality assessment and data extraction using Gwet's first-order agreement coefficient (AC1) statistic, rather than Cohen's *kappa*, as the latter tends to underestimate reliability when there is an asymmetric distribution between agreements and disagreements.[33]

The number of citations for each review (listed by Google Scholar) was translated into an 'average-per-year' score by dividing the number of citations by the years since publication. A citation matrix was created to determine the overlap of the evidence base among the included reviews. As proposed by Pieper and colleagues,[18] the overlap is expressed as 'corrected covered area' (CCA). The CCA is calculated as: $(N - r)/((r \times c) - r)$, where '*r*' is the number of unique citations (representing the rows of the matrix), 'c' is the number of reviews (columns), and N is the total number of citations in the area ($r \times c$). CCA thresholds are: 0% = none; <5% = slight; <10% = moderate; <15% = high; >15% = very high overlap.

Data synthesis

Review characteristics, methodological quality, and data on the overlap in the evidence base were summarized using frequency tables and descriptive statistics, with reviews divided into inductively developed categories (see results). Evidence on the effects of technologies on weight-related outcomes and related behaviors was narratively synthesized, reflecting information regarding study design and quality where review authors commented on these aspects.

Results

Included/excluded articles

The search and reference lists generated 6128 hits. After duplicate removal, titles and abstracts of 4540 records were screened for inclusion, and 668 articles were reviewed in full-text (92% agreement, AC1 = 0.91, 95% CI = 0.89 to 0.92). Of these, 457 were categorized (90% agreement, AC1 = 0.82, 95% CI = 0.79 to 0.85) as described elsewhere;[13] 413 articles were subsequently excluded from this overview with reasons (Supplementary table 2). This left 44 articles reporting systematic reviews, 28 of which (63%) were published since 2013, the first in 2007[34] (PRISMA[28] flow diagram, figure 1).

Review characteristics

Data extraction from meta-analyses demonstrated good reliability (85% agreement, AC1 = 0.83, 95% CI = 0.73 to 0.93). Based on the extracted data, reviews were inductively categorized according to *review type*, *behaviour focus* and *technology focus* (table 1). Review type included: 'scoping reviews' ($n = 5$, 11%), 'narrative syntheses' ($n = 28$, 64%) and 'meta-analyses' ($n = 11$, 25%), depending on the type of data and approach to synthesis of findings from research studies. We defined 'scoping reviews' as systematic reviews describing the field of study and highlighting potential gaps without formally synthesizing findings, such as are commonly undertaken before conducting systematic reviews of effectiveness, which have more focused research questions, detailed data extraction, and study quality assessment.[35,36] 'Narrative syntheses' were defined as systematic reviews providing only a text-based, or vote-counting summary of findings. 'Meta-analyses' were systematic reviews including quantitative estimates of effects. Behavior focus covered: (i) 'Multiple health behaviors', for reviews of interventions addressing various health behaviors, including - but not restricted to - PA, diet and/or weight loss; (ii) 'Weight management', for interventions focused on weight-related outcomes (e.g. changes in BMI, weight); and (ii) 'PA'. Technology focus included: (i) 'eHealth', for reviews reporting on various technologies, including - but not restricted to - Internet, mobile phones and Web 2.0; (ii) 'Mobile', focusing on mobile technologies (text messaging, apps, etc.) alone; and (iii) 'Web 2.0', for those focusing on social media (e.g. social networking sites).

