Preventive Medicine 105 (2017) 237–249

Contents lists available at ScienceDirect

Preventive Medicine

journal homepage: www.elsevier.com/locate/ypmed

Review Article

The effectiveness of school-based physical activity interventions for adolescent girls: A systematic review and meta-analysis

Michael B. Owen[a,*], Whitney B. Curry[a], Charlotte Kerner[b], Lisa Newson[c], Stuart J. Fairclough[a,d]

[a] Physical Activity and Health Research Group, Department of Sport and Physical Activity, Edge Hill University, St Helens Road, Ormskirk, Lancashire, UK
[b] Department of Life Sciences, Brunel University, London, UK
[c] Natural Sciences and Psychology, Research Centre for Brain and Behaviour, Liverpool John Moores University, Liverpool, UK
[d] Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland

ARTICLE INFO

Keywords:
Adolescents
Girls
School
Physical activity
Intervention

ABSTRACT

Physical activity (PA) decreases during the transition from childhood to adolescence, with larger declines observed in girls. School-based interventions are considered the most promising approach for increasing adolescents' PA levels although, it is unclear which types of school-based interventions have the greatest impact. The objective of this systematic review is to assess the impact and design of school-based PA interventions targeting adolescent girls. A systematic search was conducted using four electronic databases (PubMed, Web of Science, SPORTDiscus and PsychInfo). This systematic review was registered with PROSPERO (Registration number: CRD42016037428) and PRISMA guidelines (2009) were followed throughout. Twenty studies were identified as meeting the inclusion criteria and were included in a narrative synthesis. Seventeen studies were eligible for inclusion in a meta-analysis. There was a significant small positive treatment effect for school-based PA interventions for adolescent girls ($k = 17$, $g = 0.37$, $p < 0.05$). After an outlier was removed (residual $z = 7.61$) the average treatment effect was significantly reduced, indicating a very small positive effect ($k = 16$, $g = 0.07$, $p = 0.05$). Subgroup analysis revealed very small significant effects for multi-component interventions ($k = 7$, $g = 0.09$, $p < 0.05$), interventions underpinned by theory ($k = 12$, $g = 0.07$, $p < 0.05$), and studies with a higher risk of bias ($k = 13$, $g = 0.09$, $p < 0.05$). Intervention effects were very small which indicates that changing PA behaviors in adolescent girls through school-based interventions is challenging. Multi-component interventions and interventions underpinned by theory may be the most effective approaches to positively change adolescent girls' PA.

1. Introduction

The World Health Organisation (2014) has classified physical inactivity as the fourth leading risk factor for global mortality from non-communicable diseases. Insufficient physical activity (PA) contributes towards 3.2 million deaths (5.5%) worldwide per year (World Health Organisation, 2014). A strong body of evidence indicates that regular moderate-to-vigorous physical activity (MVPA) is associated with numerous health benefits for children and young people (Chief Medical Officers, 2011). These include reduced body fat and the promotion of healthy weight, enhanced cardio-metabolic and bone health, and enhanced psychological well-being (Biddle and Asare, 2011; Janssen and Leblanc, 2010).

Though the benefits and protective effects of regular PA are well understood, insufficient PA during adolescence is a major concern (Heitzler et al., 2011; Khunti et al., 2007; Sisson et al., 2010). Inactive adolescents are more at risk of being overweight or obese and have a greater chance of developing type 2 diabetes (World Health Organisation, 2015). Additionally, physical inactivity is a major risk factor for not only poor physical health but is also associated with poor mental wellbeing (Ar-yuwat et al., 2013). More frequent engagement in PA contributes towards greater well-being and lower levels of anxiety and depressive symptoms in both sexes (McMahon et al., 2017).

According to global estimates of self-reported PA, 80% of 13–15-year-olds do not engage in 60 min of MVPA per day, with girls being less active than boys (Hallal et al., 2012). A combination of biological and psychosocial factors put adolescent girls at risk of inactivity and uptake of sedentary lifestyles (Young et al., 2014). A review of 26 longitudinal studies concluded that there was a 7% decrease in total PA per year during adolescence (Dumith et al., 2011), with the most recent studies indicating that girls' PA levels declined at a greater rate than boys'. Research assessing objectively measured PA from the

* Corresponding author.
E-mail address: Michael.Owen@edgehill.ac.uk (M.B. Owen).

http://dx.doi.org/10.1016/j.ypmed.2017.09.018
Received 26 April 2017; Received in revised form 13 September 2017; Accepted 24 September 2017
Available online 28 September 2017

International Children's Accelerometry Database (ICAD) suggests that boys were more active than girls but, both boys' and girls' MVPA levels declined steadily through adolescence (Cooper et al., 2015). There is no widely accepted explanation for this decrease in adolescent girls. However, it is suggested that alongside biological changes, lack of enjoyment, negative experiences in, and perceptions of school-based PA may be important factors (Barr-Anderson et al., 2008).

