

SYSTEMATIC REVIEW

Open Access



# Effects of Consumer-Wearable Activity Tracker-Based Programs on Objectively Measured Daily Physical Activity and Sedentary Behavior Among School-Aged Children: A Systematic Review and Meta-analysis

Carolina Casado-Robles<sup>1</sup> , Jesús Viciano<sup>1</sup> , Santiago Gujarro-Romero<sup>1</sup> and Daniel Mayorga-Vega<sup>2\*</sup>

## Abstract

**Background:** The popularity of consumer-wearable activity trackers has led the scientific community to conduct an increasing number of intervention studies integrating them to promote physical activity (PA) and to reduce sedentary behavior (SB) levels among school-aged children. Therefore, the aim of the present study was to estimate the effects of consumer-wearable activity tracker-based programs on daily objectively measured PA and SB among apparently healthy school-aged children, as well as to compare the influence of participants' and programs' characteristics.

**Methods:** Eligibility criteria were: (1) participants: apparently healthy school-aged children (< 18 years old); (2) intervention: aimed to promote PA and/or to reduce SB incorporating consumer-wearable activity trackers; (3) comparator: baseline measurements and/or a control/traditional group; (4) outcomes: objectively measured daily PA and/or SB levels; (5) study design: pre-experimental, quasi-experimental, and true-experimental trials. Relevant studies were searched from eight databases up to December 2020, as well as from four alternative modes of searching. Based on the Cochrane Risk-of-bias tool 2, the risk of bias was assessed following four domains: (1) randomization process; (2) missing outcome data; (3) measurement of the outcomes; and (4) selection of the reported results. Based on a comprehensive systematic review, meta-analyses of the Cohen's standardized mean difference ( $d$ ) and 95% confidence interval (CI) with a random-effects model were conducted to estimate the overall effects, as well as the within- and between-study subgroups analyses effects, of the programs on daily total steps, moderate-to-vigorous PA (MVPA), total PA and SB.

**Results:** Forty-four publications (i.e., 45 studies) were included in the systematic review (5,620 unique participants; mean age =  $12.85 \pm 2.84$  years) and 40 publications (i.e., 41 studies) in the meta-analysis. Programs had a mean length of  $11.78 \pm 13.17$  weeks and most used a waist-worn consumer-wearable activity tracker (77.78% waist-worn; 22.22% wrist-worn). Programs characteristics were: goal-setting strategies (64.06%); participants' logbooks (56.25%); counseling sessions (62.50%); reminders (28.13%); motivational strategies (42.19%); and exercise routine (17.19%). Results showed a statistically significant moderate favorable effect on daily total steps ( $d = 0.612$ , 95% CI 0.477–0.746), small

\*Correspondence: dmayorgavega@gmail.com

<sup>2</sup> Department of Didactic of Musical, Plastic and Corporal Expression, University of Jaen, Paraje de las Lagunillas, Campus de Las Lagunillas, Edificio Humanidades y Ciencias de la Educación I (D2-D136), 23071 Jaén, Spain

Full list of author information is available at the end of the article

favorable effect on daily MVPA ( $d = 0.220$ , 95% CI 0.134–0.307), trivial favorable effect on daily total PA ( $d = 0.151$ , 95% CI 0.038–0.264) and trivial unfavorable effect on daily SB ( $d = 0.172$ , 95% CI 0.039–0.305). Subgroups analyses showed a higher effect for daily total steps and daily MVPA levels in females and the physically inactive for daily total steps ( $p = 0.003$ – $0.044$ ). Programs with educational counseling and/or goal-setting strategies, as well as a greater number of strategies, were more effective for improving children's daily total steps, and wrist-worn activity trackers were more effective than waist-worn trackers for improving their daily MVPA levels ( $p = 0.001$ – $0.021$ ).

**Conclusions:** Consumer-wearable activity tracker-based programs seem to be effective in promoting school-aged children's daily total steps and MVPA levels, especially for females and those that are physically inactive. These programs should include specific goal-setting, educational counseling, and wrist-worn trackers as especially effective strategies. However, due to the certainty of evidence being from "low" to "moderate", future well-designed primary research studies about the topic are needed.

**PROSPERO:** CRD42020222363.

**Keywords:** Wearable activity tracker, Activity monitor, Fitness tracker, Behavior change, Self-monitoring, Intervention, Goal-setting, Children, Adolescents

## Key Points

- Consumer-wearable activity tracker-based programs improve school-aged children's daily total steps and moderate-to-vigorous physical activity levels, especially for females and the physically inactive.
- Goal-setting, educational counseling, and wrist-worn trackers have been highlighted as particularly useful strategies.
- Further research is needed to determine the effectiveness of consumer-wearable activity tracker-based programs using robust designs with low risk of bias and also to compare the effect of intervention characteristics in the same study.

## Background

Childhood and adolescence are considered to be sensitive periods of life as they are the stages in which healthy lifestyle behaviors can be formed and become established [1]. Moreover, these behaviors could influence future adult health status and behavior [2]. Two important lifestyle behaviors during the school-aged children's awake time are physical activity (PA) (i.e., any waking behavior consisting of bodily movement that requires energy expenditure, which can be categorized into a *continuum* from light to vigorous intensity) [3, 4], and sedentary behavior (SB) (i.e., any waking behavior characterized by a low energy expenditure) [4, 5]. Both PA and SB are independent key indicators of health and quality of life among school-aged children [4]. Specifically, regular moderate-to-vigorous PA (MVPA) is positively associated with several health-related markers in young people [4]. Moreover, school-aged children's daily total PA levels (i.e., light-to-vigorous PA) and daily total steps have also been demonstrated to be potential health indicators

[6]. Independently of daily PA levels, long periods of SB is considered a risky behavior associated with health problems in school-aged children [7]. The World Health Organization (WHO) [4] recommends that school-aged children achieve, on average, at least 60 min daily of MVPA across the week, and they also should limit the amount of time spent being sedentary. Unfortunately, nowadays worldwide most school-aged children (approximately 81%) are physically inactive (i.e., do not meet the PA recommendations) [8], and they spend most of their time in SB [9]. These levels are worrying because physical inactivity is one of the main risk factors for non-communicable disease and death worldwide, which makes the promotion of health-enhancing PA and SB levels (i.e., high levels of PA and low levels of SB) among young people a paramount priority public health challenge [10].

As a result of the high importance of this topic in public health, a large number of intervention programs have been carried out to promote school-aged children's health-enhancing PA and SB levels, which include a wide variety of behavior change techniques [11, 12]. Among others, self-monitoring behavior is an essential technique for PA practice promotion [13], and consumer-wearable activity trackers (e.g., smartwatches, activity wristbands, or pedometers) could be ideal devices to track school-aged children's PA and SB levels, providing them this important real-time feedback [14]. These consumer-wearable activity trackers are electronic devices worn on the body as an accessory used for monitoring and recording daily PA and fitness-related metrics and providing users with real-time behavioral feedback via the monitor display [14, 15]. They usually integrate an accelerometer or pedometer to track physical movements automatically. Their outputs are generally based on step counts, the amount and intensity of PA, energy expenditure, periods of inactivity, or heart rate [16, 17]. Furthermore,

these consumer-wearable activity trackers have become increasingly popular over recent years being reflected in a great increase in sales and strong demand from society worldwide every year [18]. Moreover, these devices are considered the most plausible activity monitors to be used in public health [19]. The popularity of these devices has led the scientific community to carry out effective studies that integrate consumer-wearable activity trackers promoting health-enhancing school-aged children's PA and SB levels [20, 21].

However, each of those effectiveness studies alone only constitutes a specific part of the total evidence about this topic. By contrast, systematic reviews and meta-analyses allow for an objective analysis of all the available evidence about a specific topic, making sense of the often conflicting results found and providing estimation with greater power and precision than each individual primary study [22, 23]. Specifically, the number of reviews focused on the use of wearable activity trackers to increase PA and reduce SB levels has grown exponentially in recent years. However, most of them have been carried out in adults [24–26] or clinical populations [16, 27], leaving little evidence on the effectiveness of wearable activity trackers in apparently healthy school-aged children.

To our knowledge, only Ridgers et al. [28], Böhm et al. [29], Cajita et al. [30], and Creaser et al. [31] reviews included intervention studies carried out with both apparently healthy school-aged children and those with diagnosed diseases (e.g., cancer or diabetes) for promoting PA with consumer-wearable activity trackers. However, those reviews have considerable limitations. Firstly, regarding inclusion criteria, the four systematic reviews included restrictions on the language of publication (only in English) and they also exclude some grey literature sources (e.g., conference abstracts, dissertations, or pilot studies) which could bias the results due to interventions with significantly larger effects being more likely to be published in English and journal papers [32–34]. Furthermore, the Böhm et al. [29] and Cajita et al. [30] reviews included date restrictions in the search (only from 2012 and 2009, respectively), and the Cajita et al. [30] review also presented a limitation regarding electronic databases selected, since they did not include main bibliographic databases (e.g., Web of Science, Scopus, or SPORTDiscus) [22, 34]. These limitations imply that the systematic reviews were not carried out considering all the scientific evidence, and therefore, the estimation of the effect is not completely good.

Moreover, these four systematic reviews included studies with self-reported measured PA levels (i.e., questionnaires). However, previous evidence has demonstrated that self-reported measures of children and adolescents are poorly correlated with objectively measured PA and SB

levels [35, 36]. Furthermore, the four systematic reviews also included studies measuring only part-time days (e.g., physical education lessons or school recess), which does not adequately reflect the effect on daily school-aged children's PA and SB levels. Lastly, both the Ridgers et al. [28] and Böhm et al. [29] systematic reviews did not include pedometer-based studies as they did not consider them as consumer-wearable activity trackers. However, following common consumer-wearable activity trackers' scientific definitions, pedometers should be included [15, 16, 37]. All the above-mentioned reasons could explain the limited evidence found (i.e., low number of publications) in their reviews, due to the inclusion of only five [28, 29] or two studies in their analyses [30]. Regarding the most recent Creaser et al.'s [31] review, it included more publications (i.e., 24 studies), but if only effectiveness studies with apparently healthy school-aged children and objectively measured daily PA and SB levels were considered, this would decrease to 11 studies. Moreover, it should be noted that Cajita et al. [30] conducted only a scoping review, whose objective was only to provide an overview of the available research evidence without producing a summary answer to a discrete research question [38]. Finally, as far as we know, there is no previous meta-analysis examining the effects of consumer-wearable activity tracker-based programs on daily objectively assessed PA and SB among school-aged children. Conducting a meta-analysis could summarize the effectiveness of those interventions in an overall statistical synthesis (effect size) rather than taking the results of each primary study separately. Therefore, it allows for improving the precision of the results by the estimation of the effect size and direction, and clarifying whether or not the effect size is consistent across studies. Thus, also based on study design and risk of bias assessment, the meta-analysis allows for assessing the strength of the evidence [34, 39].

Consequently, the main purpose of the present systematic review and meta-analysis was to estimate the effects of consumer-wearable activity tracker-based programs on daily objectively measured PA and SB among apparently healthy school-aged children. A secondary purpose was to compare the influence of school-aged children's characteristics (i.e., sex, age group, and PA status) and the intervention programs characteristics (i.e., duration, worn-type activity tracker, goal-setting, diary, counseling, reminders, motivational strategies, and exercise) on the effects of the consumer-wearable activity tracker-based programs on daily objectively measured PA and SB.

## Methods

### Protocol and Registration

The review protocol was registered with the International Prospective Register for Systematic Reviews

(PROSPERO, CRD42020222363, [https://www.crd.york.ac.uk/prospero/display\\_record.php?RecordID=222363](https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=222363)).

The present systematic review and meta-analysis was based on the methodology described in previous reference literature such as the PRISMA guidelines [40] (see Additional File 1 for the PRISMA checklist) and Cochrane Handbook for Systematic Reviews of Interventions [34], among other important references [22, 39]. Firstly, a reproducible, transparent, and comprehensive systematic review was performed to identify, select, and synthesize all the relevant studies. Then, a meta-analysis was performed to provide more precise estimates of the effects than those derived from the primary studies.

### Eligibility Criteria

The eligibility criteria for including the retrieved studies in the systematic review were the following: (1) participants: apparently healthy children and adolescents (i.e., populations with diagnosed diseases/conditions were excluded) aged from 5 to 18 years (although the study was also included if the participants were up to 24-years-old but the average age was under 18, as they are considered “young people” by the WHO [41]); (2) Intervention: studies that examined the effect of programs to promote PA and/or to reduce SB incorporating consumer-wearable activity trackers (i.e., pedometers, smartwatches, fitness wristbands, or similar; smartphones applications were not included) alone or combined with other strategies (i.e., goal-setting, diaries, counseling, reminders, motivational strategies, or exercise routine) with a duration of at least three weeks were included [42]; (3) Comparator: The design must include at least one measure to compare intervention effects (i.e., baseline measurements in single-group designs and/or a control group with no intervention or with usual treatment without consumer-wearable activity trackers). Thus, interventions with only one experimental group and only post-intervention measures were not included; (4) outcomes: studies that assessed the effect of the programs on the objectively measured daily PA and/or SB levels were included (i.e., objective measures such as the same consumer-wearable activity trackers used in the intervention or research-grade activity trackers such as accelerometers, excluding self-reported measures). Furthermore, only whole day (awake) time-based timeframes such as a whole week, weekdays, and/or weekend days were included (i.e., part-time days such as physical education lessons, school recess, or leisure time were not included); (5) study design: Any kind of experimental designs including, but not limited to, pre-experimental trials (e.g., non-controlled trials with one-group pre–post-intervention design); quasi-experimental trials (e.g., cluster-randomized controlled trials or non-randomized controlled

trials), and true-experimental trials (e.g., randomized controlled trials) were included.

### Data Sources and Search Strategy

The databases search following the search strategies and the download and collection in the reference manager was completed in December 2020. This search included the following eight electronic bibliographic databases: Web of Science™ (all databases), Scopus, PubMed, SPORTDiscus with Full Text, CINAHL, Cochrane Library, ProQuest Social Sciences Premium Collection, and ProQuest Dissertations & Theses Global™. The searches were carried out in the search field type “title, abstract, and keywords” or equivalent. The search terms used were based on three concepts: (1) consumer-wearable activity tracker; (2) intervention program, and (3) PA/SB. The terms of the same concept were combined with the Boolean operator “OR” and then the three concepts were combined using the Boolean operator “AND.” The keywords with more than one word were enclosed in quotes. No publication status, language, or date restrictions were imposed [22]. All search strategies are available in Additional File 2.

Then, additional studies were identified as follows (i.e., “snowballing”): (1) searching the reference lists of original studies, as well as some related study reviews and study protocols; (2) examining the reference citations and the researchers’ publications (first authors) in the Web of Science™ and Scopus databases; (3) contacting with the corresponding authors by email, and (4) screening the researchers’ personal lists in ResearchGate and Google Scholar (first authors). Any time a new study was found, all of these modes of searching were repeated until no new study appeared.

### Study Selection

After eliminating duplicates, the first author (CCR) undertook the potentially eligible records selection based on the screening of titles and abstracts to identify relevant studies. After that, two independent reviewers assessed the full texts for inclusion following the above-mentioned eligibility criteria (CCR/SGJ). Any disagreements regarding the inclusion of studies were resolved by consensus with a third reviewer (DMV). The inter-rater agreement between coders was substantial to almost perfect (proportion of agreement = 0.89; Cohen’s Kappa = 0.78).

### Data Extraction

From each selected study, data were coded using an ad hoc coding form developed by the research group and previously tested with a pilot sample of studies at the beginning of the review. This form included data about: (1) study characteristics (i.e., reference, publication date,

date of the data collection, study design, sequence generation, suspicion of selective outcomes, and initial and final sample size); (2) participant characteristics (i.e., sex and age); (3) outcome measures pertaining to PA and/or SB (i.e., measurement moment, measurement time, kind of measurement instrument, and measurement score and units); (4) intervention characteristics (intervention length, kind of consumer-wearable activity tracker, kind of goal-setting, diary, counseling, reminders, motivational strategies, and exercise routine); and (5) results of the intervention for each group (i.e., initial and final group size, pre- and post-intervention standard deviation, and pre- and post-intervention means score or pre-post-intervention mean difference score). Complete coding form is available in Additional File 3.