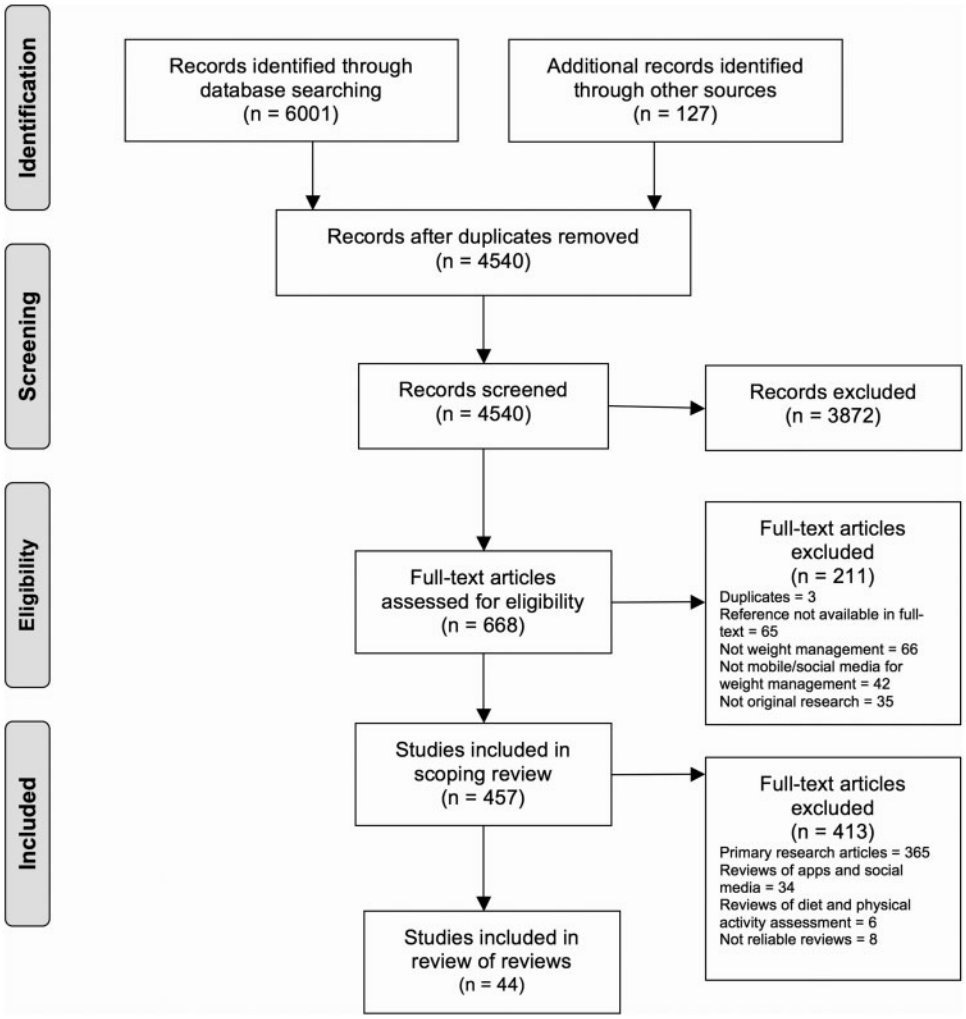

Figure 1 PRISMA 2009 flow diagram

Table 1 Frequency distribution of the reviews according to review type, behavior focus, and technology

Technology focus	eHealth in general (incl. mobile and Web 2.0)		Mobile only		Web 2.0 only		Total	
Review method and behavior focus	n	% of total	n	% of total	n	% of total	n	Total
1. General Scoping Reviews	*0*	*0%*	*4*	*9%*	*1*	*2%*	*5*	*11%*
Multiple health behaviors ([a])	0	0%	4	9%	1	2%	5	11%
3. Narrative Syntheses	*12*	*27%*	*15*	*34%*	*1*	*2%*	*28*	*64%*
Multiple health behaviors ([a])	1	2%	6	14%	0	0%	7	16%
Weight management	8	18%	4	9%	1	2%	13	30%
PA behavior change	3	7%	5	11%	0	0%	8	18%
4. Meta-analyses	*1*	*2%*	*5*	*11%*	*5*	*11%*	*11*	*25%*
Multiple health behaviors ([a])	1	2%	1	2%	1	2%	3	7%
Weight management	0	0%	3	7%	1	2%	4	9%
PA behavior change	0	0%	1	2%	3	7%	4	9%
Total	*13*	*30%*	*24*	*55%*	*7*	*16%*	*44*	*100%*

PA, physical activity; ([a]) Multiple health behaviors includes weight management, PA and/or diet but also covers additional behaviors.

All scoping reviews addressed multiple health behaviors (IDs: 7, 19, 28, 37); four out of five focused on mobile technologies. Most narrative syntheses (13/28, 46%) reported on WM interventions, with the majority of these (8/13, 62%) covering various eHealth technologies (IDs: 2, 3, 6, 13, 15, 26, 33, 41). Eight narrative syntheses (8/28, 29%) analysed PA interventions, the majority of which (5/8, 63%) focused on mobile technologies (IDs: 8-10, 34, 43), including text-messaging alone (ID: 38) or novel mobile apps (IDs: 1, 40). Of the 11 meta-analyses, four (36%) focused on WM, with three of these covering mobile technologies only (IDs: 27, 30, 39). Diet was not the specific focus of any review, but interventions targeting diet were included in reviews of 'multiple health behaviors' and in one meta-analysis explicitly analyzing Web 2.0 interventions for diet and PA (ID: 44). Detailed characteristics of the selected reviews are reported in Supplementary table 4.

Review quality

Duplicate AMSTAR quality assessment achieved 90% agreement and good reliability (AC1 = 0.80, 95% CI = 0.75 to 0.85). However, four checklist definitions produced suboptimal reliability: 'Independent duplicate study selection/data extraction' (item A2) [68% agreement, AC1 = 0.29, 95% CI = 0.01 to 0.56]; 'Declaration of conflict of interests' (A11) [80% agreement, AC1 = 0.61, 95% CI = 0.38 to 0.84], 'Appropriate use of scientific quality when formulating conclusions' (A8) [82% agreement, AC1 = 0.67, 95% CI = 0.47 to 0.88], and 'Appropriate use of methods to combine findings' (A9) [84% of agreement, AC1 = 0.68, 95% CI = 0.74 to 0.94]. These inter-rater reliability assessments, especially those below AC1 = 0.65, suggest that the item definitions were interpreted differently. All disagreements were resolved through discussion.

Supplementary table 3 presents the AMSTAR quality items and scores, with the summary statistics on overlap in the evidence base. For scoping reviews and narrative syntheses, the AMSTAR ranking is based on 10 out of 11 checklist items, excluding the assessment of publication bias (A10), which was not reported in any of these reviews. This scale adjustment affected the ranking of 10 (34%) reviews (IDs: 2, 3, 6, 10, 12, 14, 29, 33, 34, 42), but only one (ID: 6) improved from a low (E) to a medium (D) category. The citation matrix did not include scoping reviews, as not all of these provided a list of included studies.