Previous systematic reviews (Camacho-Minano et al., 2011; Voskuil et al., 2017) and a meta-analysis (Pearson et al., 2015) have assessed interventions to promote PA in adolescent girls across school and community settings. Voskuil et al. (2017) reported highly variable effect sizes, inferring that PA interventions only had a small effect on objectively measured PA in girls aged 6–18 years (Voskuil et al., 2017). Camacho-Minano et al. (2011) found overall mixed results regarding the effectiveness of PA interventions for adolescent girls but, suggested that multicomponent school-based interventions, which included PE that addressed the unique needs of girls were the most effective. Pearson et al. (2015) reported small but significant effects ($g = 0.35$, $p < 0.001$) for the effectiveness of PA interventions on girls aged 12 to 18 years. Larger effects were found for interventions which were underpinned by theory, school-based, girls only, targeted younger adolescents (ages 12 to 15), multicomponent in design, and that targeted both PA and sedentary behaviour.

Camacho-Minano et al. (2011) and Pearson et al. (2015) suggested that school-based PA interventions are the most promising setting to impact adolescent girls' PA levels. Thus, this review aims to address this gap in the literature and assess the effectiveness of girl-specific and mixed-sex school-based interventions on adolescent girls' PA. The inclusion of mixed-sex studies is novel because often reviews (Camacho-Minano et al., 2011; Voskuil et al., 2017) focus only on interventions exclusively designed for girls, when mixed-sex interventions could be equally as effective for girls. The purpose of this study was to systematically review school-based PA interventions involving adolescent girls and quantify their effect through meta-analysis.

2. Methods

This systematic review was registered with PROSPERO (Registration number: CRD42016037428). This review adhered to the PRISMA reporting guidelines for systematic reviews (Moher et al., 2009).

2.1. Search procedure

A systematic search was conducted using four electronic databases (PubMed, Web of Science, SPORTDiscus and PsychInfo). Journal articles published in English post 31/12/2004 until the date of the last search (01/12/16) were considered for review. The key words included; physical activity, physical education, sedentary behaviour, sedentary time, walking, sport, fitness, energy expenditure, school, teacher, classroom, gymnasium, sports hall, recess, playtime, break time, playground, before-school and after-school. The search strategies are detailed in the supplementary information (Supplementary Table 1). Reference lists of retrieved articles were examined for additional articles.

2.2. Inclusion and exclusion criteria

Studies were eligible if they reported the effects of school-based PA interventions on PA outcomes among adolescent girls (mean age 11–18 years), with the primary outcome being objectively measured or self-reported PA levels. Feasibility and pilot studies were included. Mixed sexed studies were included if girls' data were presented separately to boys' or if girls' data were received upon request. A school-based intervention was defined as one that occurred in the school environment. The extended school day (8 am–6 pm) was used to operationally define the school day, so as to capture school-based interventions that took place before and after formal hours (e.g., breakfast clubs, boot camps, and after-school activities). Studies could be randomised or non-randomised and only published peer-reviewed studies were reviewed. Only journal articles published post 31/12/2004 were considered after preliminary searches ('physical activity' AND 'girls' AND 'intervention') indicated that most interventions had been conducted in the last 10 years with the earliest published in 2004.

All search results were exported into a reference manager (Endnote × 7.4, Thomson Reuters) and duplicates were removed. Initially, the first author (MO) screened all titles and abstracts for obvious irrelevance, and a random sample (20%) were also checked by another author (WC). The full-text of eligible studies were then retrieved and reviewed by two authors (MO and WC). Where full texts were not readily available, the lead author was contacted and asked to provide the full text for further assessment on eligibility. If no response was received after a follow-up reminder, these studies were excluded as they could not be fully assessed for eligibility. Any disagreements were resolved in a meeting involving three authors (MO, WC, and SF).

2.3. Data extraction and synthesis

Relevant data from the selected studies were extracted by the first author (MO) and checked by the second author (WC) (see Table 1). If studies reported multiple PA outcomes, data for the primary outcome stated in the studies' aims and objectives were used. Any disagreements were resolved through a consensus discussion between MO and WC. A narrative synthesis was completed to provide a summary of school-based PA interventions for adolescent girls (11–18).