If a study consisted of two or more study arms of which one of the intervention arms did not meet the inclusion criteria, data were only extracted from the study arms that met the inclusion criteria. In the event that the studies did not report some study feature, corresponding authors were contacted to retrieve it. If means and standard deviation were not retrieved, the scores were estimated and converted by the standard error, confidence intervals, *F*, *t* or *p* values [34]. Since median and interquartile range are often used when the data are asymmetrical, these values were not converted [34]. If any other study feature was not retrieved, the information was omitted. The sample size of each group, the mean scores of the pre- and/or post-intervention or mean difference scores of each group, and the measurement score of the dependent variable/s were considered to be critical for including the selected studies from the systematic review in the meta-analysis. In order to avoid removing studies from the meta-analysis, numerical data were extracted from their figures using the WebPlotDigitizer software [34], as was done with three studies [43–45]. Coding studies were carried out independently by two researchers (CCR/DMV). When doubt or disagreement occurred, a consensus was always achieved through discussion.

### Risk of Bias and Certainty of the Evidence

Based on the Cochrane risk-of-bias tool version 2 [34], the following methodological domains were assessed: (1) risk of bias arising from the randomization process; (2) risk of bias due to missing outcome data; (3) risk of bias in measurement of the outcomes, and (4) risk of bias in the selection of the reported results. Due to the nature of the selected studies (i.e., self-monitoring interventions to promote healthy habits with objective measurements), some risk of bias criteria were not considered. Firstly, blinding of participants and personnel was not included due to participants being always aware of their assigned intervention (i.e., they

were using a consumer-wearable activity tracker), and researchers delivering the interventions were aware of the assigned intervention because they had to implement behavior change techniques in experimental groups. Therefore, the assessment of allocation concealment was also meaningless in this type of study since both participants and researchers would know which group they were assigned to during the intervention phase. Finally, blinding the outcomes assessment was not applicable because an objective method (i.e., research-grade accelerometers or consumer-wearable activity trackers) was used instead of subjective assessors or self-reported measures. Additional File 4 shows the algorithms followed for assessing methodological risk of bias in each domain.

Domains were judged and classified as “low risk,” “some concerns” or “high risk” of bias. Finally, overall risk of bias judgment was obtained as follows: “low risk” if low risk of bias was obtained for all domains; “some concerns” if at least one domain was judged as having some concerns, but not to be at high risk of bias for any domain; and “high risk” if at least one domain was judged as high risk or if two or more domains were judged as having some concerns [34].

Additionally, the overall certainty of the evidence was rated as “high,” “moderate,” “low,” or “very low” using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach [46]. This assessment was based on the following five domains: risk of bias, inconsistency, imprecision, indirectness, and publication bias. A domain was classified as “no limitation” if no reason for downgrading the evidence was found, but the domain was classified as “serious” if a reason was found for downgrading the evidence (thus, downgrading the certainty rating by one level). The reasons for considering the domains as “serious” were: (a) risk of bias: Most of the studies (i.e., >50%) were classified as “high” risk of bias; (b) inconsistency: high level of heterogeneity (i.e.,  $I^2 > 75\%$ ) was found; (c) imprecision: The confidence interval was wide including both the possibility of trivial effect (i.e.,  $d < 0.20$ ) and large effect (i.e.,  $d \geq 0.80$ ); (d) indirectness: Most studies (i.e., >50%) addressed a restricted version of the main review question in terms of population, intervention, comparator or outcomes; and (e) publication bias: Egger’s test was statistically significant and the impact of publication bias was large (i.e., the number of additional studies with no effect that would be needed to increase the *p* value over a statistically insignificant effect (i.e.,  $p > 0.05$ ) was low following the fail-safe *N* analyses; or Trim and Fill method trimmed many studies with an adjusted value for effect size different to the observed values).

### Data Analyses

The meta-analyses were performed using the software Comprehensive Meta-Analysis version 3.3.070 for Windows (Biostat, Englewood, USA). This software allows multiple data entry formats according to each study design. Therefore, considering the data provided by each study, the input method that best suited each study was chosen. Specifically, the option “one group with pre- and post-data” was used for pre-experimental trials, while the options “unmatched groups with pre- and post-data in each group,” “unmatched groups, mean change in each group,” or “unmatched groups, post-data only in each group” were used for controlled trials. The significance level was set at  $p < 0.05$ .

If a single study reported data for the whole sample and separately by different subsamples (e.g., children and adolescents), only the whole sample was used. Moreover, when in the same study there were different options for the same outcome (e.g., data from consumer-wearable activity trackers and research-grade activity trackers, or different measurement moments during the program) only the best option was selected (e.g., research-grade activity tracker or the last measurement as post-intervention). When a study had more than one activity tracker-based intervention group, each group was included in the analysis individually. The studies carried out with a small sample (defined as less than 10 participants per group) were not included in the meta-analysis [47].

### Effects Sizes Computation

A detailed description of the data analyses carried out in the present meta-analysis can be found elsewhere [39]. Meta-analyses of the Cohen's  $d$  standardized mean difference and 95% confidence interval with a random-effects model were conducted to obtain the intervention program effects on: (a) daily total steps; (b) MVPA; (c) total PA; and (d) SB. Based on Cohen's [48] benchmarks, effect size values of  $d < 0.20$  were considered as trivial, 0.20–0.49 as small, 0.50–0.79 as moderate, and  $\geq 0.80$  as large. Moreover, since daily total steps and MVPA in minutes work with meaningful scales, the Mean Difference ( $D$ ) with a random-effects model was also conducted [34, 39]. Regarding total PA and SB,  $D$  analysis could not be conducted due to the low number of studies using the same scales within each variable. Positive  $d$  and  $D$  values mean that the program favorably increased participants' daily PA, but it also means an unfavorable increase in their SB levels.

### Publication Bias

Firstly, an exhaustive systematic review was carried out to avoid availability bias. Afterward, a deep examination of the selected studies was carried out to avoid

any potential duplication of the information retrieved. Similarities between publications of the same authors, with the same values and/or the same sample size were examined. When the selected publications had full or partial duplicated information, these particular values were not analyzed. Then, to visually identify the impact of any potential publication bias, the funnel plots and the Egger's test [49] were carried out for daily total steps and MVPA. Moreover, for assessing the impact of any potential publication bias, the Orwin's fail-safe  $N$  analyses [50] (criterion for a “trivial”  $d = 0.10$ ; mean  $d$  in missing studies,  $d = 0.00$ ) [48], the Duval and Tweedie's Trim and Fill method [51] (assuming missing studies in the left of the mean), and a cumulative meta-analysis sorted by larger study size were computed [39]. Due to the limited number of studies found for total PA and SB ( $k = 8$ ), the publication bias analyses could not be carried out for these variables [47].

### Heterogeneity and Subgroups Analyses

The presence of statistical heterogeneity in the estimation of the effect sizes of the program was examined with the  $I^2$  statistic. The thresholds for its interpretation were: Values up to 40% were considered not important, up to 75% moderate, and more than 75% high heterogeneity [34].

Based on a priori hypothesized moderators, subgroups analyses were also carried out to test the effect of the intervention regarding: (a) individuals' characteristics (i.e., sex, age, and PA status); and (b) intervention program characteristics (i.e., duration, activity tracker type, goal-setting, kind of goal-setting, diary, counseling, reminders, motivational strategies, exercise routine, and number of behavior change strategies). Due to the limited number of studies found for total PA ( $k = 8$ ) and SB ( $k = 8$ ), the above-mentioned subgroups analyses were only carried out for the daily steps and MVPA variables following a partial hierarchical analysis approach. All subgroups analyses were carried out for between-study meta-analysis, while for within-study meta-analysis, only those with at least two units of analysis to compare with were performed. Finally, the influence of continuous covariates (i.e., age, PA status, duration, and number of behavior change strategies) on the intervention effect was also evaluated using meta-regression analyses.

### Sensitivity Analysis

Finally, in order to evaluate the robustness of the main results, the following sensitivity analyses were performed: Cohen's  $d$  with a fixed-effect model, Hedges'  $g$  with a random-effects model, and Cohen's  $d$  with a random-effects model separately for randomized controlled trial design or not. However, sensitivity analyses separately for studies classified by the overall risk of bias could not

be carried out due to the low number of studies classified as “Low risk” ( $k=1$  for daily total steps; and  $k=2$  for MVPA, total PA, and SB).

## Results

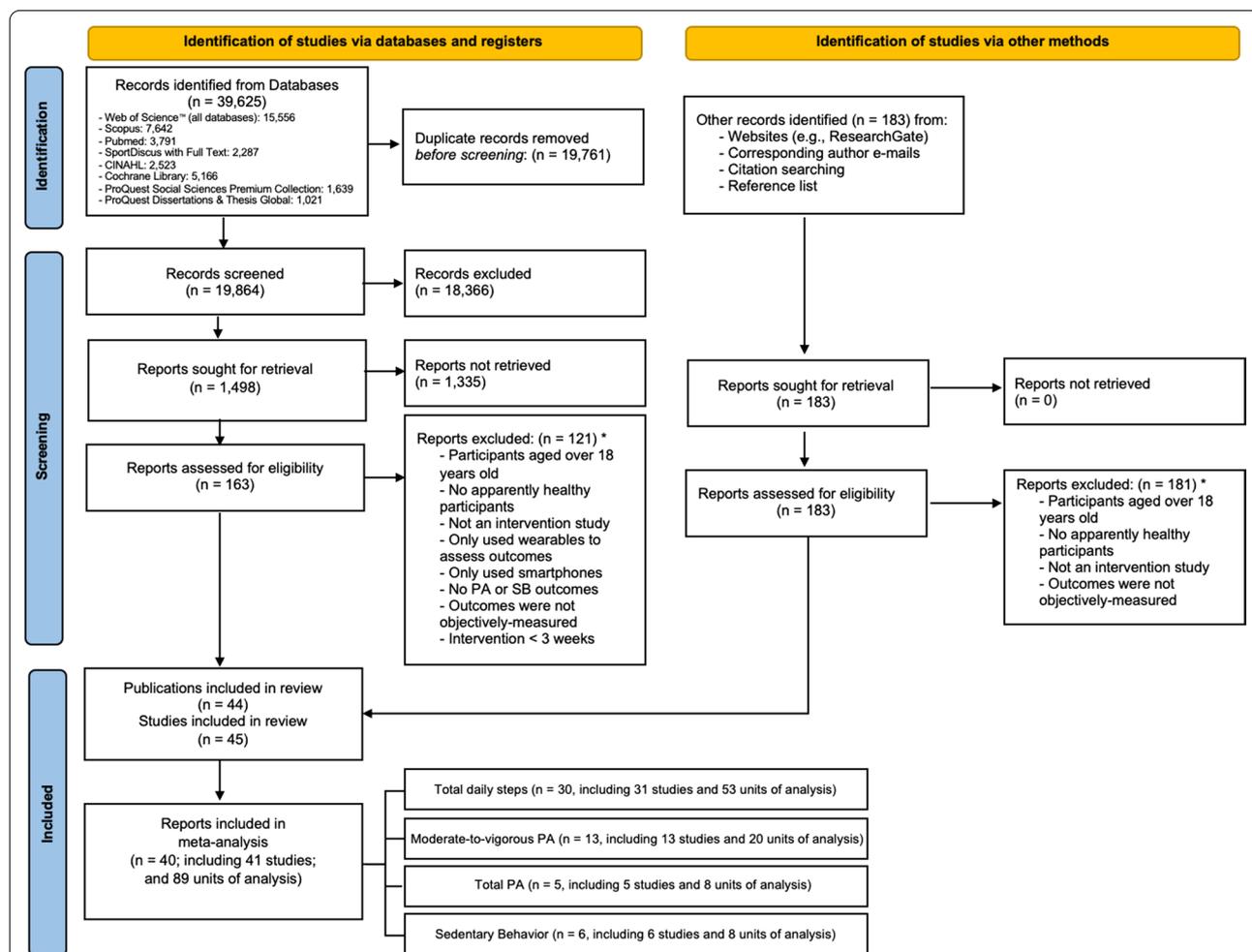
### Study Selection

Figure 1 shows the flow diagram of the study selection process. The search strategy identified 39,625 potentially relevant studies (19,864 studies after removing duplicates). Afterward, as a result of the studies of the Boolean-based database search, 183 additional records were identified through other sources. From the 1,498 records retrieved for a more detailed evaluation, 47 publications were a priori selected for meeting the selection criteria. However, only 44 unique publications (i.e., 45 studies) were finally included in the systematic review to avoid duplicated information. Finally, due to carrying out

the study with a small sample and/or not reporting the critical values, only 40 publications (i.e., 41 studies) were included in the meta-analysis.

### Study Characteristics

Table 1 presents a summary of study characteristics included in the systematic review. Across the 44 included publications (i.e., 45 studies), in total 5,620 unique participants were included in all studies (3,914=intervention; 1,706=control). Samples were composed of a median of 84.50 initial participants, ranging from 20 [52] to 496 [53]. The participants’ mean age was  $12.85 \pm 2.84$  years (range: 5–18 years). Regarding sex, 32 studies were conducted with males and females, while eight studies included only females and five studies included only males.



**Fig. 1** Flow diagram of the study selection process. PA = physical activity. Note The sum of publications included in the meta-analysis for each variable is greater than 40 because some studies reported outcomes for different variables. \*The number of reports excluded for each reason has not been included because a report could have been removed for only one reason, but at the same time could meet many other reasons for exclusion, and therefore including them would not really give accurate information on the reports removed

**Table 1** Characteristics of included studies in the systematic review

Study	Participants characteristics			Study design			Intervention characteristics					Outcome parameter (instrument) <sup>e</sup>	
	Number of participants <sup>a</sup>	Age <sup>b</sup>	Sex <sup>c</sup>	Tracker	Duration <sup>d</sup>	Group	Goal-setting	Diary	Counseling	Reminder	Motivation		Exercise
Aksornsrri [63] <sup>f</sup>	EG1: 40 (39)	11.4 (8–13)	Both	Non-CT	Wrist	4	EG1	Static	No	Yes	Yes	Yes	Steps (wearable)
Baldursdóttir [64] <sup>f</sup>	CG: 57 (54); EG1: 61 (57); EG2: 58 (57)	16.0 (15–16)	Both	CRCT	Hip	3	EG1	Adaptive	Yes	Yes	No	No	Steps (wearable)
Baldursdóttir et al. [54] <sup>f</sup>	CG: 46 (27); EG1: 38 (26)	16.0 (16–16)	Both	CRCT	Hip	3	EG1	Adaptive	No	Yes	No	No	Steps (wearable)
Bronikowski et al. [56] <sup>f</sup>	EG1: 100 (100); EG2: 93 (93)	14.8 (10–17)	Both	Non-CT	Wrist	8	EG1	Static	No	No	No	No	Steps (wearable)
Corepal et al. [65] <sup>f</sup>	CG: 82 (64); EG1: 142 (126)	13.0 (12–14)	Both	CRCT	Hip	22	EG1	Adaptive	No	No	No	No	MVPA; SB; Steps; and Total PA (accelerometer)
Corr et al. [66] <sup>f</sup>	EG1: 31 (17)	14.0 (12–16)	Females	Non-CT	Hip	6	EG1	Adaptive	Yes	Yes	Yes	Yes	Steps (wearable)
Corr and Murtagh [67] <sup>f</sup>	EG1: 31 (10)	16.0 (15–17)	Females	Non-CT	Hip	6	EG1	Adaptive	Yes	Yes	No	Yes	Steps (wearable)
Duck et al. [68] <sup>f</sup>	CG: 18 (14); EG1: 17 (13)	9.2 (9–10)	Both	CRCT	Wrist	10	EG1	No	No	No	Yes	No	MVPA (accelerometer)
Duncan et al. [69] <sup>f</sup>	EG1: 59 (59)	10.7 (10–11)	Both	Non-CT	Hip	4	EG1	Static	Yes	Yes	No	No	Steps (wearable)
Ermetici et al. [55] <sup>f</sup>	EG1: 181 (167)	12.5 (11–15)	Both	Non-CT	Hip	84	EG1	No	No	Yes	Yes	No	Steps (wearable)
Evans et al. [70] <sup>f</sup> —Study 1	EG1: 32 (32)	10.0 (10–11)	Both	Non-CT	Hip	4	EG1	Static	Yes	Yes	Yes	No	Steps (wearable)
Evans et al. [70] <sup>f</sup> —Study 2	CG: 10 (10); EG1: 19 (19); EG2: 13 (13)	12.3 (11–12)	Both	Non-RCT	Wrist	6	EG1	Static	Yes	Yes	No	No	Steps and MVPA (accelerometer)
Eyre et al. [60] <sup>f</sup>	CG: 40 (30); EG1: 94 (55)	10.0 (9–12)	Both	CRCT	Hip	6	EG1	Static	No	No	No	No	Steps (wearable)
Finkelstein et al. [71] <sup>f</sup>	CG: 138 (89); EG1: 147 (145)	8.2 (6–12)	Both	CRCT	Hip	36	EG1	Static	No	Yes	No	Yes	Steps (wearable)
Galy et al. [72] <sup>f</sup>	EG1: 24 (15)	11.9 (12–14)	Both	Non-CT	Wrist	4	EG1	No	No	Yes	No	Yes	Steps (wearable)
Gaudet et al. [43] <sup>f</sup>	CG: 23 (16); EG1: 23 (16)	13.0 (13–14)	Both	CRCT	Wrist	7	EG1	No	No	No	No	No	MVPA and SB (accelerometer); Steps and MVPA (wearable)