Scoping reviews had a median percentile rank of 10 (F, low quality) and none scored above the 40th percentile. No review reported a list of included and excluded studies (A5), formally assessed scientific quality, or consequently considered quality when formulating conclusions (A7 and A8). Only one (ID: 23) was based on an *a priori* study protocol (A1), and another (ID: 19) used a comprehensive search strategy (A3). The average AMSTAR score for the 28 narrative syntheses also indicated low quality (Median percentile rank = 30, F), with the majority (24/28, 86%) ranking low. Two were of high quality (IDs: 3, 5). Only four each reported a list of included and excluded studies (A5) [IDs: 3, 5, 20, 41], were based on a published protocol (A1) [IDs: 2, 3, 5, 10], or appropriately accounted for scientific quality when formulating conclusions (A8) [IDs: 5, 33, 34, 41]. The 11 meta-analyses were of higher, moderate quality (Median percentile rank = 64, D), with five achieving high quality scores (IDs: 21, 22, 27, 36, 44). The most frequently unreported AMSTAR items were: A1 (study protocol), A4 (grey literature), and A5 (lists of both included and excluded studies). There was a significant difference ($t = -3.15$, $df = 41.90$, $P = 0.003$) in AMSTAR scores between the 28 reviews published in the last 2 years [M(SD) = 5.21 (3.29)], scoring on average 2.4 points higher (95% CI = −3.94 to −0.86) than the 16 published before 2012 [M (SD) = 2.81 (1.76)].

Evidence base

The 39 reviews (excluding the 5 scoping reviews) covered 280 studies, which were cited 536 times collectively. This corresponds to a 2% corrected covered area [536 − 280/(280 ∗ 39) − 280], indicating a 'slight overlap' (i.e., high variability) in the evidence base across reviews.[18] Consequently, none of these individual reviews provide a summary of the whole literature they refer to collectively. The overlap varied across different review types, and behavioral and technology foci (Supplementary table 4, ST4). For reviews of 'Multiple health behaviours' (ST4, Section 1), only the six narrative syntheses focusing on mobile technologies (i.e., text messaging) [IDs: 11, 14, 17, 20, 35, 42] had high overlap (12%), covering 17 studies, mostly randomized controlled trials (RCTs). The overlap was slightly attenuated when the meta-analysis on RCTs of 'Multiple health behaviours' (ID: 22) was included (29 studies, 44 citations, 9% overlap). For 'WM reviews' (ST4, Section 2), both narrative syntheses (IDs: 1, 5, 38, 40) and meta-analyses (IDs: 27, 30, 39) focusing on mobile technologies had high overlap (15%), analyzing 24 studies, mostly RCTs. The reviews on Web 2.0 technologies showed high overlap only when the narrative synthesis (ID: 12) was considered together with the meta-analyses (IDs: 4, 44), with divergence between the two meta-analyses. For 'PA reviews' (ST4, Section 3), only the group of narrative syntheses (IDs: 8-10, 34, 43) had high overlap (12%) based on 43 studies of various designs. The coverage area increased when the meta-analysis on PA (ID: 18) was included (43 studies, 74 citations, 14% overlap).

Clear conclusions about mobile and Web 2.0 technologies could not be drawn from the identified eHealth reviews due to variability in the technologies considered, heterogeneity in studies included (CCAs below 4%) and only a small proportion of the included studies focusing specifically on mobile and Web 2.0 interventions. We, therefore, concentrated on more focused reviews (26/44, 59%) that clearly reported on effects on weight-related outcomes, PA and/or dietary behaviors. The findings from these are summarized in table 2.

Effects of technologies on weight-related outcomes

Both the narrative syntheses (IDs: 1, 5, 38, 40) and meta-analyses (IDs: 27, 30, 39) that focused on WM specifically provided consistent, strong evidence for positive effects of mobile technologies on weight loss, with at least three-quarters of included studies showing significant weight loss (0.39–4.5 kg). As table 2 shows, the majority of the studies reported on SMS-based interventions. A narrative synthesis of high quality (ID: 5) and three meta-analyses (one high and two low quality) [IDs: 27, 30, 39] which included only RCTs, reported consistent evidence (10% overlap) supporting positive effects of mobile interventions on weight loss. The meta-analyses reported significant positive pooled effects across 24 studies (15% overlap), with weighted mean differences (WMD) of −1.09 kg (95% CI = −2.12 to −0.05, 6 RCTs) [ID: 27], −2.56 kg (95% CI = −3.46 to −1.65, 13 studies, both RCTs and quasi-experiments) [ID: 39], and an effect size for weight loss of $d = 0.43$ (95% CI = 0.25 to 0.61) across 11 RCTs (ID: 30) respectively. Sensitivity analyses in the first meta-analysis (ID: 27) showed larger effects when interventions utilizing modern smartphones (WMD = −1.78 kg; 95% CI = −2.92 to −0.63), as opposed to those using PDAs (WMD = −1.05 kg; 95% CI −2.20 to 0.10), were included in the pooled estimate. However, another meta-analysis (high quality) [ID: 22], which analysed mobile-based interventions addressing multiple health behaviors, reported non-significant changes in weight and other anthropometric outcomes across seven RCTs of interventions promoting calorie reduction and PA (SMD = −2.14 kg; 95% CI = −7.05 to 2.77), and three RCTs examining calorie reduction alone (SMD = −0.10 kg; 95% CI = −0.49 to 0.69).