2.4. Risk of bias assessment

Included studies were assessed for risk of bias using a modified tool (Morton et al., 2016; Pluye et al., 2009) appropriate for PA reviews which include measures for quantitative experimental and quantitative observational studies. This adapted risk of bias assessment tool (Supplementary Table 2) used a 1–4 scoring system (i.e., 1 = weak, 2 = moderate, 3 = strong and 4 = very strong) at study level as a combined risk of bias score. A higher risk of bias score indicates better methodological quality with a lower risk of bias score indicating poorer methodological quality. Risk of bias was scored on the presence or absence of each criteria respectively (sequence generation and/or randomisation, concealment and/or blinding, complete outcome data and/or low withdrawal/drop-out (< 20%), appropriate outcome measure). Studies were scored on what was reported in the current article or if they cited a previously published protocol paper which was examined for further information.

2.5. Meta-analysis

Meta-analytic procedures were conducted in R (https://cran.r-

Table 1
Data extraction procedure.

Study characteristics	(a) Author, year of publication, country (b) Aims and objectives of study (c) Participant characteristics (d) Study design (e) Intervention content
Theory underpinning intervention	(f) Any theory or model that the authors suggest underpins the intervention, including non-behaviour change theories
PA measurement tool	(g) Any measurement tool used to collect PA data, including outcome measure of PA
Primary PA findings	(h) Key findings of each study in relation to PA change due to the intervention

project.org) using the metafor package (Viechtbauer, 2010). Studies were included in the meta-analysis if they employed a pre-post control group design. Pre-post intervention PA levels were used as few studies included post-intervention follow up data. The meta-analyses effect size selected was Hedge's *g*, which provides a correction factor for smaller sample sizes ($k < 20$). Meta-analyses were conducted using random effects models to reflect the likelihood of different effect sizes underlying the studies due to the diversity of the included interventions and their implementation (Borenstein et al., 2010). Heterogeneity was assessed using Cochrane's Q-statistic and I^2 (Higgins et al., 2003). The Q-statistic and corresponding p value provide a calculation of variance between study effects. A significant Q value indicates systematic differences between the individual studies which might influence the results. I^2 is represented as a percentage with a value of 0% indicating no dispersion and larger values indicating gradual increases in heterogeneity (i.e., 25% = low, 50% = moderate, 75% = high level of heterogeneity (Higgins et al., 2003). Subgroup analyses were performed on possible moderators of the average intervention effect. These were: physical activity measurement method (objective vs. self-report), intervention duration (short vs long), risk of bias (*/** vs. ***/****), intervention design (single component vs. multi-component), presence of underpinning theory (yes vs. no), and the target sex (girls only vs. mixed sex).

Outliers were identified to evaluate the influence of extreme values on the overall treatment effect. Studies with an inflated residual value approximately two standard deviations ($z = \pm 1.98$) above or below the average treatment effect were considered outliers. Publication bias was estimated by examining asymmetry of funnel plots (effect size vs. standard error) where asymmetry is indicative of publication bias (Sterne and Egger, 2001). Following these visual inspections, the trim and fill procedure (Duval and Tweedie, 2000a, 2000b), Orwin's fail safe number (Orwin, 1983) and Egger's regression test (Egger et al., 1997) were used to confirm the presence or absence of publication bias.

3. Results

3.1. Literature search

In total, 9383 records were identified. After screening and eligibility assessments, 20 records met the inclusion criteria for the narrative synthesis (Fig. 1).

3.2. Participant characteristics

Table 2 provides an overview of participant and study characteristics. In this review, the 20 studies evaluated a total sample of 10,755 girls across the interventions (Mean age = 12.88 years). Four studies reported mixed samples where girls' data were extracted (Bronikowski and Bronikowska, 2011; Haerens et al., 2006; How et al., 2013; Loucaides et al., 2009), with the remaining sixteen studies including girls only samples. The majority of studies were with girls aged 11–14 years, with only three studies (Dudley et al., 2010; Schofield et al., 2005; Taymoori et al., 2008) involving girls aged 15–17 years. Nine studies recruited girls only with no set eligibility criteria stated (Bronikowski and Bronikowska, 2011; Dewar et al., 2014; Fairclough and Stratton, 2005; Haerens et al., 2006; How et al., 2013; Huberty et al., 2014; Jago et al., 2015; Jago et al., 2012; Loucaides et al., 2009; Martin and Fairclough, 2008; Pate et al., 2005). For the remaining eleven studies, four were mixed-sex interventions but reported boys' and girls' PA outcomes separately (Bronikowski and Bronikowska, 2011; Haerens et al., 2006; How et al., 2013; Loucaides et al., 2009). Two studies stated that girls had to be enrolled in two semesters of PE (Jones et al., 2008; Young et al., 2006), two targeted low active girls (Robbins et al., 2006; Schofield et al., 2005), one targeted girls with low PA enjoyment (Dudley et al., 2010), one targeted girls at the preparation stage of exercise behaviour change, and one targeted girls who did not meet national recommendations for MVPA (Robbins et al., 2012). Seventeen studies contained participant numbers < 1000, with the smallest sample being 15 participants (Martin and Fairclough, 2008). Three studies contained > 1000 participants (Haerens et al., 2006; Pate et al., 2005; Webber et al., 2008), with the largest sample being 3502 participants (Webber et al., 2008).