**Table 1** (continued)

Study	Participants characteristics			Study design			Intervention characteristics					Outcome parameter (instrument) <sup>e</sup>	
	Number of participants <sup>a</sup>	Age <sup>b</sup>	Sex <sup>c</sup>	Tracker	Duration <sup>d</sup>	Group	Goal-setting	Diary	Counseling	Reminder	Motivation		Exercise
Grao-Cruces et al. [73] <sup>f</sup>	EG1: 142 (66)	11.4 (11–12)	Both	Non-CT	Hip	EG1	Static	Yes	No	No	Yes	No	Steps (wearable)
Groffik et al. [74] <sup>f</sup>	EG1: 71 (64)	17.9 (17–18)	Both	Non-CT	Hip	EG1	No	Yes	No	No	No	No	Steps (wearable)
Guagliano et al. [21] <sup>f</sup>	CG: 25 (23); EG1: 24 (18); EG2: 24 (15)	9.3 (7–11)	Both	CRCT	Hip	EG1	No	No	Yes	No	Yes	No	MVPA and SB (accelerometer)
Hardman et al. [75]	CG: 18 (15); EG1: 14 (14)	10.6 (10–11)	Females	CRCT	Hip	EG1	Adaptive	Yes	Yes	Yes	Yes	No	Steps (wearable)
Hardman et al. [76] <sup>f</sup>	EG1: 81 (51); EG2: 119 (67); EG3: 186 (118)	9.0 (7–11)	Both	Non-CT	Hip	EG1	No	No	No	No	No	No	Steps (wearable)
Horne et al. [44] <sup>f</sup>	CG: 53 (51); EG1: 47 (38)	10.0 (9–11)	Both	CRCT	Hip	EG1	Adaptive	Yes	Yes	Yes	Yes	No	Steps (wearable)
Jago et al. [62] <sup>f</sup> —Phase 1	CG: 55 (51); EG1: 65 (64)	13.0 (10–14)	Males	CRCT	Hip	EG1	Static	No	Yes	No	No	Yes	SB, MVPA; and Total PA (accelerometer)
Jago et al. [62] <sup>f</sup> —Phase 2	CG: 130 (88); EG1: 103 (86)												
Jauho et al. [77]	CG: 139 (38); EG1: 137 (25)	17.9 (17–18)	Males	RCT	Wrist	EG1	No	No	No	No	No	No	MVPA and SB (wearable)
Kantanista et al. [78] <sup>f</sup>	EG1: 26 (26); EG2: 56 (56)	17.2 (16–18)	Females	Non-CT	Hip	EG1	No	No	No	No	No	No	Steps (wearable)
Kerner et al. [79] <sup>f</sup>	EG1: 62 (28)	14.5 (14–15)	Both	Non-CT	Wrist	EG1	Adaptive	No	No	No	No	No	MVPA (accelerometer)
Larsen et al. [80] <sup>f</sup>	EG1: 21 (19)	14.7 (12–18)	Females	Non-CT	Hip	EG1	No	Yes	Yes	Yes	Yes	No	MVPA (accelerometer)
Leinonen et al. [53] <sup>f</sup>	CG: 246 (80); EG1: 250 (87)	17.8 (17–18)	Males	RCT	Wrist	EG1	Static	No	Yes	No	Yes	No	MVPA and SB (wearable)
Linck [81] <sup>f</sup>	CG: 22 (17); EG1: 22 (18)	16.6 (14–18)	Females	RCT	Hip	EG1	Adaptive	Yes	No	No	No	No	Steps (wearable)
Lubans and Morgan [82] <sup>f</sup>	CG: 66 (52); EG1: 50 (45)	14.2 (14–15)	Both	CRCT	Hip	EG1	Adaptive	Yes	Yes	No	Yes	Yes	Steps (wearable)

**Table 1** (continued)

Study	Participants characteristics			Study design			Intervention characteristics					Outcome parameter (instrument) <sup>e</sup>	
	Number of participants <sup>a</sup>	Age <sup>b</sup>	Sex <sup>c</sup>	Tracker	Duration <sup>d</sup>	Group	Goal-setting	Diary	Counseling	Reminder	Motivation		Exercise
Lubans et al. [83] <sup>f</sup>	CG: 66 (52); EG1: 58 (50)	14.1 (13–18)	Both	CRCT	24	EG1	Adaptive	Yes	Yes	Yes	Yes	Yes	Steps (wearable)
Lubans et al. [61] <sup>f</sup>	CG: 50 (50); EG1: 50 (50)	14.3 (14–15)	Mailes	CRCT	24	EG1	Adaptive	Yes	Yes	No	No	Yes	Steps (wearable)
Macias-Cervantes et al. [84]	CG: 38 (30); EG1: 38 (32)	7.8 (6–9)	Both	RCT	12	EG1	Adaptive	No	Yes	No	No	Yes	Steps (wearable)
Morris et al. [85] <sup>f</sup>	CG: 72 (31); EG1: 82 (52)	9.9 (9–10)	Both	CRCT	6	EG1	Adaptive	No	No	No	No	No	MVPA; SB; and Total PA (accelerometer)
Newton et al. [86] <sup>f</sup>	EG1: 14 (14); EG2: 13 (13)	8.7 (6–10)	Both	Non-CT	12	EG1	Adaptive	No	No	No	Yes	No	Steps (wearable)
Remmert et al. [52]	EG1: 10 (6); EG2: 10 (9)	12 (12–12)	Both	Non-CT	12	EG1	No	No	Yes	Yes	Yes	No	MVPA (accelerometer); Steps and MVPA (wearable)
Routen et al. [45] <sup>f</sup>	EG1: 17 (14); EG2: 26 (20); EG3: 25 (16)	11.2 (11–12)	Both	Non-CT	3	EG1	No	No	No	No	No	Yes	MVPA and Total PA (accelerometer)
Schofield et al. [87] <sup>f</sup>	CG: 30 (24); EG1: 27 (23); EG2: 28 (21)	15.8 (15–18)	Females	CRCT	12	EG1	Adaptive	Yes	Yes	Yes	No	No	Steps (wearable)
Shapiro et al. [88] <sup>f</sup>	EG1: 18 (13); EG2: 18 (7)	8.7 (5–13)	Both	Non-CT	8	EG1	Adaptive	Yes	Yes	Yes	No	No	Steps (wearable)
Shimon and Petlichkoff [89] <sup>f</sup>	CG: 62 (36); EG1: 72 (43); EG2: 60 (34)	13.1 (12–14)	Both	CRCT	4	EG1	Adaptive	Yes	Yes	No	No	No	Steps (wearable)
Shore et al. [90] <sup>f</sup>	EG1: 57 (46); EG2: 56 (46)	11.9 (11–12)	Both	Non-CT	6	EG1	Static	Yes	Yes	Yes	No	No	Steps (wearable)
						EG2	No	Yes	No	No	No	No	
						EG1	Static	Yes	Yes	Yes	Yes	Yes	Steps (wearable)
						EG2	No	Yes	No	No	No	No	
						EG1	Static	Yes	Yes	Yes	Yes	Yes	Steps (wearable)
						EG2	No	Yes	No	No	No	No	

**Table 1** (continued)

Study	Participants characteristics			Study design			Intervention characteristics						Outcome parameter (instrument) <sup>e</sup>
	Number of participants <sup>a</sup>	Age <sup>b</sup>	Sex <sup>c</sup>	Tracker	Duration <sup>d</sup>	Group	Goal-setting	Diary	Counseling	Reminder	Motivation	Exercise	
Smith et al. [91] <sup>f</sup>	CG: 180 (154); EG1: 181 (139)	12.7 (12–14)	Males	CRCT		EG1	Adaptive	Yes	Yes	Yes	Yes	Yes	Total PA and MVPA (accelerometer)
Thompson et al. [92] <sup>f</sup>	CG: 40 (34); EG1: 40 (36); EG2: 40 (31); EG3: 40 (37)	15.2 (14–17)	Both	RCT		EG1	No	No	Yes	No	No	No	MVPA and Steps (accelerometer)
Wang [93] <sup>f</sup>	EG1: 34 (24); EG2: 32 (22)	13.5 (13–14)	Females	Non-CT		EG1	Adaptive	Yes	Yes	No	No	No	Steps (wearable)
Zizzi et al. [94] <sup>f</sup>	EG1: 84 (68); EG2: 81 (60)	16 (14–18)	Both	Non-CT		EG1	No	Yes	Yes	No	No	No	Steps (wearable)
						EG2	No	Yes	No	No	No	No	

CG = control group; EG = experimental group; CT = controlled trial; CRCT = cluster randomized controlled trial; RCT = randomized controlled trial; PA = physical activity; MVPA = moderate-to-vigorous physical activity; SB = sedentary behavior.

<sup>a</sup> Number of participants are reported as "group: initial number (final number)"

<sup>b</sup> Age is reported as "mean (minimum–maximum)" in years

<sup>c</sup> Both means that includes males and females

<sup>d</sup> Duration is reported in number of weeks

<sup>e</sup> Wearable correspond to the same wrist-worn or hip-worn consumer-wearable activity trackers used in the intervention. Accelerometer correspond to research-grade activity trackers

<sup>f</sup> Studies included in the meta-analysis

Regarding the design of the studies, five were true-experimental trials (i.e., a randomized controlled trial design), 19 were quasi-experimental trials (i.e., a cluster-randomized controlled design or non-randomized controlled design), and 21 were pre-experimental trials (i.e., non-controlled design). Regarding the measurement instrument, 27 used a hip-worn consumer-wearable activity tracker, five a wrist-worn consumer-wearable activity tracker, and 11 a research-grade activity tracker, while two studies reported both wrist-worn consumer-wearable activity tracker and research-grade activity tracker measures.

The consumer-wearable activity tracker-based programs included had a mean length of  $11.78 \pm 13.17$  weeks, but varied considerably from three [54] to 84 weeks [55]. Specifically, 44.44% of the included studies lasted more than eight weeks. Most of them used a waist-worn consumer-wearable activity tracker as a motivational tool in the program ( $n=35$  waist-worn; and  $n=10$  wrist-worn). Regarding the intervention characteristics, from the 64 experimental groups formed (along with the 45 studies), they included: A goal-setting strategy [20.31% (i.e., 13) and 43.75% (i.e., 28) of experimental groups used a static goal and an adaptive goal, respectively]; participants' logbooks such as any diary in which school-aged children filled in their daily steps or minutes involved in PA, or any mobile app in which they synchronized their daily PA data and checked their progress [56.25% (i.e., 36) of experimental groups]; educational counseling sessions about PA-related knowledge, like benefits of regular PA practice, international guidelines, or how to overcome barriers to PA practice [62.50% (i.e., 40) of experimental groups]; reminders to persuade participants to move [28.13% (i.e., 18) of experimental groups]; motivational strategies such as the support of social networks like Facebook or the inclusion of challenges and competitions between groups [42.19% (i.e., 27) of experimental groups]; or exercise routines [17.19% (i.e., 11) of experimental groups].

#### **Risk of Bias and Certainty of the Evidence**

Additional File 5 shows the risk of bias assessment for each included unit of analysis. For total daily steps, 60.38% ( $n=32$ ) of units of analysis were assessed as overall "high risk," 37.74% ( $n=20$ ) as "some concerns," and only 1.88% ( $n=1$ ) as overall "low risk." Regarding MVPA, 80.00% ( $n=16$ ) was assessed as overall "high risk," 10.00% ( $n=2$ ) as "some concerns," and 10.00% ( $n=2$ ) as overall "low risk." Concerning total PA, 75.00% ( $n=6$ ) was assessed as overall "high risk" and 25.00% ( $n=2$ ) as "some concerns." Finally, for SB, 62.50% ( $n=5$ ) was assessed as overall "high risk," 12.50% ( $n=1$ ) as "some concerns" and 25.00% ( $n=2$ ) as overall "low risk."

Particularly, the risk of bias arising from the randomization process was the most problematic domain, with 21.35% of units of analysis being classified as "high risk" for not being a random design and 44.94% classified as "some concerns" due to being a non-controlled design or presenting imbalances regarding sample size or PA status between control and experimental groups. Moreover, the bias in measurement of the outcome domain was also problematic with half of the unit of analysis classified as "high risk" (13.48%) or "some concerns" (37.08%) due to inappropriate method of measuring (e.g., studies using consumer-wearables as measurement instrument instead of research-grade).

Moreover, based on the GRADE assessment (Additional File 6), the effect size of daily total steps was classified as "low" certainty of evidence due to both the risk of bias and inconsistency domains presenting serious limitations (i.e., most studies were classified as "high risk" of bias, and a substantial level of heterogeneity was found, respectively). Regarding MVPA, total PA and SB, the effect size was classified as "moderate" certainty of evidence due to serious limitations in the risk of bias domain (most studies were classified as "high risk" of bias).

#### **Publication Bias**

##### ***Avoiding Duplicated Information***

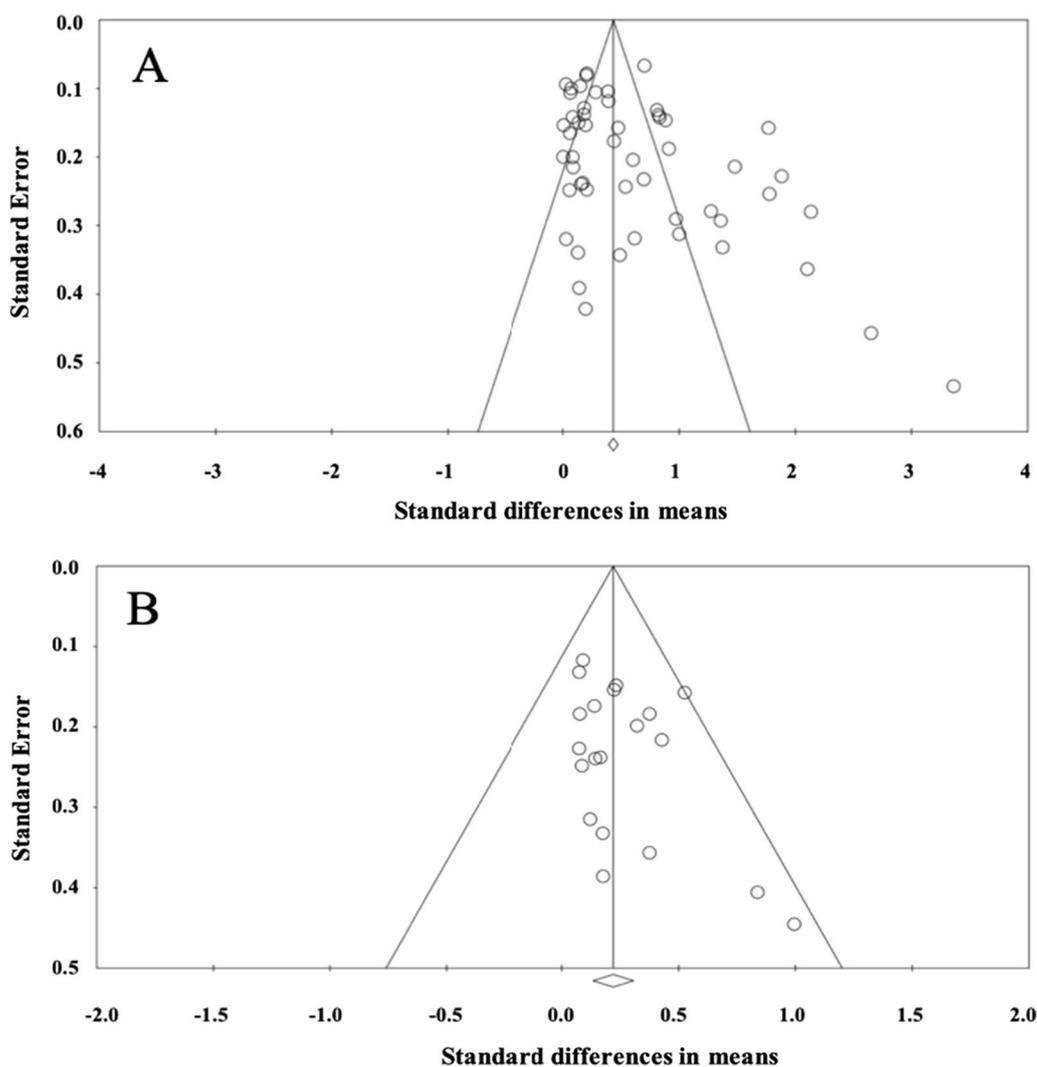
Although three publications met the selection criteria [56–58], they were not included in the systematic review and meta-analysis because the same data had been reported in other publications.