Findings on the effects of Web 2.0 technologies on weight-related outcomes are mixed and the evidence limited despite building on a coherent set of RCTs (45 studies, 10% overlap), covered by one narrative synthesis (ID: 12) and two meta-analyses (IDs: 4, 44) that focused specifically on WM. The narrative synthesis of low quality (ID: 12), which included 20 RCTs, concluded that effects are unknown due to a lack of studies appropriately isolating social media from other web-based intervention components. One meta-analysis of low quality (ID: 4) included 12 RCTs and reported significant pooled effects on BMI, but the authors reported inconsistent measures of effects in the appendices (WMD vs. standardized mean difference) and there were no significant effects on weight (eight studies), waist circumference (five studies), or body fat (four studies). Also, the second meta-analysis of high quality (ID: 44) found no significant pooled effects on weight (10 RCTs), but all studies were rated as having 'unclear' or 'high risk' of bias, and mostly compared Web 2.0 interventions to other active interventions (e.g. websites). Large variability in effect sizes for weight was also reported in another meta-analysis, of low quality, on Web 2.0 technologies for multiple health behaviors (ID: 31): across three studies assessing weight-related outcomes, the effects were

Table 2 Summary of the evidence for WM, PA and diet

Behavior/Technology Focus	Type of review/ synthesis	No. reviews	Quality of reviews	Design/quality of included studies[c]	Size & overlap of evidence base	Findings and overall conclusions
Effects on weight-related outcomes						
Mobile	Narrative	4	One high, three low	One RCTs only; Two RCTs and quasi-experimentsOne Any study design Two assessed quality and majority judged high quality	27 studies CCA = 15%, high	Significant positive effects (e.g. between 1.5 and 4.5kg weight loss) reported in at least 3/4 of studies of SMS/ app-based interventions included in each review. One high quality review of RCTs, of which majority at low risk of bias, concluded that there is strong evidence for effectiveness.
	Meta-analysis	3	One high, two low	Two RCTs only; One RCTs and quasi-experiments All assessed study quality: 1 found low-moderate quality; 2 rated about half of studies as high quality	24 studies CCA = 15%, high	Significant, moderate-sized pooled effects (1.09-2.17kg weight loss) reported across all reviews, though most studies of SMS interventions. Larger effects from more modern technologies and an app reported in one high quality review, but RCTs included in this review were rated as low to moderate quality.
	All	7	Two high, five low	Three RCTs only; Three RCTs and quasi-experiments; One Any design; Five assessed study quality with variable findings	43 studies CCA = 10%, high	Strong evidence of effectiveness: Two high quality reviews of RCTs of variable methodological quality suggest positive effects, but findings mainly relate to SMS-based interventions; so need further high quality research on modern mobile technologies, such as mobile apps.
Web 2.0	Narrative	1	Low	RCTs only Study quality assessed as moderate-high quality	20 studies	Only one study isolated effects of social media, in others embedded in larger interventions, so difficult to draw any conclusions.
	Meta-analysis	2[a]	One high, one low	Both RCTs only Both assessed study quality: one rated all studies as unclear or high risk of bias; one rated half of studies as high quality	34 studies CCA = 0%, no overlap	In one review, significant positive pooled effects on BMI, but not on other weight-related outcomes. In other, only 2 of 22 studies showed significant effects on primary weight-related outcomes with 0.1 median effect size overall and non-significant pooled effects on weight
	All	3	Two low, one high	All RCTs only Two of three assessed study quality, with mixed findings	45 studies CCA=10%, high	Lack of evidence on effectiveness: Positive effects in one review on some BMI but not on weight and other anthropometric measures in this or another review; mixed findings may reflect failure to isolate social media components in many studies and active comparison groups.
Effects on PA						
Mobile	Narrative	5	All low	Two RCTs and quasi-experiments; Three Any design One assessed study quality as moderate-high	43 studies CCA=12%, high	Only one review of mostly SMS interventions found significant positive effects amongst the majority of included studies, in another isolating RCTs less than half showed significant effects.
	Meta-analysis	2[b]	Both low	One RCTs only; One any design Both assessed study quality and rated more than half of included studies as high quality	21 studies CCA = 10%, moderate	One review reported significant, moderate-sized pooled effect (g = 0.54) from a range of study designs examining SMS and app-based interventions. The second review reported generally positive effects from 5 studies, but did not report pooled estimates of the effects.
	All	6	All low	One RCTs only; Two RCTs and quasi-experiments; Four any design Three of seven assessed study quality, and rated this as moderate-high for most studies	43 studies CCA=14%, high	Weak evidence of effectiveness: Although most reviews reported some positive findings, the low quality of all these reviews, range of study designs included and lack of study quality assessment precludes any definitive conclusions about effects on PA

(continued)