3.3. Study characteristics

Eight studies were conducted in the USA (Huberty et al., 2014; Jones et al., 2008; Pate et al., 2005; Robbins et al., 2006; Robbins et al., 2012; Spruijt-Metz et al., 2008; Webber et al., 2008; Young et al., 2006), with four studies from the UK (Fairclough and Stratton, 2005; Jago et al., 2015; Jago et al., 2012; Martin and Fairclough, 2008), and four from Australia (Dewar et al., 2014; Dudley et al., 2010; How et al., 2013; Schofield et al., 2005). There were: fourteen randomised controlled trials (RCTs) (Bronikowski and Bronikowska, 2011; Dudley et al., 2010; Haerens et al., 2006; How et al., 2013; Jago et al., 2015; Jago et al., 2012; Jones et al., 2008; Pate et al., 2005; Robbins et al., 2006; Spruijt-Metz et al., 2008; Taymoori et al., 2008; Webber et al., 2008; Young et al., 2006) including three cluster RCTs (Dewar et al., 2014; Jago et al., 2015; Jago et al., 2012), and one pilot RCT (Dudley et al., 2010); five quasi-experimental studies (Fairclough and Stratton, 2005; Loucaides et al., 2009; Martin and Fairclough, 2008; Robbins et al., 2012; Schofield et al., 2005); and one case-crossover study (Huberty et al., 2014). Five studies had PA measurement periods of 12 to 36 months (Bronikowski and Bronikowska, 2011; Dewar et al., 2014; Haerens et al., 2006; Jones et al., 2008; Webber et al., 2008), including two which utilised a long-term follow-up (i.e., ≥ 12 months) after the cessation of the intervention (Bronikowski and Bronikowska, 2011; Dewar et al., 2014). Eight studies had PA measurement periods of 5 to 12 months (Huberty et al., 2014; Jago et al., 2015; Jago et al., 2012; Pate et al., 2005; Robbins et al., 2012; Spruijt-Metz et al., 2008; Taymoori et al., 2008; Young et al., 2006), including four studies that incorporated short-term follow ups (i.e., ≤ 6 months post-end of intervention) (Huberty et al., 2014; Jago et al., 2015; Spruijt-Metz et al., 2008; Taymoori et al., 2008). Seven studies had measurement periods that were < 4 months and did not include follow-up measurements (Dudley et al., 2010; Fairclough and Stratton, 2005; How et al., 2013; Loucaides et al., 2009; Martin and Fairclough, 2008; Robbins et al., 2006; Schofield et al., 2005). Eight studies were published since 2010 (Bronikowski and Bronikowska, 2011; Dewar et al., 2014; Dudley et al., 2010; How et al., 2013; Huberty et al., 2014; Jago et al., 2015; Jago et al., 2012; Robbins et al., 2012).

3.4. Intervention characteristics

Ten studies reported multi-component interventions (Dewar et al., 2014; Haerens et al., 2006; Huberty et al., 2014; Jones et al., 2008; Pate et al., 2005; Robbins et al., 2006; Robbins et al., 2012; Taymoori et al., 2008; Webber et al., 2008; Young et al., 2006). Components included school environment adaptions, modified PE lessons, extra-curricular PA sessions, educational sessions, counselling sessions, and provision of further opportunities to be physically active (e.g., lunch and break time PA clubs). Ten studies reported single-component interventions. Four of these were modified PE lessons (Bronikowski and Bronikowska, 2011; Dudley et al., 2010; Fairclough and Stratton, 2005; How et al., 2013; Martin and Fairclough, 2008), three were after-school dance interventions (Jago et al., 2015; Jago et al., 2012), two were educational-based interventions (Schofield et al., 2005; Spruijt-Metz et al., 2008) and one was a modified playground intervention (Loucaides et al., 2009). Eighteen of the interventions provided an opportunity for the participants to engage in PA, such as modified active PE lessons, lunchtime PA sessions and after-school PA clubs. Twelve of the interventions incorporated an educational component. Ten interventions lasted for < 4 months in total duration (Dudley et al., 2010; Fairclough and