##### ***Identifying Publication Bias***

The visual assessment of funnel plots suggested that there was publication bias for daily total steps and MVPA (Fig. 2). In this sense, the results of the Egger's test were statistically significant for daily total steps ( $p=0.001$ ) and MVPA ( $p=0.045$ ).

##### ***Assessing the Impact of Publication Bias***

The results of Orwin's fail-safe  $N$  analyses showed that the number of missing studies needed to bring the mean  $d$  values under a trivial value was unlikely, especially if the percentage of unlocated/located studies was considered, being 177 (333.96%) for daily total steps and 25 (125.00%) for daily MVPA. Furthermore, the results of Duval and Tweedie's Trim and Fill method did not trim any study for daily steps and only two studies for MVPA in the standardized mean difference  $d$  analyses. Regarding mean difference  $D$  analyses, no study for daily steps and only one study for daily MVPA were trimmed. Moreover, the adjusted value for MVPA was similar to the observed values ( $d=0.220$  vs. 0.213;

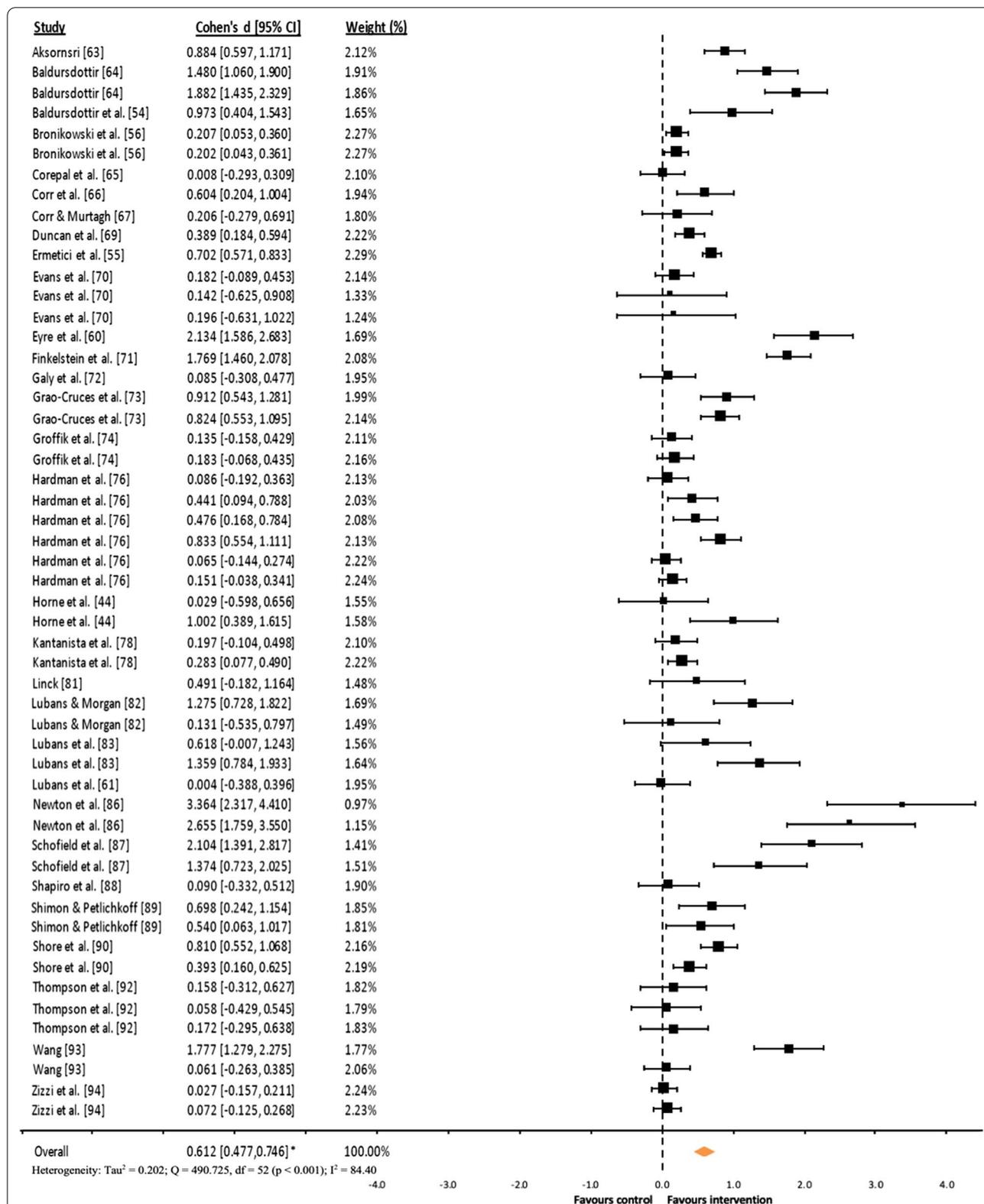


**Fig. 2** Funnel plots of standard error by standard differences in means (*d*) comparing consumer-wearable activity tracker-based programs effects on **A** daily total steps, and **B** moderate-to-vigorous physical activity

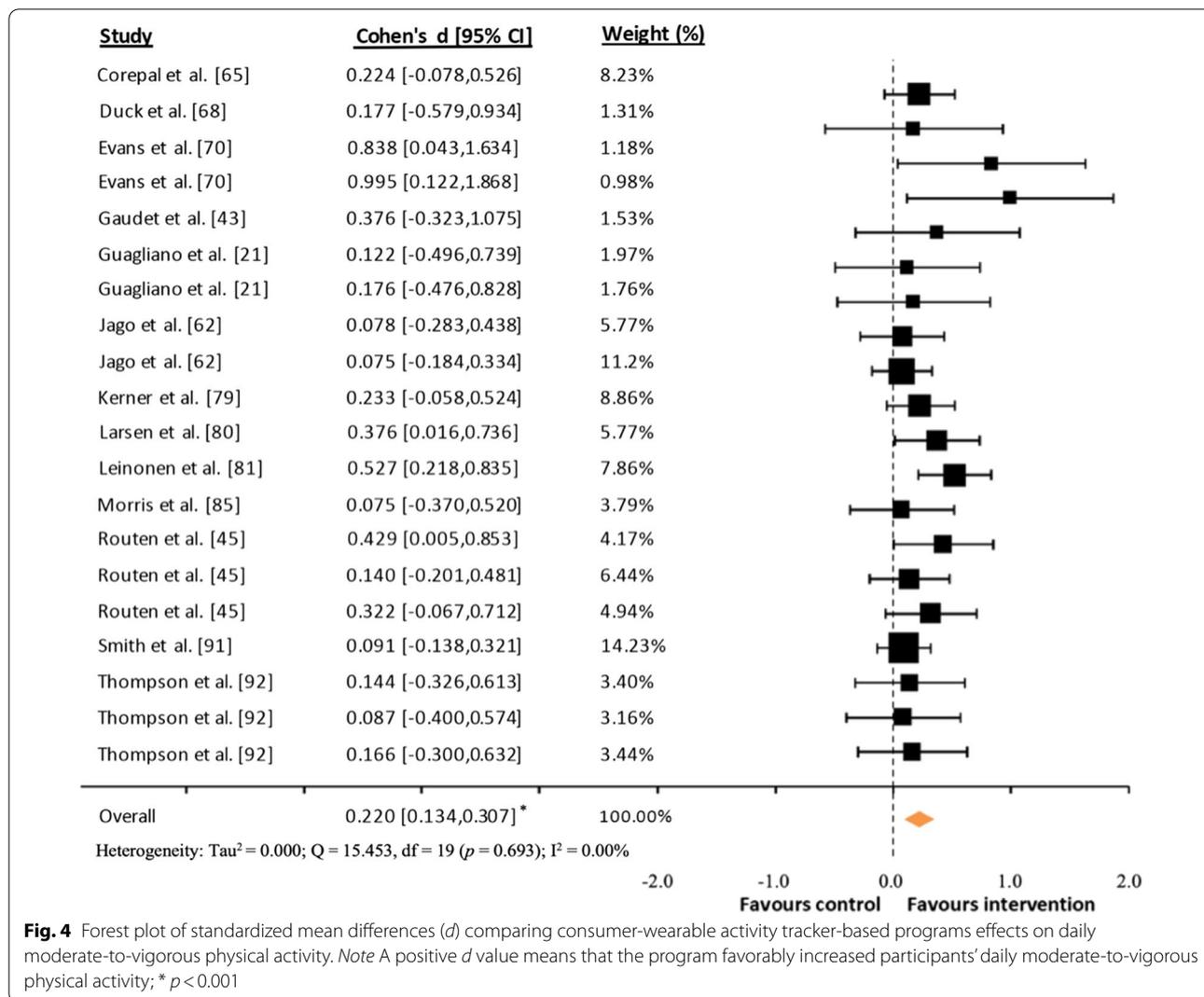
$D=5.583$  vs.  $4.824$ ). Finally, regarding the cumulative meta-analysis plots sorted by larger study (Additional File 7), after some fluctuations in the first studies which may be due to chance [59], a fairly constant estimate of the effect over sample size was observed for total steps and MVPA. Although a large effect in the first primary study was found, the summary value was decreased after the 14th-to-15th study and the 2nd-to-3rd study for daily total steps and MVPA, respectively. Furthermore, neither a transient loss of formal significance nor a complete reversal of the initial association was found. Finally, it is worth mentioning that the addition of new primary studies did not materially change the estimates, so the final effect size values for these intervention programs seem to be quite robust.

**Effects Sizes**

The results of the effects sizes showed that consumer-wearable activity tracker-based programs had a statistically significant moderate favorable effect on daily total steps ( $d=0.612$ , 95% CI  $0.477-0.746$ ,  $p<0.001$ ; Fig. 3), small favorable effect on daily levels of MVPA ( $d=0.220$ , 95% CI  $0.134-0.307$ ,  $p<0.001$ ; Fig. 4), and trivial favorable effect on daily levels of total PA ( $d=0.151$ , 95% CI  $0.038-0.264$ ,  $p=0.009$ ; Fig. 5). Moreover, the results of effect sizes carried out with mean difference showed significant favorable effects on daily total steps ( $D=1,692.792$ , 95% CI  $1,322.522-2,063.061$ ,  $p<0.001$ ) and MVPA ( $D=5.583$ , 95% CI  $2.527-8.640$ ,  $p<0.001$ ). However, the programs had a statistically significant trivial unfavorable effect on



**Fig. 3** Forest plot of standardized mean differences (*d*) comparing consumer-wearable activity tracker-based programs effects on daily total steps. Note A positive *d* value means that the program favorably increased participants' daily total steps; \* *p* < 0.001



daily SB (*d* = 0.172, 95% CI 0.039–0.305, *p* = 0.011; Fig. 6).

Furthermore, regarding the heterogeneity of the results, for daily MVPA (only for *d* analysis), total PA and SB heterogeneity was not found (*I*<sup>2</sup> = 0.00%). However, high heterogeneity was found for daily total steps (*I*<sup>2</sup> = 89.40% and 93.36% for *d* and *D* analyses, respectively), and moderate heterogeneity for daily MVPA (*I*<sup>2</sup> = 54.23%, only for *D* analysis). Therefore, together with the fact that the number of studies with daily total PA and SB was low (*k* = 8), follow-up subgroups analyses were conducted only for daily total steps and MVPA.

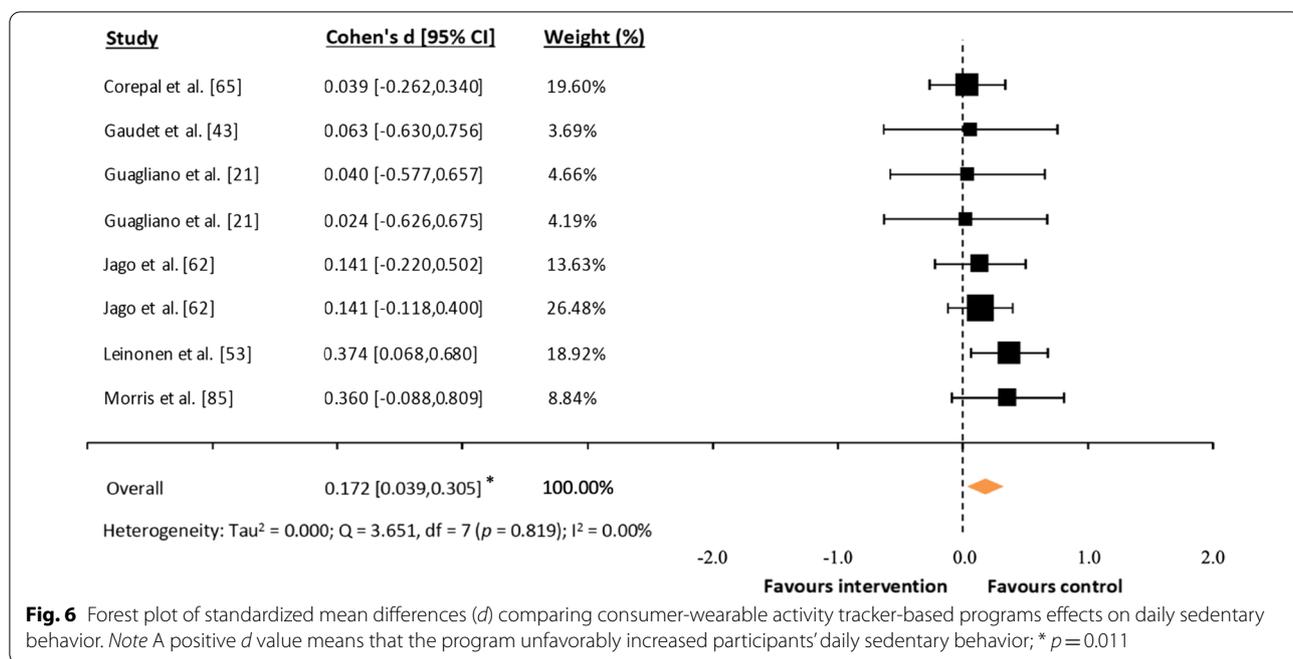
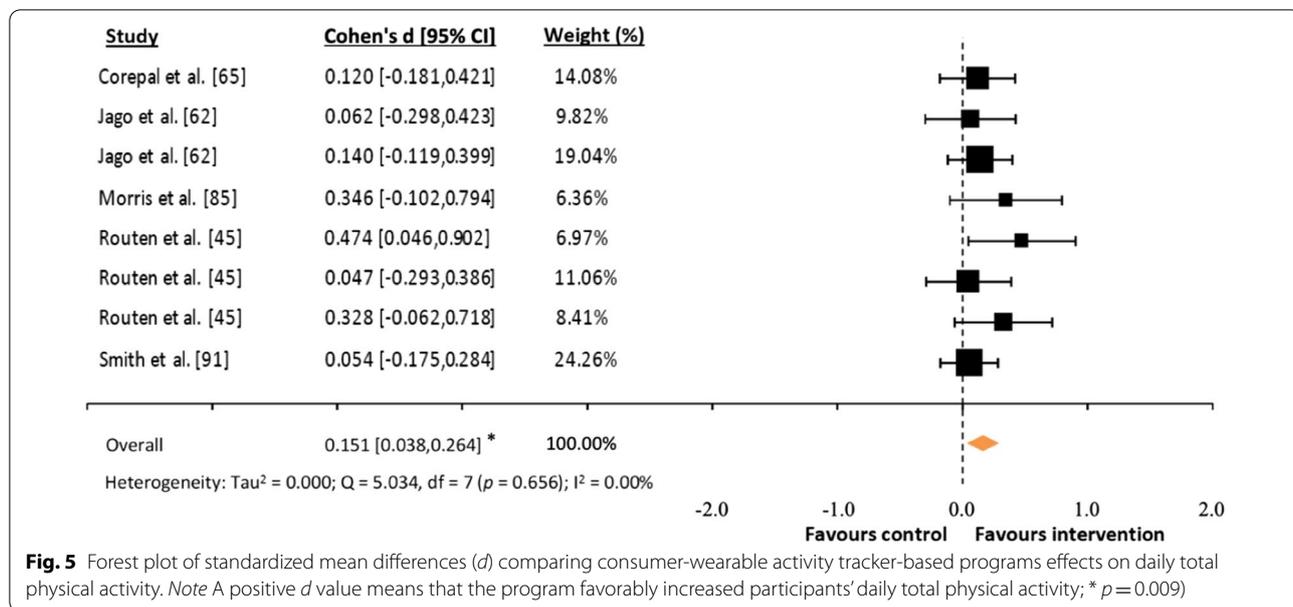
**Subgroups Analyses**

**Subgroups Analyses for Daily Total Steps**

Additional File 8 shows the results of the within-study subgroups analyses for the effect of the consumer-wearable

activity tracker-based programs on the daily total steps among school-aged children. The initial values of school-aged children's accomplishment with PA recommendations was the only individuals' characteristic that influenced the effect of the programs, being more effective in those physically inactive than in those physically active (*d* = 1.206 vs. 0.107; *p* < 0.001). Moreover, heterogeneity analyses showed that the effect of the program separately for accomplishment with PA recommendations was homogeneous (*I*<sup>2</sup> = 0.00%).