Table 2 Continued

Behavior/Technology Focus	Type of review/ synthesis	No. reviews	Quality of reviews	Design/quality of included studies[c]	Size & overlap of evidence base	Findings and overall conclusions
Web 2.0	Meta-analysis	3[a]	All high	RCTs only All assessed study quality: two rated most studies as low risk of bias; one rated all studies as unclear or high risk of bias	33 studies CCA = 2%, slight	Lack of evidence on effectiveness: Reviews do not report effective results or use definitions for including studies that are not conclusive. One review reported no significant pooled effects, another reported significant pooled effects on fitness and PA, but this and the remaining review of only one RCT included other 'remote' interventions (e.g. telephone calls) preventing conclusions about social media, and highlighting the need for clearer definitions of Web 2.0 interventions.
Effects on diet						
Mobile	Meta-analysis	1[b]	Low	RCTs only Study quality assessment found half of studies at low risk of bias	12 studies	Weak evidence of effectiveness: Generally positive findings from small numbers of primary research studies (max 3) examining effects on fruit and vegetable and calorie intake, plus marginal effects in one study on sugar and fat intake. More review evidence of good quality is needed to establish effects on dietary behaviors.
Web 2.0	Meta-analysis	1[a]	High	RCTs only All studies were rated as unclear or high risk of bias	Five studies	Weak evidence of effectiveness: Significant, small decrease in fat consumption across five studies. More evidence from studies of good quality is needed to ascertain the effects of the technologies on dietary outcomes.

a: The meta-analysis by Williams et al. (ID = 44) was used to inform the results on weight-related, PA and diet outcomes.
b: The meta-analysis by Lyzwinski (ID = 30) was used to inform the results on weight-related, PA and diet outcomes.
c: Study design of included studies: 'RCTs only', systematic reviews that included only randomized controlled trials; 'RCTs & Quasi-experiments', systematic reviews that included pseudo-randomized, quasi-experiments, as defined by the authors themselves; 'Any design', systematic reviews that included primary research studies of any type of design, including case control, before-after, pilot trials, RCTs, uncontrolled CT.

negligible (0.00, 95% CI = −0.40 to 0.40), small (0.28, 95% CI = −0.17 to 0.73), and large (0.85, 95% CI = 0.13 to 1.56).

Effects of technologies on PA and dietary outcomes

There is evidence for generally positive effects on PA. For mobile technologies, five narrative syntheses (low quality) [IDs: 8–10, 34, 43] all reported positive effects in over half of 43 studies, with high overlap (CCA = 12%). This is confirmed by the results of a meta-analysis (ID: 18) showing overall positive pooled effects on PA for mobile-based interventions ($g = 0.54$; 95% CI = 0.17–0.91, 11 studies). However none of these reviews isolated the effects from RCTs.

For Web 2.0, findings from three meta-analyses of high quality (IDs: 21, 36, 44) are mixed. One meta-analysis (ID: 44) showed no significant changes in PA (SMD = 0.13; 95% CI = −0.04 to 0.30; 12 RCTs), in line with findings in a similar narrative synthesis (ID: 12). In contrast, two high quality meta-analyses (IDs: 21, 36) of 'remote and Web 2.0 interventions' found positive effects on cardiovascular fitness (SMD = 0.40; 95% CI = 0.04 to 0.76; 2 RCTs) and self-reported PA (SMD = 0.20; 95% CI = 0.11–0.28; 9 RCTs) compared to control conditions, but no significant differences compared to face-to-face interventions (SMD = −0.02; 95% CI = -0.30 to 0.26, only one RCT). However, these reviews had slight overlap (CCA = 2%), suggesting that they analysed independent sets of studies that were not always relevant to Web 2.0 technologies. In fact, 'telephone coaching' interventions were included as remote interventions alongside Web 2.0, which obviously do not rely on the Internet.

Diet was not a specific focus for any reviews and few reported effects of technologies on dietary behaviors. Hence, there is limited evidence on this. For mobile technologies, a meta-analysis of low quality that focused on WM (ID: 30) included 12 RCTs reporting positive effects on fruit and vegetable intake (2 out of 3 studies), and calorie intake (1 out of 2 studies) plus marginally significant reductions in sugar and fat intake in one included study ($P = 0.05$). Two studies measuring changes in energy-dense foods and eating behavior inventory scores also reported significant effects favoring mobile interventions. A further meta-analysis of high quality (ID: 44) reported significant decreases in dietary fat consumption across five RCTs employing Web 2.0 technologies (SMD = −0.35; 95% CI −0.68 to −0.02).

Discussion

This is the first systematic overview of reviews to assess the methodological quality of systematic reviews on mobile and Web 2.0 technologies for weight management and synthesize review evidence on effectiveness. We identified 44 'reliable systematic reviews',[27] published between 2007 and 2014, which reported on the use of these technologies for promoting WM and PA. The literature has predominantly assessed mobile-based interventions, with 24 out of 44 reviews focusing solely on mobile technologies, particularly text messaging (IDs: 10, 11, 14, 20, 35, 38, 39, 42, 43), with others also including mobile apps (IDs: 1, 2, 9, 22, 17). Only seven reviews covered Web 2.0 technologies (5 meta-analyses and 2 narrative syntheses). This is consistent with the evidence from primary research published in the past 10 years, which has mostly focused on mobile technologies, with social media being a source of data for content analyses rather than a medium for experimental interventions.[13] About a third of the reviews analysed interventions employing various 'eHealth' technologies, including computers, mobile and web-based approaches, reinforcing evidence for the superiority of interactive and web-based interventions over non-technology-based interventions.[11,12]