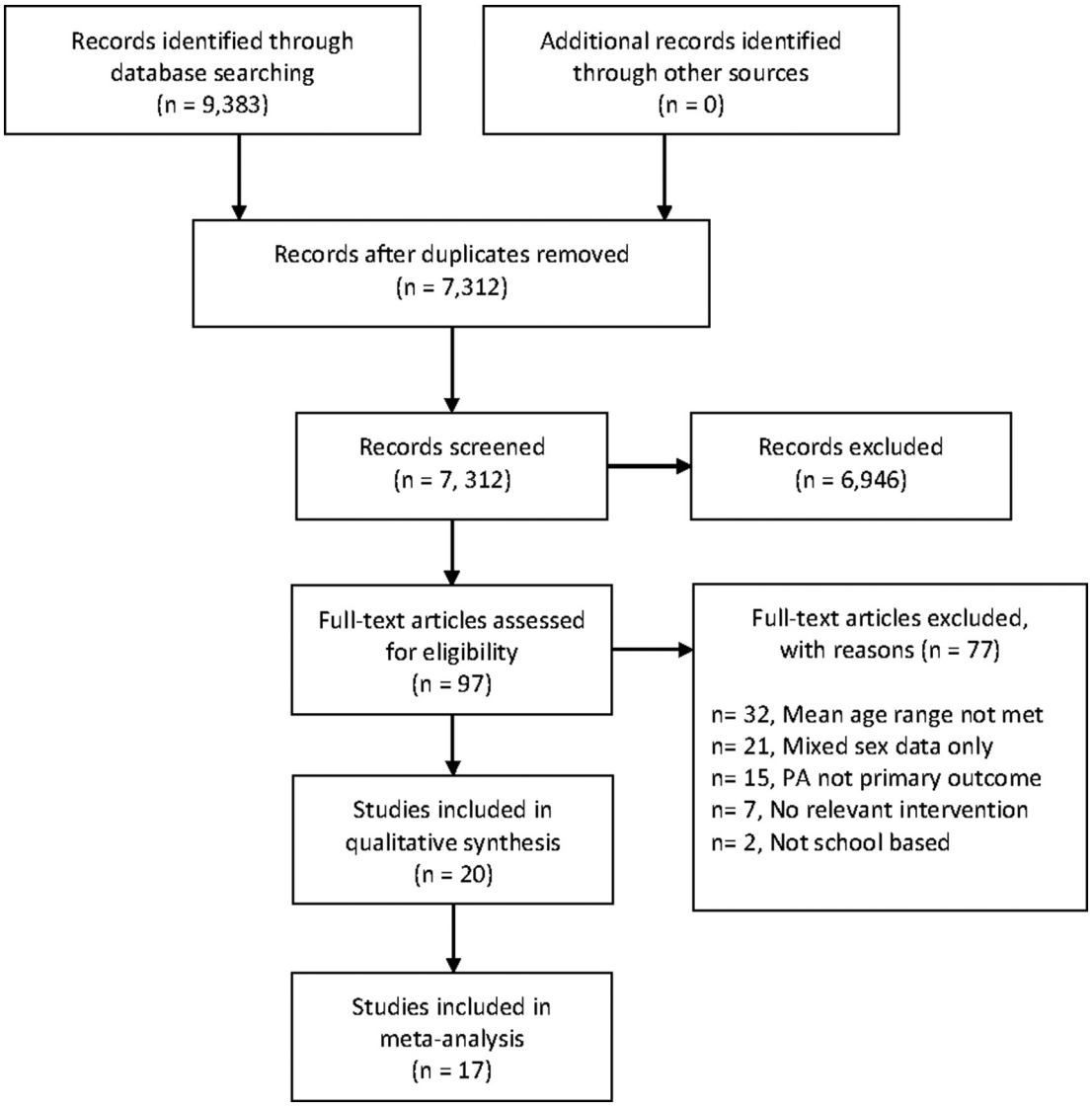

Fig. 1. PRISMA flow diagram (Moher et al., 2009) to show each stage of the systematic eligibility process.