Furthermore, Additional File 9 shows the results of the between-study subgroups analyses. Regarding the influence of the individuals' characteristics, subgroups analysis showed that the program had statistically significantly more effect in females than in males (*d* = 0.636 vs. 0.266; *p* = 0.044) and in physically inactive (i.e., not meeting the daily PA recommendations)



than in those that were physically active ( $d = 0.795$  vs.  $0.397$ ;  $p = 0.003$ ) for increasing daily total steps. As regards the influence of the intervention characteristics, programs with some kind of counseling were more effective than those without counseling ( $d = 0.711$  vs.  $0.407$ ;  $p = 0.003$ ); likewise, programs that included goal-setting were more effective than those that did not ( $d = 0.770$  vs.  $0.243$ ;  $p < 0.001$ ) for increasing

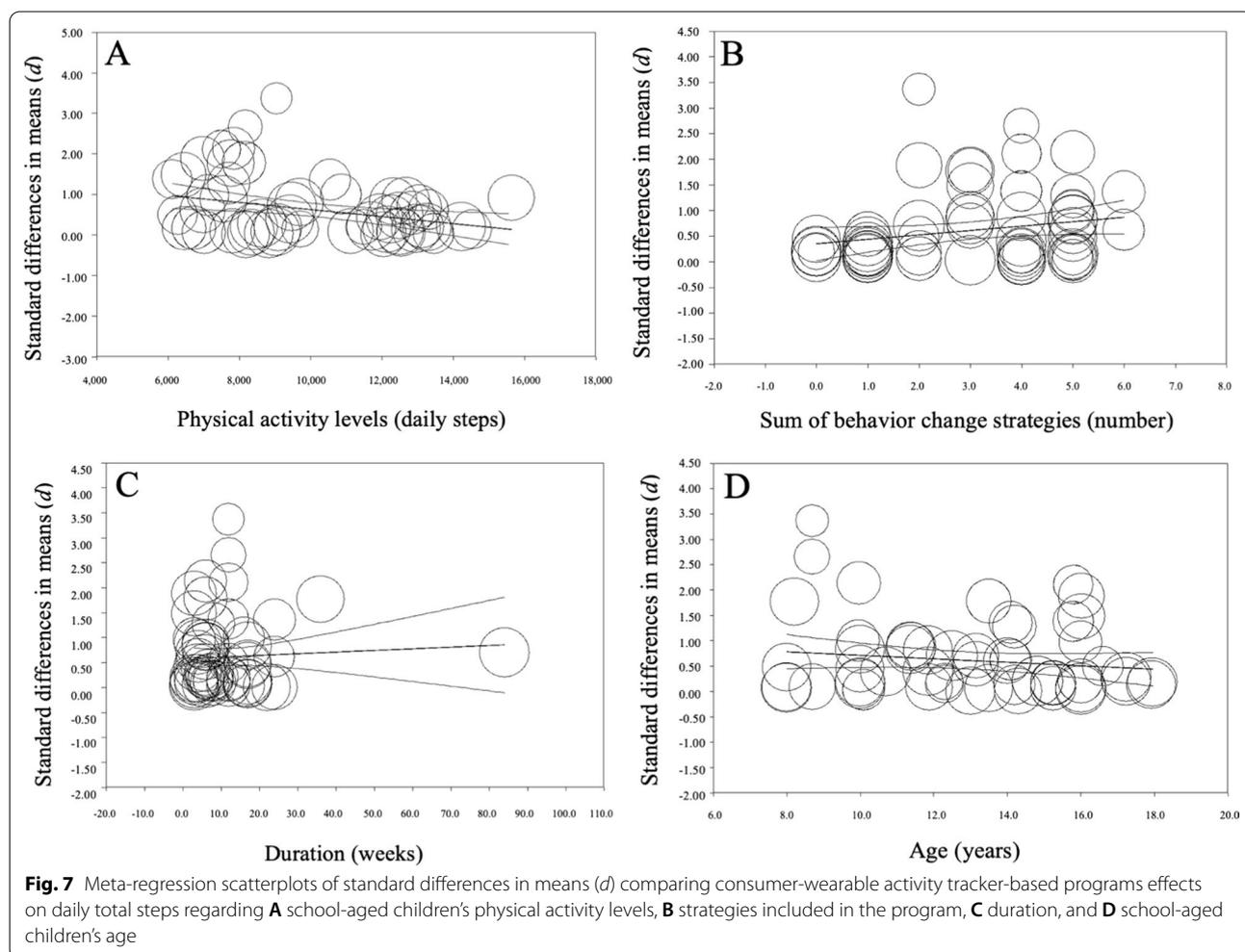
school-aged children's daily total steps. However, for the rest of the between-study subgroups comparisons no statistically significant differences were found ( $p > 0.05$ ). According to heterogeneity analyses, the effect of these programs separately for sex, PA status, goal-setting and counseling was still moderate-to-high heterogeneous ( $I^2 = 66.08-92.54\%$ ).

Moreover, meta-regression analysis showed that the effect of the intervention program was statistically significantly associated with PA status and number of strategies, with the effect being higher for daily total steps in less active school-aged children ( $Q=10.83$ ;  $p=0.001$ ;  $R^2=0.07$ ;  $I^2=88.58\%$ ), and when a greater number of strategies were included in programs ( $Q=5.19$ ;  $p=0.023$ ;  $R^2=0.02$ ;  $I^2=89.02\%$ ) (Fig. 7). However, no statistically significant associations were found between intervention duration ( $Q=0.38$ ;  $p=0.536$ ;  $R^2=0.00$ ;  $I^2=89.04\%$ ) or age ( $Q=2.17$ ;  $p=0.141$ ;  $R^2=0.01$ ;  $I^2=89.17\%$ ) and school-aged children's daily total steps. Furthermore, meta-regression analysis of a more complex model including both significant explanatory variables together (i.e., PA status and number of strategies included), showed a statistically significant association with intervention effect on daily total steps, explaining a higher percentage of the variance than both explanatory variables alone ( $Q=16.55$ ,  $p<0.001$ ;  $R^2=0.09$ ;  $I^2=89.40\%$ ).

**Subgroups Analyses for Daily Moderate-to-Vigorous Physical Activity**

Additional File 10 shows the results of the within-study subgroups analyses for the effect of the consumer-wearable activity tracker-based programs on the daily MVPA levels among school-aged children. Results showed no statistically significant differences for any subgroups comparisons (i.e., sex, goal-setting, diary, and reminders).

Furthermore, Additional File 11 shows the results of the between-study subgroups analyses. Regarding the influence of the school-aged children's characteristics, subgroup analyses showed that the program had significantly more effect in physically inactive participants than in those physically active ( $d=0.404$  vs.  $0.170$ ;  $p=0.046$ ) for increasing daily MVPA levels. As regards the influence of the intervention programs' characteristics, programs carried out with a wrist-worn activity tracker were more effective than those with a waist-worn activity tracker ( $d=0.413$  vs.  $0.167$ ;  $p=0.021$ ); moreover, programs that did not include any exercise routine were



**Fig. 7** Meta-regression scatterplots of standard differences in means ( $d$ ) comparing consumer-wearable activity tracker-based programs effects on daily total steps regarding **A** school-aged children's physical activity levels, **B** strategies included in the program, **C** duration, and **D** school-aged children's age

more effective than those that did ( $d=0.283$  vs.  $0.083$ ;  $p=0.036$ ). However, for the rest of the between-subgroups comparisons no statistically significant differences were found ( $p>0.05$ ). According to heterogeneity analyses, results by subgroups showed homogeneous results ( $I^2=0.00-29.06\%$ ).

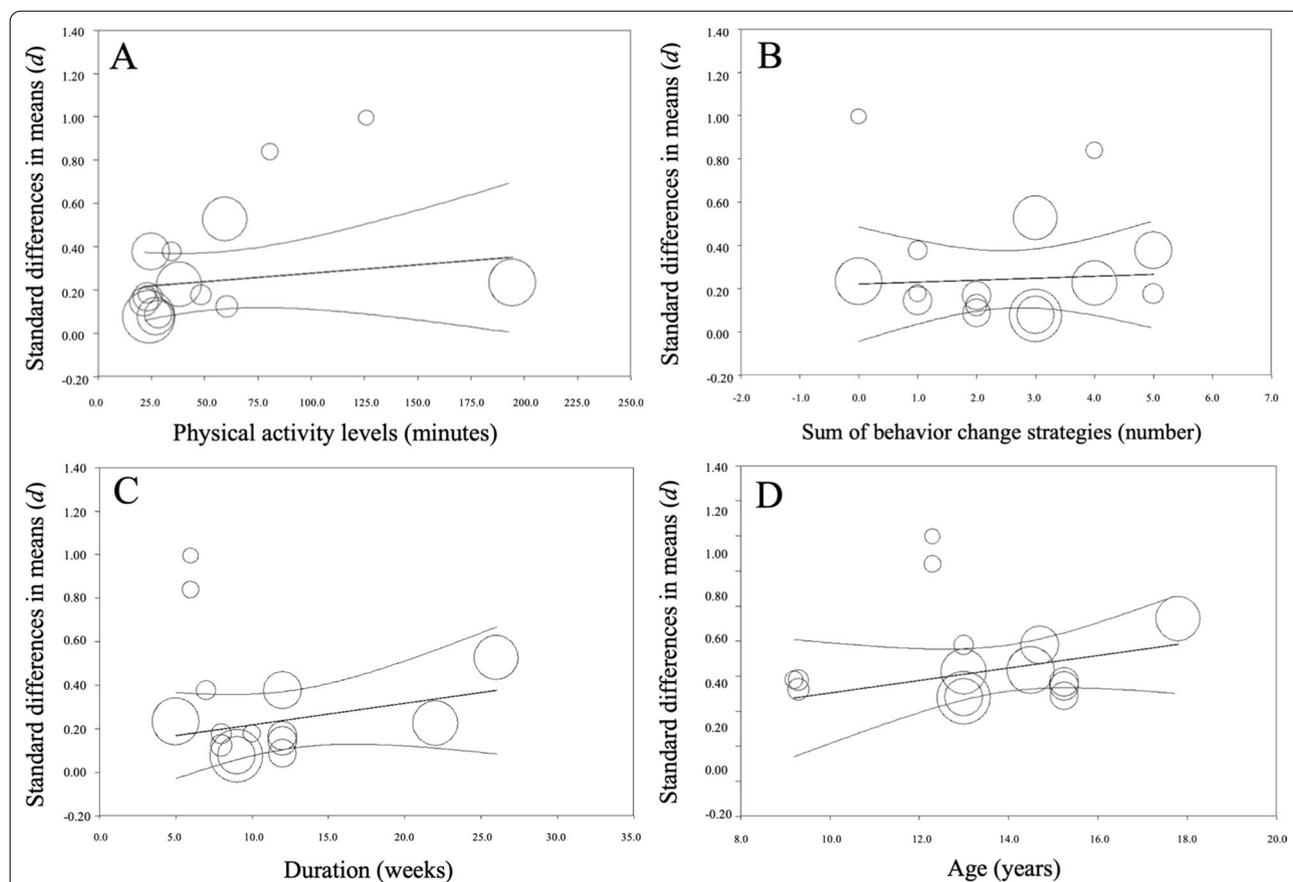
Moreover, meta-regression analyses did not show statistically significant associations between PA status ( $Q=0.65$ ;  $p=0.420$ ;  $R^2=0.00$ ;  $I^2=0.00\%$ ), number of strategies ( $Q=0.06$ ;  $p=0.805$ ;  $R^2=0.00$ ;  $I^2=0.00\%$ ), intervention duration ( $Q=1.55$ ;  $p=0.213$ ;  $R^2=0.00$ ;  $I^2=0.00\%$ ), or participants' age ( $Q=1.84$ ;  $p=0.175$ ;  $R^2=0.00$ ;  $I^2=0.00\%$ ) and school-aged children's daily MVPA levels (Fig. 8).

**Sensitivity Analyses**

The results of the sensitivity analyses for the overall effects sizes carried out with Hedges'  $g$  random-effects model for daily total steps ( $g=0.600$ , 95% CI  $0.467-0.732$ ,  $p<0.001$ ), MVPA ( $g=0.218$ , 95% CI  $0.133-0.302$ ,  $p<0.001$ ), total PA ( $g=0.150$ , 95% CI  $0.039-0.261$ ,

$p=0.008$ ) and SB ( $g=0.171$ , 95% CI  $0.038-0.303$ ,  $p=0.011$ ) showed the same results as the main analysis carried out with Cohen's  $d$  with a random-effects model. Similarly, overall effects sizes carried out with Cohen's  $d$  with a fixed-effect model for daily total steps ( $d=0.434$ , 95% CI  $0.392-0.475$ ,  $p<0.001$ ), MVPA ( $d=0.220$ , 95% CI  $0.134-0.307$ ,  $p<0.001$ ), total PA ( $d=0.151$ , 95% CI  $0.038-0.264$ ,  $p=0.009$ ) and SB ( $d=0.172$ , 95% CI  $0.039-0.305$ ,  $p=0.011$ ) also showed the same results as the main analysis with the random-effects model.

Finally, results of the Cohen's  $d$  with a random-effects model separately for randomized controlled trial design compared with non-randomized controlled trial design showed greater effects in randomized controlled trials for daily total steps ( $d=0.859$  vs.  $0.465$ ;  $p=0.004$ ). However, no statistically significant differences were found for MVPA ( $p=0.104$ ) or total PA ( $p=0.328$ ). Regarding SB, all the included studies were randomized controlled trials and the comparison could not be performed.



**Fig. 8** Meta-regression scatterplots of standard differences in means ( $d$ ) comparing consumer-wearable activity tracker-based programs effects on daily moderate-to-vigorous physical activity regarding **A** school-aged children's physical activity levels, **B** strategies included in the program, **C** duration, and **D** school-aged children's age

## Discussion

### Overall Effects of Consumer-Wearable Activity Tracker-Based Programs

The present systematic review and meta-analysis synthesizes the evidence to date about the effectiveness of consumer-wearable activity tracker-based programs on daily objectively measured PA and SB among apparently healthy school-aged children [20, 21, 43–45, 52–55, 60–94]. The overall results showed that the consumer-wearable activity tracker-based programs brought about significant moderate improvements in school-aged children's daily total steps, and small but significant improvements in daily MVPA levels after the intervention program. However, regarding daily levels of total PA the effect was trivial. Therefore, the use of a consumer-wearable activity tracker as a motivation tool for young people is strongly recommended to reduce the high levels of physical inactivity among school-aged children [8]. However, the intervention programs seem not to reduce the school-aged children's SB.

These results agree with similar previous meta-analyses carried out in adults [24–26]. Firstly, regarding daily total steps, all previous reviews found improvements, although they seem greater in the present systematic review ( $d=0.612$  vs.  $0.240$ – $0.449$ ;  $D=1,692.79$  vs.  $950.54$ ). Regarding changes in actual units, adding 1,000 steps per day, which is equivalent to a 10% increase within the recommended 10,000 steps per day [19, 95], has been shown to be a significant and clinically meaningful change related to a substantial reduction of the risk of all-cause mortality in the adult population [96, 97]. Therefore, although there is no such evidence in studies carried out with school-aged children, the increase of 1,692.79 daily steps obtained in the present meta-analysis might be considered an important and meaningful change considering the existing evidence with adults. Moreover, with reference to MVPA, similar improvements in daily levels of MVPA were obtained ( $d=0.220$  vs.  $0.270$ ;  $D=5.583$  vs.  $6.160$ ). With reference to changes in actual units, to our knowledge, there is no clinical evidence of a link between changes in minutes involved in MVPA and health outcomes in any population. However, extrapolating the results obtained as a percentage of the international recommendations (i.e., 60 min of MVPA per day [4]), the obtained results (i.e., 5.6 min) represent almost a 10% increase. Therefore, considering the same reasoning explained above with daily total steps, these changes in daily levels of MVPA may also be considered clinically significant.