We assessed the methodological quality of the reviews using the AMSTAR checklist, which is a standardized instrument designed for this purpose.[29,30] Overall, quality was low, with only seven out of 44 (16%) reviews being of high quality: three meta-analyses (IDs: 21, 22, 36) and two narrative syntheses (IDs: 3, 5) of A quality, and two meta-analyses of B quality (IDs: 27, 44). Reviews published in the last 2 years had significantly higher AMSTAR scores than those previously available, with meta-analyses scoring higher than narrative syntheses and scoping reviews. Some items of the checklist [e.g. assessing risk of bias (A10) and the scientific quality of included studies (A7)] are more relevant to meta-analyses than scoping reviews or narrative syntheses. Indeed, none of these two types of reviews assessed risk of bias. Nevertheless, most low quality scores were due to authors not clarifying whether their work was based on a protocol (A1), not including 'grey literature' (A3), not including lists of both included and excluded studies (A5), or not declaring competing interests and sources of funding (A11). While the latter two might be lacking due to reporting constraints, the former are important elements to ensure research accountability and reproducibility, which should not be neglected even in scoping reviews. Low AMSTAR scores do not mean that a review has no value. Some low scoring reviews provided coherent and consistent coverage of the evidence base and were highly cited (IDs: 14, 20, 24, 33).

To summarize the research evidence we relied entirely on the information provided by authors of the systematic reviews, which were our primary unit of analysis. If the authors did not provide details on the quality, sample size, study design, significance and size of the effects, we could not provide an interpretation with respect to these features. We used citation matrices to estimate overlap among reviews and identify whether the evidence drawn on by similar reviews was homogeneous and coherent. These useful matrices[18] revealed that there is a homogeneous and consistent body of evidence supporting the effectiveness of mobile technologies for WM: highly overlapping narrative syntheses (IDs: 1, 5, 38, 40) and meta-analyses (IDs: 27, 30, 39), though low quality, overall reported positive effects from the generally high quality RCTs they included. Contrasting results on weight reported in one meta-analysis (ID: 22) might be explained by the differing focus of the review and different set of studies included.

In contrast, the effects of Web 2.0 on weight-related outcomes remain uncertain as findings are mixed. Even though the included primary research studies were all RCTs, a narrative synthesis of low quality (ID: 12) and a meta-analysis of low quality (ID: 4) reported no significant effects on weight, body fat and waist circumference, whereas a second meta-analysis of high quality (ID: 44) estimated small, but significant pooled effects only for BMI, but inconsistently reported the outcome (% BMI change) and effect sizes (WMD and mean differences) in the paper and in the supplementary materials, making the results unclear. The evidence on Web 2.0 is weak and inconclusive: more primary research is needed to appropriately isolate social media components from generic web-based interventions and reviews are needed to more accurately identify Web 2.0-based interventions.

There is suggestive evidence on the effectiveness of mobile technologies on PA, from five highly overlapping narrative syntheses (IDs: 8–10, 34, 43) and one meta-analysis (ID: 18). However results cannot be deemed definitive, as all these reviews included studies of different designs, used different outcomes, and few isolated the effects of technology. For Web 2.0 technologies, effects on PA also remain suggestive, as the three meta-analyses (IDs: 21, 36, 44), despite being of high quality, provided divergent results, drawing upon a non-overlapping evidence base. Future reviews should use clearer definitions of Web 2.0 technologies in order to identify and compare more homogeneous primary research studies.

Review evidence for effects on diet is limited for both mobile and Web 2.0 technologies and the findings vary for different dietary outcomes (e.g. fruit and vegetable, calorie, fat and sugar intake). More primary research studies and better quality reviews, on mobile technologies in particular, are therefore needed to

evaluate the effects of these technologies on a broad range of outcomes.

Reviews that reported on general eHealth technologies or that focused on Web 2.0 technologies provided a more heterogeneous picture of the literature. General eHealth reviews, especially those of multiple health behaviors, tended to include studies utilizing a wide range of different technologies and only few were relevant to the specific behaviors under study. Web 2.0 reviews had little or no overlap with each other and with the eHealth reviews, suggesting large variability in the evidence base. This is due to different definitions and inclusion criteria being used in eHealth and Web 2.0 reviews, regardless of their methodological quality. For example, the two Cochrane reviews of high quality (IDs: 21, 25) on both 'remote and Web 2.0 interventions' for PA conducted a meta-analysis among telephone-based studies, hence reporting the effects solely of 'remote' interventions that were not included in another meta-analysis on 'social media' (ID: 44). The former reviews did not use any technology-related keywords in their search strategies nor a clear definition of Web 2.0, such as Kaplan and Haenlein's definition,[6] which was used in the latter.

Reviews and meta-analyses on Web 2.0 technologies may also not identify significant findings because social media components are embedded in multi-component interventions (i.e., their effects were not isolated) or because reviews compared their effects with other active interventions. The effectiveness of mobile and Web 2.0 technologies depends on what these interventions are compared to; so greater attention to comparators in primary evaluations is needed. Clearer, accurate definitions and inclusion criteria are needed to improve the consistency of reporting of eHealth reviews in general. Some researchers have developed an eHealth-version of the CONSORT[37] statement for reporting RCTs (i.e., the CONSORT-EHEALTH[38]), but no equivalent currently exists for systematic reviews. Future research could be undertaken to apply this approach to the review literature, by developing eHealth versions of PRISMA or AMSTAR guidelines, for example, by conducting a consensus-seeking Delphi study to develop an equivalent standard for reporting systematic reviews in eHealth.