Stratton, 2005; How et al., 2013; Huberty et al., 2014; Jago et al., 2015; Loucaides et al., 2009; Martin and Fairclough, 2008; Robbins et al., 2006; Schofield et al., 2005; Spruijt-Metz et al., 2008), with the shortest intervention period being reported as 5–7 days (Spruijt-Metz et al., 2008). Five interventions lasted 6–10 months (Jago et al., 2012; Pate et al., 2005; Robbins et al., 2012; Taymoori et al., 2008; Young et al., 2006), and five lasted for 12–36 months (Bronikowski and Bronikowska, 2011; Dewar et al., 2014; Haerens et al., 2006; Jones et al., 2008; Webber et al., 2008).

3.5. Intervention delivery

Thirteen of the interventions were delivered by school staff including PE teachers (Bronikowski and Bronikowska, 2011; Dewar et al., 2014; Dudley et al., 2010; Fairclough and Stratton, 2005; Haerens et al., 2006; How et al., 2013; Huberty et al., 2014; Jones et al., 2008; Martin and Fairclough, 2008; Pate et al., 2005; Spruijt-Metz et al., 2008; Webber et al., 2008; Young et al., 2006). Two were delivered by dance instructors (Jago et al., 2015; Jago et al., 2012), who taught dance-specific sessions. Two were delivered by a research team (Schofield et al., 2005; Taymoori et al., 2008), one was delivered by the school nurse and physical activity club instructors (Robbins et al., 2012), and one was delivered through a combination of an online advice programme, a paediatric nurse and a phone-based research assistant (Robbins et al., 2006). One intervention was a playground modification which had no direct deliverer (Loucaides et al., 2009).

3.6. Outcome measures

Five methods were used to measure PA (Table 2). PA was objectively measured with accelerometers in ten studies (Dewar et al., 2014; Dudley et al., 2010; Haerens et al., 2006; How et al., 2013; Huberty et al., 2014; Jago et al., 2015; Jago et al., 2012; Martin and Fairclough, 2008; Robbins et al., 2012; Webber et al., 2008), and subjectively measured through self-report questionnaires in nine studies (Bronikowski and Bronikowska, 2011; Dewar et al., 2014; Haerens et al., 2006; Jones et al., 2008; Pate et al., 2005; Robbins et al., 2006; Spruijt-Metz et al., 2008; Taymoori et al., 2008; Young et al., 2006). Two studies combined self-report and accelerometers (Dewar et al., 2014; Haerens et al., 2006), one study used pedometers (Loucaides et al., 2009), one study combined pedometers and self-reported PA (Schofield et al., 2005), and one study used heart rate (HR) and direct observation (Fairclough and Stratton, 2005). Seven out of eight studies published from 2010 onwards utilised accelerometers (Dewar et al., 2014; Dudley et al., 2010; How et al., 2013; Huberty et al., 2014; Jago et al., 2015; Jago et al., 2012; Robbins et al., 2012). Eight out of twelve

Table 2
Study characteristics, and key findings from each intervention.