The greater improvement in school-aged children's daily total steps than in their MVPA levels may be due to the kind of goal established in the program. Most of the studies included in the systematic review with a

goal-setting strategy set only a step-based goal (30 of 37 studies), while only five studies established both step-based and minutes of total PA-based goals [62, 87] or SB-based goals [88, 91]. Therefore, all the motivational strategies included, such as reminders [44, 54, 66], counseling sessions [70], or rewards [65, 73] were carried out around this goal of increasing the number of steps. Moreover, the reason for relying mainly on the number of steps as the reference output for goal-setting may be due to steps having the advantage of being easier to understand and interpret by school-aged children compared to MVPA minutes [19]. In addition, in many studies, the consumer-wearable activity-tracker used were pedometers, which only show feedback about the number of steps. Therefore, similar to the evidence found with adults [24–26], consumer-wearable activity tracker-based programs seem to be effective for improving objectively measured daily PA, especially for daily total steps, although the effects seem to be greater for school-aged children. This greater effect may be due to the fact that school-aged children have a higher affinity with new technologies playing an important role in their daily life, and therefore these kinds of technology-based interventions may be more interesting for them [98, 99]. In addition, during the period of childhood and early adolescence, school-aged children are still forming their daily habits and they are more sensitive to changing their PA behavior which could explain these greater intervention effects, while in adulthood the stability of PA patterns is high and more difficult to be changed [100].

Regarding total PA, results showed a trivial favorable effect ( $d=0.151$ , 95% CI  $0.038$ – $0.264$ ). These results are curious because daily total steps is an indicator of total PA [101], and therefore, results were expected to be similar. However, it may be because very slow steps were included in daily total steps, but they do not reach the threshold to be considered light PA, or because fewer studies are assessing total PA compared to studies assessing daily total steps. Nevertheless, in the present meta-analysis, only the study by Corepal et al. [65] measured both variables in the same study, without any clear relationship between both outputs. Therefore, future studies including the evaluation of both variables (i.e., daily total steps and total PA) would be very interesting to establish a real cause–effect relationship between them.

Besides, regarding SB, similar to the trivial unfavorable effect obtained in the present systematic review ( $d=0.172$ ; 95% CI  $0.039$ – $0.305$ ), no real differences were found in any previous review [24, 26]. Firstly, this may be due to the fact that most of the programs that evaluate SB only used strategies to encourage and support PA behavior change (i.e., goals, tips and challenges, behavioral incentives or reinforcement messages based

only on PA practice) and were not specifically designed to reduce SB [21, 43, 62, 65, 77, 85]. Only Leinonen et al. [53] specifically included some SB-based strategies, such as feedback showing a thumb either up or down if the day included over two hours of sedentary (sitting) periods or not, and rewards regarding decrement in weekly school-aged children's sedentary time. The lack of specific focus on reducing SB may have contributed to this trivial unfavorable effect. Moreover, these results may also present significant measurement bias due to all SB-studies (except Morris et al. [85]) analyzed raw time involved in SB per day instead of valid wear time-based standardized scores (e.g., percentage of time of each day engaged in SB of the total valid wear time; standardized mean SB in minutes) [102, 103]. In addition, it must be considered that for accelerometer-based measures only a valid minimum time per day is established (normally 600 min), but that standardized values are not taken into account [104]. This is even more accentuated for consumer-based wearables for which valid wear time cannot be known and, therefore, it cannot be controlled if school-aged children wore the wearable for a long enough time, or if they were more motivated to wear it for more time in the baseline or post-intervention measure. In this sense, it is easier for school-aged children with more valid time to have higher registered time involved in SB, so SB time outcomes could be especially affected by potential systematic valid wear-time variation between measurement moments (i.e., pre–post-intervention measures) or groups [102]. This could also affect daily total steps and daily MVPA levels although to a much lesser extent than for SB [102]. For instance, Gaudet et al. [43] showed that school-aged children in the experimental group had approximately 48 min more wear time per day in comparison with the control group, which may directly affect their differences regarding time involved in SB. However, most studies [21, 62, 65, 77] only reported the minimum time per day needed to be included in the study but did not report the actual mean valid wear time per day, and it is important that this be reported in order to compare compliance between groups. Finally, it should be noted that there are very few studies measuring school-aged children's SB in comparison with studies measuring steps and MVPA, which implies a wider confidence interval and consequently greater uncertainty about the real value.

#### **Influence of Participants' Characteristics**

According to the results of the present meta-analysis, a significant relationship has been observed for PA status from the within-study subgroups analysis, the intervention being much more effective for improving daily total steps in school-aged children who were physically inactive in the baseline measure than in those who were

physically active. Furthermore, this positive influence was also found for improving objectively measured daily MVPA levels, although only from the between-study analysis (i.e., an observational relationship). A potential reason for this influence could be that school-aged children who already are physically active before the intervention are motivated enough for PA practice without the need for extra motivation with the proposed intervention [105]. Furthermore, it is difficult to further increase school-aged children's PA levels when the baseline levels are high, just as has occurred with the improvement of physical fitness levels [106, 107]. Therefore, consumer-wearable activity tracker-based programs may be an especially appropriate strategy for less active school-aged children [82, 85]. In line with the recommendation by Love et al.'s [108] systematic review, future studies should analyze their results distinguishing by school-aged children's baseline PA profiles to correctly identify the intervention impact because all participants do not react in the same way to the intervention. Regarding school-aged children who already are physically active before the intervention, it may be necessary to study which specific strategies should be implemented in these interventions to help them maintain their PA levels, or even to continue increasing them, therefore obtaining greater health-related benefits [4]. Finally, regarding school-aged children's age, it does not seem to affect the intervention program effect.

Furthermore, programs seem to be more effective in females than in males for improving daily total steps, which could be related to PA status since females tend to be more physically inactive than males throughout childhood and adolescence [109, 110]. Therefore, sex-specific interventions could be considered in future research like Böhm et al. [29] suggested, although the conclusions in relation to the effectiveness of sex-specific interventions should be taken with caution because the success of those programs was not obtained by the within-study subgroups analysis (i.e., cause–effect relationships), but rather by the between-study analysis, which only establishes observational relationships [20, 44, 69, 73, 74, 83]. Furthermore, the between-study subgroups analysis results still showed moderate-to-high heterogeneity in each subgroup (i.e., males and females), so it is likely that there were differences in other strategies of the intervention or participants' characteristics that could influence its effect, even more than sex itself. Additionally, these differences are even more accentuated because the analysis includes twice as many publications carried out with females than with males. However, the number of studies including this within-study sex comparison and which allow establishing cause–effect relationships is very limited ( $k=7$ ), and for this reason, the observational

relationships of the between-study comparisons have been considered in the present meta-analysis.

#### **Influence of Intervention Programs' Characteristics**

Firstly, overall meta-regression analysis results showed a higher effect for daily total steps when a greater number of strategies were included in the programs. Therefore, multi-dimensional interventions that include most of these strategies seem to be preferable for mediating PA behavior. That agrees with some psychological theories which include most of those strategies as positive mediators influencing PA behavior. For instance, the Social Cognitive Theory [111], highlighted PA-related knowledge included in counseling sessions, positive reinforcement such as reminders with encouraging messages about PA practice (e.g., "You're in charge! Make the choice to meet your step goal today!"), and the importance of setting achievable goals as determining factors in the design of PA promotion interventions that could lead to behavior change in school-aged children. In addition, the Self-Determination Theory [112], emphasizes the need for relatedness included in most motivational strategies (e.g., teamwork or the use of social networks), and the perceived competence reflected in the reinforcement reminders praising their efforts (e.g., "You can meet your step goal; just keep stepping!"), or the evolution that school-aged children could observe in the diary they filled in during the program, as necessary to increase their intrinsic motivation. Furthermore, recent systematic reviews highlight that using multi-strategy approaches as behavior change techniques show better PA outcomes than singular change approaches [113, 114]. For instance, the PA-related knowledge provided in the counseling sessions, the autonomy support environment by the consumer-wearable activity tracker feedback, the inclusion of additional motivational strategies like social networks, setting goals of moderate difficulty, or sending reminders with encouraging messages about PA practice have been shown to positively change school-aged children's PA behavior [115, 116]. However, it should be noted that the explained variance was low ( $R^2=0.02$ ), as well as the results being highly heterogeneous ( $I^2=89.02$ ).

Furthermore, there was considerable heterogeneity in the strategies included in the reviewed studies. Most of the programs used a goal-setting strategy, participants' logbooks, educational counseling sessions, and/or some kind of motivational strategy. Nevertheless, the inclusion of reminders to persuade participants to move or exercise more was a less frequently included strategy. The influence of these intervention program characteristics

in school-aged children's daily total steps and MVPA has been analyzed in the present meta-analysis. Firstly, interventions including some kind of counseling [54, 60, 71] and/or goal-setting techniques [73, 86, 90], in addition to consumer-wearable activity trackers, were highlighted as more effective than those without them for improving school-aged children's daily total steps. These results are in accordance with previous studies' recommendations about the inclusion of these explicit strategies (e.g., advice about PA benefits, strategies to reduce SB and increase PA, resolution of barriers to PA practice, or goal-setting strategies based on the international guidelines) which make students feel that they are making an informed decision about their health in any kind of program for PA promotion [11, 13] and specifically in wearable-based programs [27, 115]. Nevertheless, apparently contradictory results showed that consumer-wearable activity trackers-based programs which did not include any exercise routine seem to be more effective than those that included it for improving school-aged children's daily MVPA levels. However, analyzing the kind of exercise routine included, had some limitations. Firstly, Jago et al. [62] and Smith et al. [91] included a low frequency of supervised PA sessions (one 20-min PA session per week, and only six 20-min lessons in 20 weeks, respectively) with which it is very difficult to positively affect the school-aged children's daily PA levels. Secondly, most of the activities included by Jago et al. [62] did not have a direct relationship with increasing the school-aged children's number of steps or minutes involved in MVPA (i.e., stretching, technical drills, or strengthening tasks), and also it should be noted that despite including this exercise routine, they did not include other strategies that may be even more important than this one (e.g., reminders, diary, or motivational strategies). Finally, it is also important to denote that the analysis included 17 units of analysis without an exercise routine vs. three with an exercise routine [62, 91], so given this marked difference in the sample, results should be interpreted with caution. Moreover, it must be considered that only observational relationships have been obtained from the between-study subgroups analyses due to the low number of studies compared in the within-study subgroups analyses. Furthermore, the results regarding school-aged children's daily total steps separated by each subgroup still showed a high level of heterogeneity which implies differences in other intervention characteristics. Therefore, future studies should include different intervention groups that compared some intervention characteristics (e.g., one experimental group including counseling and another without counseling) to establish causal-effect relationships between intervention characteristics and the effect of the programs.

Regarding the kind of consumer-wearable activity tracker, waist-worn trackers such as pedometers are more common than wrist-worn trackers. This could mainly be due to the fact that waist-worn trackers such as traditional step counters with digital displays, have featured in scientific research since approximately 1996, being an accepted method for assessing PA and a tool for walking interventions [117]. On the contrary, wrist-worn trackers have burst onto the market in the last decade and, therefore, their scientific evidence base is still scarce [14, 18]. Moreover, the meta-analysis results showed that programs carried out with a wrist-worn activity tracker seem to be more effective than those carried out with waist-worn trackers for improving school-aged children's daily MVPA levels. This may be due to wrist-worn trackers having several advantages compared to waist-worn trackers, such as reporting real-time feedback that can be easily checked on their wrist or touch screens [14]. Moreover, unlike more traditional waist-based trackers, which only monitor and display simple feedback about PA levels, wrist-based trackers are much more interactive since the user is able to set reminders, notifications, or congratulatory messages upon reaching the proposed goal [118]. Finally, wrist devices have shown greater wear time compliance which could mean that if they wear it for a longer time school-aged children could interact more with its features [119].

#### **Risk of Bias and Certainty of the Evidence**

Firstly, based on the methodological risk of bias assessment, most studies were classified as "high risk" or "some concerns," leaving only two studies classified as overall "low risk" [21, 81]. Therefore, this may have resulted in a biased assessment of the intervention effect, underestimating or overestimating the true intervention effect, which meant downgrading the GRADE certainty rating by one level for all outcomes (i.e., daily total steps, MVPA, total PA, and SB) regarding the risk of bias domain [34, 46]. In reference to the study designs, it is interesting to highlight that only 51.11% of the included studies are true or cluster-randomized controlled trials, which are markedly far stronger interventions to demonstrate effect significance [120]. However, sensitivity analysis showed no differences for school-aged children's daily MVPA, total PA, and SB levels between study designs, but much greater effects were found in randomized controlled trials for school-aged children's daily total steps than non-randomized trials.

Moreover, regarding daily total steps outcomes, a substantial level of heterogeneity was found, even in the follow-up subgroups analyses (except when separating by accomplishment of PA recommendations) and

it meant downgrading the GRADE certainty rating by another level regarding inconsistency domain for daily total steps. This is most likely because it is the PA outcome that includes the largest number of studies and, consequently, the greatest variety in the types of intervention applied when compared with other outcome measures. Finally, regarding publication bias, although the funnel plots and Egger's test suggested publication bias for daily total steps and daily MVPA levels, its impact seems to be very low given the unlikely number of "lost" studies suggested by the fail-safe  $N$  analyses. Furthermore, the Trim and Fill method did not trim any study for daily steps and only two studies for MVPA, resulting in an adjusted value similar to the observed values ( $d=0.220$  vs.  $0.213$ ).

For all the above-mentioned reasons, it is important to highlight the "low" certainty of evidence found for daily total steps, which means that the confidence in the effect estimate is limited and the true effect may be different from the estimated effect [121]. Regarding MVPA, total PA and SB outcomes, "moderate" certainty of evidence was found, so the true effect is likely close to the estimated effect, but there is a possibility that it is substantially different [121]. Therefore, the findings of the present meta-analysis should be considered with caution and firmer conclusions should await the accumulation of a larger high-quality number of primary studies.

#### **Strengths and Limitations**

Regarding the strengths of the present systematic review, numerous measures to avoid, or at least to reduce, publication bias were followed (e.g., the inclusion of a great range of bibliographic databases from different disciplines and complementary search strategies, or not restricting the search by the language, type or date of publication). Then, several exploratory analyses were conducted to identify and assess the impact of any potential publication bias (e.g., funnel plots, or Orwin's fail-safe  $N$  analyses), as well as sensitivity analyses (e.g., Hedges'  $g$  with a random-effects model or Cohen's  $d$  with a random-effects model separately for randomized controlled trial design or not) to verify the robustness of the results. Furthermore, the present review was focused only on objective measurements, which have shown high validity to measure PA and SB levels in comparisons with self-reported measures [35, 36]. Lastly, to our knowledge, to date this is the first systematic review and meta-analysis about the effects of consumer-based activity tracker-based programs on objectively measured PA and SB levels within apparently healthy school-aged children, including analyzing the influence of the intervention programs' characteristics and school-aged children's

characteristics on the effects. This meta-analysis summarizes the effectiveness of those interventions in an overall statistical synthesis, improving the precision of the results by the estimation of the effect size and direction, and clarifying whether or not the effect size is consistent across studies.

However, the present systematic review and meta-analysis is not without limitations. First, although randomized controlled trials have higher methodological quality, the present systematic review includes several study designs. As expected, there were not many consumer-based activity tracker-based studies with a high level of quality design for improving the different PA-related behaviors. Therefore, a reason for including several designs is to provide evidence of the effects of interventions for which only a small number of randomized controlled trials are available, drawing on the “best available evidence” rather than the “highest tier” of evidence [34]. Nevertheless, sensitivity analyses were also performed comparing randomized controlled trials and non-randomized trials, showing no differences for school-aged children’s daily MVPA, total PA, and SB levels. Furthermore, even greater effects in school-aged children’s daily total steps were found in randomized controlled trials. Second, although the inclusion of a wide range of intervention types, populations, sample size, and study designs had some advantages regarding the generalizability of conclusions, it also means a high level of heterogeneity. For instance within intervention types, multiple behavior change strategies such as goal-setting (even including different kinds like static or adaptive goals) or extra motivational strategies (e.g., social networks or social support) were usually combined in the same study. Therefore, it makes the independent contribution of any intervention features and, therefore, making strong conclusions from the intervention difficult to establish. However, in addition to the overall effect size, subgroups analyses and meta-regression of the a priori hypothesized moderators were also performed. Therefore, not only general effect results are provided, but also results for each specific group based on the characteristics of the programs and school-aged children. Third, the present systematic review investigated effectiveness at the end of the program (i.e., short-term), but future studies should investigate long-term effectiveness to assess actual behavioral changes some months after the program. However, due to the very limited evidence, this was not performed in the present systematic review. Finally, coding some study outcomes was problematic due to authors not reporting them. Although authors were contacted, many of

them did not reply and the particular study outcome had to be omitted. However, this is a common problem in most systematic reviews [34], and a great effort was made in contacting authors, recalculating data, or estimating values from figures. Finally, in some cases, consumer-based activity trackers were used both as a motivational instrument during the intervention and to objectively measure PA, which could affect results by increasing their actual PA levels in the control group or during baseline assessments.