Strengths and limitations

A strength of this overview is the assessment of methodological quality combined with metrics to evaluate overlap in the evidence base of reviews, which is rarely undertaken in overviews.[18] Our approach allowed us to provide a comprehensive evaluation of the relevant review literature. A limitation is that we included only reviews published in English. We minimized bias in the search, selection, data extraction and analysis of data, but it is possible that some reviews were not retrieved by our search strategy. Another limitation relates to the AMSTAR checklist. Despite being reported as valid and reliable,[14,29,31] the use of dichotomous items make the evaluation imprecise; it also penalizes reviews that do not report items fundamental for the reproducibility of results (e.g. review protocol, comprehensive search strategies), to the same extent as reviews lacking details that might be due to space limitations or publication guidelines (e.g. conflict of interests, lists of included/excluded studies). Also, the assessment of publication bias, risk of bias and heterogeneity of studies do not apply to systematic reviews that do not perform meta-analyses. We excluded the publication bias item (A10) from the AMSTAR score for narrative syntheses and scoping reviews, but this did not improve the overall quality of these. Nonetheless clearer methodological assessments could be achieved by assigning different weights to more or less important quality criteria.

Conclusion

There is coherent evidence suggesting positive effects of mobile technologies on WM. Evidence is less conclusive with regard to using mobile technologies for promoting PA or diet, and is lacking with regard to using Web 2.0 technologies for promoting WM, PA or diet. Definitive conclusions are limited due to variability in the number, quality and design of studies included in these reviews, and to generally low methodological quality scores for the reviews themselves. The heterogeneity in studies included in the reviews results from wide variability in definitions and inclusion criteria for Web 2.0 technologies, which leads to conflicting results drawn from a limited number of studies. More and better quality primary research employing Web 2.0 technologies is also needed.

Better quality reviews, especially meta-analyses, using detailed data extraction techniques, clearer definitions of technologies, interventions and comparators are needed to advance our understanding of the effectiveness of these technologies, and identify directions for future research. Investing in the refinement of these technologies for WM and PA may be justified but future work should establish when and for whom mobile and Web 2.0 technologies may support effective behavior changes.

Supplementary data

Supplementary data are available at *EURPUB* online.

Acknowledgements

This overview was partially supported by a mobility grant awarded by the Swiss National Science Foundation (SNSF) to the first author (ref: 152290) and by the UK National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care of the South West Peninsula (PenCLAHRC). The views expressed in the paper are those of the authors and do not necessarily reflect those of the SNSF, NIHR or the UK Department of Health. These organizations had no role in the design of the study, in the collection, analysis, and interpretation of data, in the report writing and in the decision to submit the paper for publication. The authors wish to thank Dr. Helen Ryland for her valuable support in selecting the papers for the scoping review and for this paper, as well as the information specialist (Morwenna Rogers) and expert reviewer (Rebecca Whear) and the Evidence Synthesis team of the University of Exeter Medical School for their support in the development of the search strategy.

Funding

Swiss National Science Foundation (SNSF) to the first author (Mobility grant to the first author: ref: 152290); UK National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care of the South West Peninsula (PenCLAHRC).

Conflicts of interest: None declared.

Key points

- There is consistent and coherent evidence on the effectiveness of mobile technologies for supporting weight management.
- There is evidence which is suggestive of mobile technologies being effective in improving physical activity and diet.
- Evidence is lacking with regard to whether Web 2.0 technologies can support weight management, PA and diet.

- Good quality evidence from primary research studies and reviews is needed to determine which components of mobile- and Web 2.0-based interventions are associated with larger effects.
- Clearer terminologies and definitions for technology-based interventions and reviews are recommended to improve the quality and consistency of reporting.

References

1 Matheson GO, Klügl M, Engebretsen L, et al. Prevention and management of non-communicable disease: the IOC consensus statement, Lausanne 2013. *Br J Sports Med* 2013;47:1003–11.

2 NICE, National Institute for Health and Clinical Excellence. Preventing obesity and helping people to manage their weight [Internet]. NICE, National Institute for Health and Clinical Excellence; 2013. Available at: https://www.nice.org.uk/advice/lgb9/chapter/.

3 Dietz WH, Baur LA, Hall K, et al. Management of obesity: improvement of health-care training and systems for prevention and care. The Lancet [Internet]. 2015 Feb [cited 2015 Mar 21]; Available at: http://linkinghub.elsevier.com/retrieve/pii/S0140673614617487.

4 Eysenbach G. What is e-health? *J Med Internet Res* 2001;3:e20.

5 Neuhauser L, Kreps GL. eHealth communication and behavior change: promise and performance. *Soc Semiot* 2010;20:9–27.

6 Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. *Bus Horiz* 2010;53:59–68.

7 Kemp S. Digital, Social & Mobile Worldwide in 2015 [Internet]. We Are Social. 2015 [cited 2015 Jun 30]. Available at: http://wearesocial.net/blog/2015/01/digital-social-mobile-worldwide-2015/. (30 June 2015, date last accessed).