Study	Design & country	Underpinning theory	Participants	Intervention duration & measurement period	PA measurement method & PA outcome measure	Key findings
1. Bronikowski and Bronikowska (2011)	RCT, Poland	Hellison's model of teaching responsibility through PA	n = 170; mean age = 13.22 (0.3) Mixed sex study	15-Month intervention & 30-month study from baseline post-intervention (month 15 to follow-up (month 30).	Self-report & frequency of weekly leisure-time PA	Significantly increased trends in the frequency of undertaking leisure time PA in INT groups for girls ($p < 0.01$), differences sustained in the 15-month follow-up after cessation of the intervention.
2. Dewar et al. (2014)	Cluster RCT, Australia	Social cognitive theory	n = 357; mean age = 13.2 (0.5) Girls only study	12-month intervention & 24-month study from baseline to post-intervention (12 months) and follow-up (month 24).	Accelerometry and self-report & % of MVPA per valid day	No observed improvements for PA levels. Self-report data shows girls in the INT group had a significantly greater reduction in sedentary activities (− 56.4 min/day; $p < 0.05$).
3. Dudley et al. (2010)	Pilot RCT, Australia	Social cognitive theory	n = 38; mean age = 16.5 (0.2) Girls with low levels of PA enjoyment only	11-week intervention & 3 month study from baseline to post-intervention.	Accelerometry & Accelerometry counts	There was a non-significant smaller decline in participation in PA during school sport for INT group compared to CON group.
4. Fairclough and Stratton (2005)	Quasi-experimental design, England	None specified	n = 26; mean age = 12.4 (0.4) Girls only study	5-week intervention & 6-week study from baseline to post-intervention.	Direct observation and HR monitor & % of lesson time in MVPA	INT group engaged in significantly more MVPA in PE lesson than those in the CON lesson (18.5% vs 13.5%; $p < 0.05$). INT group engaged in MVPA for an average of 11.9% more lesson time than the CON group.
5. Haerens et al. (2006)	RCT, Belgium	The theory of planned behaviour and the trans-theoretical model	n = 1039; mean age = 13.1 (0.8) Mixed sex study	24-Month intervention & 24-month study from baseline to 12 and 24 post-baseline.	Accelerometry and self-report & minutes of total PA per day	Time spent in PA of light intensity decreased significantly less for girls in the INT groups (− 2 min/day) compared with the CON group (− 20 min/day, $p < 0.05$) at 2 years post-baseline.
6. How et al. (2013)	RCT, Australia	Self-determination theory	n = 125; age = year 8 (13–14 years) Mixed sex study	15-week intervention & 15-week study from pre-intervention to post-intervention	Accelerometry & % of lesson time in MVPA	Girls who chose option 3 INT group (24.5%; design own lessons based on advice/guidelines) were significantly (both $p < 0.01$) more physically active for a greater percentage of time than CON (19.1%; standard lesson) and option 2 INT group (16.5%; 'PE development officer')
7. Huberty et al. (2014)	Case-crossover design, USA	None specified	n = 59; mean age = 11.3 (0.7) Girls only study	12-week intervention & 7/8 month study from baseline to mid-1, mid-2, post-intervention and 3 months follow-up.	Accelerometry & Total MVPA minutes per day	INT group was associated with a statistically significant ($p < 0.05$) increase in MVPA compared to CON group for girls aged 11–13 years = (1.5 min, 95CI 0.4 to 2.6)
8. Jago et al. (2012)	Cluster RCT, England	None specified	n = 203; Age = Year 7 (11–12 years) Girls only study	9-week intervention & 5-month study from baseline to time 1 (week 8 or 9) and time 2 (3 months follow-up).	Accelerometry & Weekday MVPA minutes	At time 1 there was a − 6.8 difference in MVPA week day minutes for the INT group compare to the CON incentive group (95CI 18 to 4). At time 2 there was an 8.7 difference between INT group compared to CON incentive group (95CI 6 to 12). Wide confidence intervals suggest potential positive but not significant intervention effects.
9. Jago et al. (2015)	Cluster RCT, England	Self-determination theory	n = 571; Age = Year 7 (11–12 years) Girls only study	8-Month intervention & 12-month study from baseline to time 1 (17–20 weeks) and time 2 (52 weeks).	Accelerometry & Weekday MVPA minutes	No evidence that the after school dance programme had any significant effect on weekday MVPA levels, overall PA or PA during the afterschool period. However, during the afterschool period on dance

(*continued on next page*)

Table 2 (*continued*)