## Conclusions

The present findings suggest that consumer-wearable activity tracker-based programs within school-aged children have a statistically significant moderate favorable effect on daily total steps and a small, but favorable effect on objectively measured daily levels of MVPA. However, the favorable effect of the programs on school-aged children’s objectively measured total PA levels and the unfavorable effect in SB was trivial, although statistically significant. The findings of this systematic review suggest that programs are more effective in females for increasing daily total steps, and in physically inactive subjects for increasing both daily total steps and daily MVPA levels. Moreover, the inclusion of a greater number of strategies in the programs had a higher effect on school-aged children’s daily total steps. It should be highlighted that prompting specific goal-setting and the inclusion of educational counseling sessions are particularly useful strategies to include in consumer-wearable activity tracker-based programs designed to promote school-aged children’s daily total steps. Furthermore, regarding the kind of consumer-wearable activity tracker, programs were more effective at increasing school-aged children’s daily MVPA levels if a wrist-worn activity tracker was used. However, since only observational relationships have been obtained from the between-study subgroups analyses due to the low number of studies, all the above-mentioned recommendations regarding intervention strategies should be taken with caution. However, due to the certainty of evidence being from “low” to “moderate,” further primary research is needed to determine the effectiveness of these programs using robust designs with low risk of bias, and which compare the effect of different intervention characteristics in the same study. Consumer-wearable activity tracker-based programs (particularly those including goal-setting, educational counseling, and wrist-worn trackers) seem to be effective for promoting school-aged children’s daily total steps and MVPA, especially for females and those who are physically inactive.

## Abbreviations

PA: Physical activity; SB: Sedentary behavior; MVPA: Moderate-to-vigorous physical activity; WHO: World Health Organization.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40798-021-00407-6>.

**Additional file 1.** Prisma checklist.

**Additional file 2.** Search strategies.

**Additional file 3.** Coding form followed in the present systematic review.

**Additional file 4.** Algorithms followed for assessing methodological risk-of-bias in each domain.

**Additional file 5.** Risk of bias assessment: (a) Daily steps units of analysis; (b) Moderate-to-vigorous physical activity units of analysis; (c) Total physical activity unit of analysis; and (d) Sedentary behavior unit of analysis. Each row corresponds to a unit of analysis. Green symbols represent a "Low risk of bias", yellow symbols represent "Some concerns", and red symbols represent a "High risk of bias".

**Additional file 6.** Certainty of the evidence assessment.

**Additional file 7.** Results of the cumulative meta-analyses by study size for: (a) Daily total steps; and (b) Moderate-to-vigorous physical activity.

**Additional file 8.** Results of the within-study subgroups analyses for the effect of the consumer-wearable activity tracker-based programs on the daily total steps among school-aged children.

**Additional file 9.** Results of the between-study subgroups analyses for the effect of the consumer-wearable activity tracker-based programs on the daily total steps among school-aged children.

**Additional file 10.** Results of the within-study subgroups analyses for the effect of the consumer-wearable activity tracker-based programs on the daily moderate-to-vigorous physical activity among school-aged children.

**Additional file 11.** Results of the between-study subgroups analyses for the effect of the consumer-wearable activity tracker-based programs on the daily moderate-to-vigorous physical activity among school-aged children.

## Acknowledgements

Authors would like to thank all of the intervention authors for their willingness, time and efforts to contribute to this analysis through the provision of data. We also thank Aliisa Hatten for the English revision of the manuscript.

## Authors' Contributions

CCR, JV and DMV contributed to the conception and design of the study. CCR and SGR were responsible for collecting the data. CCR and DMV carried out the meta-analysis and were responsible for the analysis and interpretation of the data. CCR and DMV drafted the manuscript. All authors contributed to editing and revising the manuscript. All authors read and approved the final manuscript.

## Funding

Carolina Casado-Robles is supported by a research grant from the Spanish Ministry of Universities [Grant Number: FPU16/03314]. This publication is part of the WearFit project (Reference B-SEJ-029-UGR18), funded by the FEDER/Regional Government of Andalusia-Ministry of Economy and Knowledge. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Availability of Data and Materials

The dataset and code used for meta-analysis are available from the corresponding author [DMV] on request.

## Declarations

### Ethics Approval and Consent to Participate

Not applicable.

### Consent for Publication

Not applicable.

### Competing Interests

Carolina Casado-Robles, Jesús Viciano, Santiago Guijarro-Romero and Daniel Mayorga-Vega declare that they have no conflicts of interest relevant to the content of this review.

### Author details

<sup>1</sup>Department of Physical Education and Sport, University of Granada, Granada, Spain. <sup>2</sup>Department of Didactic of Musical, Plastic and Corporal Expression, University of Jaen, Paraje de las Lagunillas, Campus de Las Lagunillas, Edificio Humanidades y Ciencias de la Educación I (D2-D136), 23071 Jaén, Spain.

Received: 31 August 2021 Accepted: 29 December 2021

Published online: 31 January 2022

## References

- Al-Hazzaa HM, Abahussain NA, Al-Sobayel HI, Qahwaji DM, Musaiger AO. Physical activity, sedentary behaviors and dietary habits among Saudi adolescents relative to age, gender and region. *Int J Behav Nutr Phys Act.* 2011;8:140. <https://doi.org/10.1186/1479-5868-8-140>.
- Longmuir PE, Colley RC, Wherley VA, Tremblay MS. Canadian society for exercise physiology position stand: benefit and risk for promoting childhood physical activity. *Appl Physiol Nutr Metab.* 2014;39:1271. <https://doi.org/10.1139/apnm-2014-0074>.
- Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep.* 1985;100:126–31.
- World Health Organization. WHO guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020.
- Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, et al. Sedentary Behavior Research Network (SBRN)—terminology consensus project process and outcome. *Int J Behav Nutr Phys Act.* 2017;14:75. <https://doi.org/10.1186/s12966-017-0525-8>.
- Poitras VJ, Gray CE, Borghese MM, Carson V, Chaput J-P, Janssen I, et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Appl Physiol Nutr Metab Physiol Appl Nutr Metab.* 2016;41:S197–239. <https://doi.org/10.1139/apnm-2015-0663>.
- Carson V, Hunter S, Kuzik N, Gray CE, Poitras VJ, Chaput JP, et al. Systematic review of sedentary behaviour and health indicators in school-aged children and youth: an update. *Appl Physiol Nutr Metab.* 2016;41:S240–65. <https://doi.org/10.1139/apnm-2015-0630>.
- Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *Lancet Child Adolesc Health.* 2020;4:23–35. [https://doi.org/10.1016/S2352-4642\(19\)30323-2](https://doi.org/10.1016/S2352-4642(19)30323-2).
- Arundell L, Fletcher E, Salmon J, Veitch J, Hinkley T. A systematic review of the prevalence of sedentary behavior during the after-school period among children aged 5–18 years. *Int J Behav Nutr Phys Act.* 2016;13:93. <https://doi.org/10.1186/s12966-016-0419-1>.
- World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. Geneva: World Health Organization; 2018.
- Dobbins M, Husson H, DeCorby K, LaRocca RL. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database Syst Rev.* 2013;2:CD007651. <https://doi.org/10.1002/14651858.CD007651.pub2>.
- Lonsdale C, Rosenkranz RR, Peralta LR, Bennie A, Fahey P, Lubans DR. A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical

- education lessons. *Prev Med.* 2013;56:152–61. <https://doi.org/10.1016/j.ypmed.2012.12.004>.
13. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol.* 2009;28:690–701. <https://doi.org/10.1037/a0016136>.
  14. Strath SJ, Rowley TW. Wearables for promoting physical activity. *Clin Chem.* 2018;64:53–63. <https://doi.org/10.1373/clinchem.2017.272369>.
  15. Ruiz DC, Goransson A. Professional android wearables. New York: Wiley; 2015.
  16. Franssen WMA, Franssen GHLM, Spaas J, Solmi F, Eijnde BO. Can consumer wearable activity tracker-based interventions improve physical activity and cardiometabolic health in patients with chronic diseases? A systematic review and meta-analysis of randomised controlled trials. *Int J Behav Nutr Phys Act.* 2020. <https://doi.org/10.1186/s12966-020-00955-2>.
  17. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. *JMIR MHealth UHealth.* 2016. <https://doi.org/10.2196/mhealth.4225>.
  18. International Data Corporation. Shipments of Wearable Devices Leap to 125 Million Units, Up 35.1% in the Third Quarter, According to IDC. 2020. <https://www.idc.com/getdoc.jsp?containerId=prUS47067820>. Accessed 15 Jul 2021.
  19. Tudor-Locke C, Craig CL, Beets MW, Belton S, Cardon GM, Duncan S, et al. How many steps/day are enough? For children and adolescents. *Int J Behav Nutr Phys Act.* 2011;8:78. <https://doi.org/10.1186/1479-5868-8-78>.
  20. Bronikowski M, Bronikowska M, Glapa A. Do they need goals or support? A report from a goal-setting intervention using physical activity monitors in youth. *Int J Environ Res Public Health.* 2016;13:9. <https://doi.org/10.3390/ijerph13090914>.
  21. Guagliano JM, Armitage SM, Brown HE, Coombes E, Fusco F, Hughes C, et al. A whole family-based physical activity promotion intervention: findings from the families reporting every step to health (FRESH) pilot randomised controlled trial. *Int J Behav Nutr Phys Act.* 2020;17:120. <https://doi.org/10.1186/s12966-020-01025-3>.
  22. Cooper H, Hedges LV, Valentine JC. The handbook of research synthesis and meta-analysis. 3rd ed. New York: Russell Sage Foundation; 2019.
  23. Schmidt FL, Hunter JE. Methods of meta-analysis: correcting error and bias in research findings. 3rd ed. London: SAGE Publications; 2021.
  24. Brickwood K-J, Watson G, O'Brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. *JMIR MHealth UHealth.* 2019. <https://doi.org/10.2196/11819>.
  25. Tang MSS, Moore K, McGavigan A, Clark RA, Ganesan AN. Effectiveness of wearable trackers on physical activity in healthy adults: systematic review and meta-analysis of randomized controlled trials. *JMIR MHealth UHealth.* 2020. <https://doi.org/10.2196/15576>.
  26. Ringeval M, Wagner G, Denford J, Paré G, Kitsiou S. Fitbit-based interventions for healthy lifestyle outcomes: systematic review and meta-analysis. *J Med Internet Res.* 2020;22: e23954. <https://doi.org/10.2196/23954>.
  27. Ha L, Mizrahi D, Wakefield CE, Cohn RJ, Simar D, Signorelli C. The use of activity trackers in interventions for childhood cancer patients and survivors: a systematic review. *J Adolesc Young Adult Oncol.* 2020. <https://doi.org/10.1089/jayao.2020.0099>.
  28. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. *JMIR MHealth UHealth.* 2016. <https://doi.org/10.2196/mhealth.6540>.
  29. Böhm B, Karwiese SD, Böhm H, Oberhoffer R. Effects of mobile health including wearable activity trackers to increase physical activity outcomes among healthy children and adolescents: systematic review. *JMIR MHealth UHealth.* 2019;7: e8298. <https://doi.org/10.2196/mhealth.8298>.
  30. Cajita MI, Kline CE, Burke LE, Bigini EG, Imes CC. Feasible but not yet efficacious: a scoping review of wearable activity monitors in interventions targeting physical activity, sedentary behavior, and sleep. *Curr Epidemiol Rep.* 2020;7:25–38. <https://doi.org/10.1007/s40471-020-00229-2>.
  31. Creaser AV, Clemes SA, Costa S, Hall J, Ridgers ND, Barber SE, et al. The acceptability, feasibility, and effectiveness of wearable activity trackers for increasing physical activity in children and adolescents: a systematic review. *Int J Environ Res Public Health.* 2021;18:6211. <https://doi.org/10.3390/ijerph18126211>.
  32. Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet.* 1997;350:326–9. [https://doi.org/10.1016/S0140-6736\(97\)02419-7](https://doi.org/10.1016/S0140-6736(97)02419-7).
  33. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess.* 2003;7:1–90. <https://doi.org/10.3310/hta7410>.
  34. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. *Cochrane handbook for systematic reviews of interventions.* 6.2. New York: Wiley; 2021.
  35. Prince SA, Cardilli L, Reed JL, Saunders TJ, Kite C, Douillette K, et al. A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act.* 2020;17:31. <https://doi.org/10.1186/s12966-020-00938-3>.
  36. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med.* 2003;37:197–206. <https://doi.org/10.1136/bjsm.37.3.197>.
  37. Liu JY-W, Kor PPK, Chan CPY, Kwan RYC, Sze-Ki D. The effectiveness of a wearable activity tracker (WAT)-based intervention to improve physical activity levels in sedentary older adults: a systematic review and meta-analysis. *Arch Gerontol Geriatr.* 2020. <https://doi.org/10.1016/j.archger.2020.104211>.
  38. Sucharew H. Methods for research evidence synthesis: the scoping review approach. *J Hosp Med.* 2019;14:416. <https://doi.org/10.12788/jhm.3248>.
  39. Borenstein MM, Hedges L, Higgins J, Rothstein H. *Introduction to meta-analysis.* 2nd ed. New York: Wiley; 2009.
  40. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372: n71. <https://doi.org/10.1136/bmj.n71>.
  41. World Health Organization. *The health of youth.* Geneva: World Health Organization; 1989.
  42. Bromberg S. *Consumer wristband activity monitors as a simple and inexpensive tool for remote heart failure monitoring.* Toronto: University of Toronto; 2015.
  43. Gaudet J, Gallant F, Bélanger M. A bit of fit: minimalist intervention in adolescents based on a physical activity tracker. *JMIR MHealth UHealth.* 2017. <https://doi.org/10.2196/mhealth.7647>.
  44. Horne PJ, Hardman CA, Lowe CF, Rowlands AV. Increasing children's physical activity: a peer modelling, rewards and pedometer-based intervention. *Eur J Clin Nutr.* 2009;63:191–8. <https://doi.org/10.1038/sj.ejcn.1602915>.
  45. Routen A, Upton D, Edwards M, Peters D. The effect of pedometer step goal, feedback and self-monitoring interventions on accelerometer-measured physical activity in children. *Grad J Sport Exerc Phys Educ Res.* 2014;2:37–53.
  46. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol.* 2011;64:383–94. <https://doi.org/10.1016/j.jclinepi.2010.04.026>.
  47. Mayorga-Vega D, Bocanegra-Parrilla R, Ornelas M, Viciano J. Criterion-related validity of the distance- and time-based walk/run field tests for estimating cardiorespiratory fitness: a systematic review and meta-analysis. *PLoS ONE.* 2016;11: e0151671. <https://doi.org/10.1371/journal.pone.0151671>.
  48. Cohen J. A power primer. *Psychol Bull.* 1992;112:155–9. <https://doi.org/10.1037/0033-2909.112.1.155>.
  49. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315:629. <https://doi.org/10.1136/bmj.315.7109.629>.
  50. Orwin RG. A fail-safe N for effect size in meta-analysis. *J Educ Stat.* 1983;8:157–9. <https://doi.org/10.2307/1164923>.

51. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56:455–63. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>.
52. Remmert JE, Woodworth A, Chau L, Schumacher LM, Butryn ML, Schneider M. Pilot trial of an acceptance-based behavioral intervention to promote physical activity among adolescents. *J Sch Nurs*. 2019;35:449–61. <https://doi.org/10.1177/1059840518786782>.
53. Leinonen A-M, Pyky R, Ahola R, Kangas M, Siirtola P, Luoto T, et al. Feasibility of gamified mobile service aimed at physical activation in young men: population-based randomized controlled study (mopo). *JMIR MHealth UHealth*. 2017. <https://doi.org/10.2196/mhealth.6675>.
54. Baldursdóttir B, Taehinen RE, Sigfusdóttir ID, Krettek A, Valdimarsdóttir HB. Impact of a physical activity intervention on adolescents' subjective sleep quality: a pilot study. *Glob Health Promot*. 2017;24:14–22. <https://doi.org/10.1177/1757975915626112>.
55. Ermetici F, Zelaschi RF, Briganti S, Dozio E, Gaeta M, Ambrogi F, et al. Association between a school-based intervention and adiposity outcomes in adolescents: the Italian "eAT" project. *Obesity*. 2016;24:687–95. <https://doi.org/10.1002/oby.21365>.
56. Bronikowski M, Bronikowska M, Maciaszek J, Glapa A. Maybe it is not a goal that matters: a report from a physical activity intervention in youth. *J Sports Med Phys Fitness*. 2018;58:348–55. <https://doi.org/10.23736/S0022-4707.16.06611-1>.
57. Kantanista A, Bronikowski M, Laudanska-Krzeminska I, Osinski W. Effects of pedometer-based walking intervention on psychological and biological variables in adolescent girls: does achieving goals make a difference? *Med Sport*. 2014;67:261–72.
58. Shore SM. The H.Y.P.P.E. initiative: a school-based physical activity promotion program. Ann Arbor: Temple University; 2010.
59. Rothstein H, Sutton A, Borenstein M. *Publication bias in meta-analysis: prevention, assessment and adjustments*. New York: Wiley; 2006.
60. Eyre ELJ, Cox VM, Birch SL, Duncan MJ. An integrated curriculum approach to increasing habitual physical activity in deprived South Asian children. *Eur J Sport Sci*. 2016;16:381–90. <https://doi.org/10.1080/17461391.2015.1062565>.
61. Lubans DR, Morgan PJ, Aguiar EJ, Callister R. Randomized controlled trial of the Physical Activity Leaders (PALs) program for adolescent boys from disadvantaged secondary schools. *Prev Med*. 2011;52:239–46. <https://doi.org/10.1016/j.ypmed.2011.01.009>.
62. Jago R, Baranowski T, Baranowski K, Thompson D, Cullen KW, Watson K, et al. Fit for Life Boy Scout badge: outcome evaluation of a troop and Internet intervention. *Prev Med*. 2006;42:181–7. <https://doi.org/10.1016/j.ypmed.2005.12.010>.
63. Aksornsri A. A pedometer-physical activity program to change self-efficacy and physical activity in Thai school-Age children. Ann Arbor: Saint Louis University; 2019.
64. Baldursdóttir B. *Physical activity and well-being among adolescents*. Gothenburg: University of Gothenburg; 2016.
65. Corepal R, Best P, O'Neill R, Kee F, Badham J, Dunne L, et al. A feasibility study of "The StepSmart Challenge" to promote physical activity in adolescents. *Pilot Feasibility Stud*. 2019. <https://doi.org/10.1186/s40814-019-0523-5>.
66. Corr M, McMullen J, Morgan P, Barnes A, Murtagh E. Supporting our lifelong engagement: mothers and teens exercising (SOLE MATES); a feasibility trial. *Women Health*. 2019. <https://doi.org/10.1080/03630242.2019.1688446>.
67. Corr M, Murtagh E. 'No one ever asked us': a feasibility study assessing the co-creation of a physical activity programme with adolescent girls. *Glob Health Promot*. 2020;27:34–43. <https://doi.org/10.1177/1757975919853784>.
68. Duck AA, Hall KC, Klamm M, Temple M, Robinson JC. Physical activity and fitness: the feasibility and preliminary effectiveness of wearable activity tracker technology incorporating altruistic motivation in youth. *J Spec Pediatr Nurs*. 2021;26: e12313. <https://doi.org/10.1111/jspn.12313>.
69. Duncan M, Birch S, Woodfield L. Efficacy of an integrated school curriculum pedometer intervention to enhance physical activity and to reduce weight status in children. *Eur Phys Educ Rev*. 2012;18:396–407. <https://doi.org/10.1177/1356336X12450799>.
70. Evans EW, Abrantes AM, Chen E, Jelalian E. Using novel technology within a school-based setting to increase physical activity: a pilot study in school-age children from a low-income, Urban Community. *BioMed Res Int*. 2017. <https://doi.org/10.1155/2017/4271483>.
71. Finkelstein EA, Tan YT, Malhotra R, Lee CF, Goh SS, Saw SM. A cluster randomized controlled trial of an incentive-based outdoor physical activity program. *J Pediatr*. 2013;163:167–72. <https://doi.org/10.1016/j.jpeds.2013.01.009>.
72. Galy O, Yacef K, Caillaud C. Improving pacific adolescents' physical activity toward international recommendations: exploratory study of a digital education app coupled with activity trackers. *JMIR MHealth UHealth*. 2019;7: e14854. <https://doi.org/10.2196/14854>.
73. Grao-Cruces A, Ruiz-López R, Moral-García J-E, Ruiz-Ariza A, Martínez-López EJ. Effects of a steps/day programme with evaluation in physical education on body mass index in schoolchildren 11–12 years of age. *Kinesiology*. 2016;48:132–41. <https://doi.org/10.26582/k.48.1.2>.
74. Groffik D, Frömel K, Pelclová J. Pedometers as a method for modification of physical activity in students. *J Hum Kinet*. 2008;20:131–8. <https://doi.org/10.2478/v10078-008-0025-7>.
75. Hardman CA, Horne PJ, Lowe CF. A home-based intervention to increase physical activity in girls: the Fit "n" Fun Dudes program. *J Exerc Sci Fit*. 2009;7:1–8. [https://doi.org/10.1016/S1728-869X\(09\)60001-0](https://doi.org/10.1016/S1728-869X(09)60001-0).
76. Hardman CA, Horne PJ, Fergus LC. Effects of rewards, peer-modelling and pedometer targets on children's physical activity: a school-based intervention study. *Psychol Health*. 2011;26:3–21. <https://doi.org/10.1080/08870440903318119>.
77. Jauho A-M, Pyky R, Ahola R, Kangas M, Virtanen P, Korpelainen R, et al. Effect of wrist-worn activity monitor feedback on physical activity behavior: a randomized controlled trial in Finnish young men. *Prev Med Rep*. 2015;2:628–34. <https://doi.org/10.1016/j.pmedr.2015.07.005>.
78. Kantanista A, Bronikowski M, Laudanska-Krzeminska I, Krol-Zielinska M, Osinski W. Positive effect of pedometer-based walking intervention on body image and physical activity enjoyment in adolescent girls. *Biomed Hum Kinet*. 2017;9:34–42. <https://doi.org/10.1515/bhk-2017-0006>.
79. Kerner C, Burrows A, McGrane B. Health wearables in adolescents: implications for body satisfaction, motivation and physical activity. *Int J Health Promot Educ*. 2019;57:191–202. <https://doi.org/10.1080/14635240.2019.1581641>.
80. Larsen B, Benitez T, Cano M, Dunsiger SS, Marcus BH, Mendoza-Vasconez A, et al. Web-based physical activity intervention for latina adolescents: feasibility, acceptability, and potential efficacy of the Niñas Saludables Study. *J Med Internet Res*. 2018;20: e170. <https://doi.org/10.2196/jmir.9206>.
81. Linck DT. *Female teens step it up with the Fitbit Zip: a randomized controlled pilot study*. Augusta: Augusta University; 2018.
82. Lubans D, Morgan P. Evaluation of an extra-curricular school sport programme promoting lifestyle and lifetime activity for adolescents. *J Sports Sci*. 2008;26:519–29. <https://doi.org/10.1080/02640410701624549>.
83. Lubans DR, Morgan PJ, Callister R, Collins CE. Effects of integrating pedometers, parental materials, and E-mail support within an extracurricular school sport intervention. *J Adolesc Health*. 2009;44:176–83. <https://doi.org/10.1016/j.jadohealth.2008.06.020>.
84. Macias-Cervantes MH, Malacara JM, Garay-Sevilla ME, Díaz-Cisneros FJ. Effect of recreational physical activity on insulin levels in Mexican/Hispanic children. *Eur J Pediatr*. 2009;168:1195–202. <https://doi.org/10.1007/s00431-008-0907-7>.
85. Morris JL, Daly-Smith A, Defeyter MA, McKenna J, Zwolinsky S, Lloyd S, et al. A pedometer-based physically active learning intervention: the importance of using preintervention physical activity categories to assess effectiveness. *Pediatr Exerc Sci*. 2019;31:356–62. <https://doi.org/10.1123/pes.2018-0128>.
86. Newton RLJ, Marker AM, Allen HR, Machtmes R, Han H, Johnson WD, et al. Parent-targeted mobile phone intervention to increase physical activity in sedentary children: randomized pilot trial. *JMIR MHealth UHealth*. 2014;2: e48. <https://doi.org/10.2196/mhealth.3420>.
87. Schofield L, Mummery WK, Schofield G. Effects of a controlled pedometer-intervention trial for low-active adolescent girls. *Med Sci Sports Exerc*. 2005;37:1414–20. <https://doi.org/10.1249/01.mss.0000174889.89600.e3>.
88. Shapiro JR, Bauer S, Hamer RM, Kordy H, Ward D, Bulik CM. Use of text messaging for monitoring sugar-sweetened beverages, physical

- activity, and screen time in children: a pilot study. *J Nutr Educ Behav*. 2008;40:385–91. <https://doi.org/10.1016/j.jneb.2007.09.014>.
89. Shimon JM, Petlichkoff LM. Impact of pedometer use and self-regulation strategies on junior high school physical education students' daily step counts. *J Phys Act Health*. 2009;6:178–84. <https://doi.org/10.1123/jpah.6.2.178>.
  90. Shore SM, Sachs ML, DuCette JP, Libonati JR. Step-count promotion through a school-based intervention. *Clin Nurs Res*. 2014;23:402–20. <https://doi.org/10.1177/1054773813485240>.
  91. Smith JJ, Morgan PJ, Plotnikoff RC, Dally KA, Salmon J, Okely AD, et al. Smart-phone obesity prevention trial for adolescent boys in low-income communities: the ATLAS RCT. *Pediatrics*. 2014;134:e723–31. <https://doi.org/10.1542/peds.2014-1012>.
  92. Thompson D, Cantu D, Ramirez B, Cullen KW, Baranowski T, Mendoza J, et al. Texting to increase adolescent physical activity: feasibility assessment. *Am J Health Behav*. 2016;40:472–83. <https://doi.org/10.5993/AJHB.40.4.9>.
  93. Wang S-H. The effects of goal setting on female middle school students' physical activity levels and motivation toward exercise. *Ann Arbor: The Florida State University*; 2004.
  94. Zizzi S, Vitullo E, Rye J, O'hara-Tompkins N, Abildso C, Fisher B, et al. Impact of a three-week pedometer intervention on high school students' daily step counts and perceptions of physical activity. *Am J Health Educ*. 2006;37:35–40. <https://doi.org/10.1080/19325037.2006.10598875>.
  95. Mayorga-Vega D, Casado-Robles C, López-Fernández I, Viciana J. A comparison of the utility of different step-indices to translate the physical activity recommendation in adolescents. *J Sports Sci*. 2021;39(4):469–79. <https://doi.org/10.1080/02640414.2020.1826667>.
  96. Hall KS, Hyde ET, Bassett DR, Carlson SA, Carnethon MR, Ekelund U, et al. Systematic review of the prospective association of daily step counts with risk of mortality, cardiovascular disease, and dysglycemia. *Int J Behav Nutr Phys Act*. 2020;17:78. <https://doi.org/10.1186/s12966-020-00978-9>.
  97. Ramakrishnan R, He J-R, Ponsonby A-L, Woodward M, Rahimi K, Blair SN, et al. Objectively measured physical activity and all cause mortality: a systematic review and meta-analysis. *Prev Med*. 2021;143:106356. <https://doi.org/10.1016/j.jpmed.2020.106356>.
  98. 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. <https://doi.org/10.2196/jmir.1533>.
  99. Carlin A, Murphy MH, Gallagher AM. Current influences and approaches to promote future physical activity in 11–13 year olds: a focus group study. *BMC Public Health*. 2015;15:1270. <https://doi.org/10.1186/s12889-015-2601-9>.
  100. Telama R, Yang X, Leskinen E, Kankaanpää A, Hirvensalo M, Tammelin T, et al. Tracking of physical activity from early childhood through youth into adulthood. *Med Sci Sports Exerc*. 2014. <https://doi.org/10.1249/MSS.000000000000181>.
  101. Tudor-Locke C, Craig CL, Thyfault JP, Spence JC. A step-defined sedentary lifestyle index: < 5000 steps/day. *Appl Physiol Nutr Metab*. 2013;38:100–14. <https://doi.org/10.1139/apnm-2012-0235>.
  102. Katapally TR, Muhajarine N. Towards uniform accelerometry analysis: a standardization methodology to minimize measurement bias due to systematic accelerometer wear-time variation. *J Sports Sci Med*. 2014;13:379.
  103. Mayorga-Vega D, Martínez-Baena A, Viciana J. Does school physical education really contribute to accelerometer-measured daily physical activity and non sedentary behaviour in high school students? *J Sports Sci*. 2018;36:1913–22. <https://doi.org/10.1080/02640414.2018.1425967>.
  104. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Med*. 2017;47:1821–45. <https://doi.org/10.1007/s40279-017-0716-0>.
  105. Owen KB, Smith J, Lubans D, Ng J, Lonsdale C. Self-determined motivation and physical activity in children and adolescents: a systematic review and meta-analysis. *Prev Med*. 2014;67:270–9. <https://doi.org/10.1016/j.jpmed.2014.07.033>.
  106. Mayorga-Vega D, Viciana J. Las clases de educación física solo mejoran la capacidad cardiorespiratoria de los alumnos con menor condición física: un estudio de intervención controlado. *Nutr Hosp*. 2015;32:330–5. <https://doi.org/10.3305/nh.2015.32.1.8919>.
  107. Guijarro-Romero S, Mayorga-Vega D, Casado-Robles C, Viciana J. An intermittent physical fitness teaching unit only improves cardiorespiratory fitness levels of students with an unhealthy physical fitness profile. *Retos*. 2020;38:8–15. <https://doi.org/10.47197/retos.v38i38.73605>.
  108. Love RE, Adams J, van Sluijs EMF. Equity effects of children's physical activity interventions: a systematic scoping review. *Int J Behav Nutr Phys Act*. 2017;14:134. <https://doi.org/10.1186/s12966-017-0586-8>.
  109. Riddoch CJ, Bo Andersen L, Wedderkopp N, Harro M, Klasson-Heggebo L, Sardinha LB, et al. Physical activity levels and patterns of 9- and 15-year-old European children. *Med Sci Sports Exerc*. 2004;36:86–92. <https://doi.org/10.1249/01.MSS.0000106174.43932.92>.
  110. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. *Med Sci Sports Exerc*. 2000;32:963–75. <https://doi.org/10.1097/00005768-200005000-00014>.
  111. Bandura A. Health promotion by Social Cognitive Means. *Health Educ Behav*. 2004;31:143–64. <https://doi.org/10.1177/1090198104263660>.
  112. Ryan RM, Deci EL. Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. *Contemp Educ Psychol*. 2020;61: 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>.
  113. Rhodes RE, Janssen I, Bredin SSD, Warburton DER, Bauman A. Physical activity: health impact, prevalence, correlates and interventions. *Psychol Health*. 2017;32:942–75. <https://doi.org/10.1080/08870446.2017.1325486>.
  114. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, 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*. 2013;46:81–95. <https://doi.org/10.1007/s12160-013-9486-6>.
  115. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. *Sports Med*. 2014;44:671–86. <https://doi.org/10.1007/s40279-014-0142-5>.
  116. Cairney J, Dudley D, Kwan M, Bulten R, Kriellaars D. Physical literacy, physical activity and health: toward an evidence-informed conceptual model. *Sports Med*. 2019;49:371–83. <https://doi.org/10.1007/s40279-019-01063-3>.
  117. Bassett DR, Toth LP, LaMunion SR, Crouter SE. Step counting: a review of measurement considerations and health-related applications. *Sports Med*. 2017;47:1303–15. <https://doi.org/10.1007/s40279-016-0663-1>.
  118. Richardson S, Mackinnon D. Becoming your own device: self-tracking challenges in the workplace. *Can J Sociol*. 2018;43:265–90. <https://doi.org/10.29173/cjs28974>.
  119. Fairclough SJ, Noonan R, Rowlands AV, Van Hees V, Knowles Z, Boddy LM. Wear compliance and activity in children wearing wrist- and hip-mounted accelerometers. *Med Sci Sports Exerc*. 2016. <https://doi.org/10.1249/MSS.0000000000000771>.
  120. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *Evid Based Med*. 2016;21:125. <https://doi.org/10.1136/ebmed-2016-110401>.
  121. Balslem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64:401–6. <https://doi.org/10.1016/j.jclinepi.2010.07.015>.

## Publisher's Note

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