8 Campbell A, Choudhury T. From Smart to Cognitive Phones. *IEEE Pervasive Comput* 2012;11:7–11.

9 Chen Y-F, Madan J, Welton N, et al. Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: a systematic review and network meta-analysis. *Health Technol Assess Winch Engl* 2012;16:1–205.

10 Lau PWC, Lau EY, Wong DP, Ransdell L. A Systematic Review of Information and Communication Technology–Based Interventions for Promoting Physical Activity Behavior Change in Children and Adolescents. *J Med Internet Res* 2011;13:e48.

11 Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. *Obes Rev* 2010;11:306–21.

12 Wieland LS, Falzon L, Sciamanna CN, et al. Interactive computer-based interventions for weight loss or weight maintenance in overweight or obese people. *Cochrane Database Syst Rev* 2012 Aug 15;8:CD007675. doi: 10.1002/14651858.CD007675.pub2.

13 Bardus M, Smith J, Samaha L, Abraham C. Mobile Phone and Web 2.0 Technologies for Weight Management: A Systematic Scoping Review. *J Med Internet Res* 2015 Nov 16;17:e259.

14 Kohl LF, Crutzen R, de Vries NK. Online Prevention Aimed at Lifestyle Behaviors: A Systematic Review of Reviews. *J Med Internet Res* 2013 Jul 16;15:e146.

15 Elbert NJ, van Os-Medendorp H, van Renselaar W, et al. Effectiveness and Cost-Effectiveness of eHealth Interventions in Somatic Diseases: A Systematic Review of Systematic Reviews and Meta-Analyses. *J Med Internet Res* 2014 Apr 16;16:e110.

16 Tang J, Abraham C, Greaves C, Yates T. Self-Directed Interventions to Promote Weight Loss: A Systematic Review of Reviews. *J Med Internet Res* 2014 Feb 19;16:e58.

17 Pieper D, Buechter R, Jerinic P, Eikermann M. Overviews of reviews often have limited rigor: a systematic review. *J Clin Epidemiol* 2012 Dec;65:1267–73.

18 Pieper D, Antoine S-L, Mathes T, et al. Systematic review finds overlapping reviews were not mentioned in every other overview. *J Clin Epidemiol* 2014 Apr;67:368–75.

19 Bardus M, Smith J, Abraham C, Hillsdon M. Smartphone and Web 2.0 interventions for weight management. *PROSPERO* 2014; CRD42014010323.

20 Hartmann-Boyce J, Johns DJ, Jebb SA, Aveyard P, Behavioural Weight Management Review Group. Effect of behavioural techniques and delivery mode on effectiveness of weight management: systematic review, meta-analysis and meta-regression. *Obes Rev* 2014;15:598–609.

21 Bert F, Giacometti M, Gualano MR, Siliquini R. Smartphones and Health Promotion: A Review of the Evidence. *J Med Syst* 2014;38:9995.

22 Marcano-Belisario JS, Greenfield G, Huckvale K, et al. Smartphone and tablet self-management apps for asthma. *Cochrane Database Syst Rev* 2013 Nov 27;11:CD010013. doi: 10.1002/14651858.CD010013.pub2.

23 Bardus M. The Web 2.0 and Social Media Technologies for Pervasive Health Communication: Are they Effective? *Stud Commun Sci* 2011;11:119–36.

24 Chou WS, Prestin A, Lyons C, Wen K. Web 2.0 for health promotion: reviewing the current evidence. *Am J Public Health* 2013;103:e9–18.

25 Foster C, Richards J, Thorogood M, et al. Remote and web 2.0 interventions for promoting physical activity. *Cochrane Database Syst Rev* 2013 Sep 30;9:CD010395.

26 Free C, Phillips G, Watson L, et al. The Effectiveness of Mobile-Health Technologies to Improve Health Care Service Delivery Processes: A Systematic Review and Meta-Analysis. *PLoS Med* 2013;10:e1001363.

27 NICE, National Institute for Health and Clinical Excellence. NHS Evidence Process and Methods Manual. 2015.

28 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 2009;6:

29 Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 2009;62:1013–20.

30 Sharif MO, Janjua-Sharif FN, Sharif FNJ, et al. Systematic reviews explained: AMSTAR-how to tell the good from the bad and the ugly. *Oral Health Dent Manag* 2013;12:9–16.

31 Pieper D, Buechter RB, Li L, et al. Systematic review found AMSTAR, but not R(evised)-AMSTAR, to have good measurement properties. *J Clin Epidemiol* 2015;68:574–83.

32 Kung J, Chiappelli F, Cajulis OO. From systematic reviews to clinical recommendations for evidence-based health care: Validation of Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) for grading of clinical relevance. *Open Dent J* 2010;4:84–91.

33 Gwet KL. *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters*, 3rd edn. Gaithersburg, MD: Advanced Analytics, LLC; 2012. 294 p.

34 Neville L, O'Hara B, Milat A. Computer-tailored physical activity behavior change interventions targeting adults: a systematic review. *Int J Behav Nutr Phys Act* 2009;6:30.

35 Armstrong R, Hall BJ, Doyle J, Waters E. ''Scoping the scope'' of a cochrane review. *J Public Health*2011;33:147–50.

36 Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*2005;8:19–32.

37 Schulz K, Altman D, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010;8:18.

38 Eysenbach G. CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. *J Med Internet Res* 2011 Dec 31;13:e126. doi: 10.2196/jmir.1923.