Study	Design & country	Underpinning theory	Participants	Intervention duration & measurement period	PA measurement method & PA outcome measure	Key findings
						days versus non-dance days' girls obtained 15 min more LPA, 4.7 min more MVPA and 258 more accelerometer counts.
10. Jones et al. (2008)	RCT, USA	Social cognitive theory and the trans-theoretical model	n = 718; mean age = 11.6 (0.4) Girls only study must be enrolled in 2 semesters of PE	18-month intervention & 18-month study from baseline to interim-intervention (month 6/7/8) to follow-up (month 18).	Self-report & total MVPA minutes per day	INT group had higher means for overall total daily minutes of PA and daily MVPA minutes at follow-up compared to CON group. But, only total daily minutes of VPA were significantly higher at follow-up for INT (difference = 6 min, 95% CI = 5.82–6.18, p = 0.05) compared to CON. A 45.4% increase in VPA minutes from baseline for INT group (CON = 4.1% decrease).
11. Loucaides et al. (2009)	Quasi-experimental design, Cyprus	None specified	n = 114; mean age = 11.1 (0.3) Mixed sex study	4-week intervention & 5–6 week study from pre-intervention and 4 weeks post-intervention.	Pedometer & Step count	Small but non-significant increase in mean steps observed during 20-min break period in INT 2 school 852 (384) to 1004 (525) from baseline to post INT. compared to slight decreases in both CON 1055 (421) to 962 (466) and INT 1 school 1224 (403) to 1150(339).
12. Martin and Fairclough (2008)	Quasi-experimental design, England	None specified	n = 15; age = year 7 (11–12 years) Girls only study	4-week intervention & 8-week study from pre-intervention (1–4 weeks) to post-intervention (week 8).	Accelerometry & % of lesson time in MVPA	Girls engaged in MVPA pre-INT for 29.7% (16.6 min) of lesson time, which increased to 34.9% (19.3 min) during intervention lessons ($p < 0.05$).
13. Pate et al. (2005)	RCT, USA	Social cognitive theory	n = 2744; mean age = 13.6 (0.6) Girls only study	8–10 month intervention (1 school year) & 12-month study from baseline (spring 8th grade) to follow-up (spring 9th grade).	Self-report & 30-min blocks of MVPA per day	Increases observed in self-reported ≥ two 30 min blocks of MVPA per day for INT group from baseline to post INT 68.6% to 72.0% but, results were not significant. However, there were significant differences in the percentage of girls who reported regular VPA in the INT group compared to the CON group (44.5% vs 36.4%). A significant increase of 8% ($p < 0.05$).
14. Robbins et al. (2006)	RCT, USA	Pender's health promotion model and the trans-theoretical model	n = 77; Age = Grade 6–8 (11–14 years) Low active girls only study	12-week intervention & 12-week study from baseline (week 1) to post-intervention (week 12).	Self-report & minutes in MPA plus VPA	No differences in self-reported PA between the INT and CON groups. Both the INT group and CON group showed increases in minutes in MPA plus VPA across 2 weekdays and 2 weekend days but these were non-significant.
15. Robbins et al. (2012)	Quasi-experimental design, USA	Pender's health promotion model	n = 69; mean age = 11.44 (0.7) Girls < MVPA national recommendations only	6-month intervention & 6-month study from baseline to 6 months follow-up.	Accelerometry & Minutes of MVPA per hour	No statistically significant differences in PA levels for minutes of MVPA per hour for the INT or CON group. But, the differences were in the expected direction, with the INT group having slightly higher improvement in minutes of MVPA per hour (0.43) compared to CON group (0.07) from baseline to 6 months follow-up.
16. Schofield et al. (2005)	Quasi-experimental design, Australia	None specified	n = 85; mean age = 15.8 Low active girls only	12-week intervention & 12-week study from pre-intervention, mid-intervention (week 6) and post-intervention (week 12).	Pedometer and Self-Report & Step count	Pedometer INT group significantly increased their total PA (Avg mean daily steps increase of 2747), when compared

(*continued on next page*)

Table 2 (*continued*)

Study	Design & country	Underpinning theory	Participants	Intervention duration & measurement period	PA measurement method & PA outcome measure	Key findings
						with the CON group ($p < 0.05$) at post-INT.
17. Spruijt-Metz et al. (2008)	RCT, USA	Self-determination theory and the theory of meanings of behaviour	n = 459; mean age = 12.47 (0.6) Girls only study	5–7 day intervention & 6–7 month intervention from baseline (3 months prior to intervention) to follow up (3 months post-intervention).	Self-report & 30-min blocks of activity of various intensities	No significant effects on PA of any intensity; VPA, MVPA MPA or LPA. However, the intervention had a significant effect on reducing time spent on SB ($p < 0.05$).
18. Taymoori et al. (2008)	RCT, Iran	Pender's health promotion model and trans-theoretical model	n = 161; mean age = 14.79 (0.4) Girls at preparation stage of exercise behaviour change only	6-month intervention & 12-month intervention from pre-intervention to post-intervention (month 6) to 6 month follow-up (month 12).	Self-report & minutes of total PA per day	TTM and HP group increased mean minutes of PA per day from 27.16 (12.02) at pre-INT to 75.80 (27.52) at post-INT with a slight drop to 60.04 (24.87) at follow-up (both $p < 0.05$) compared to CON group. This was similar for the solely HP group increasing from 28.56 (11.30) to 73.61(28.73) at post-INT with a drop to 56.79 (27.58) at follow-up (both $p < 0.05$) compared to CON group.
19. Webber et al. (2008)	RCT, USA	Operant learning theory, social cognitive theory, organizational change theory and the diffusion of innovation model in a social-ecologic framework	n = 3504; age = grades 6–8 (11–14 years) Girls only study	36-month intervention & 36-month study from baseline to month 24 post-intervention (staff directed) to month 36 (program champion) post-intervention.	Accelerometry & Average daily minutes of MET-weighted minutes of MVPA	After the 2 years staff-directed INT, there were no differences (mean = − 0.4, 95% CI = CI = − 8.2 to 7.4) in adjusted MET-weighted minutes of MVPA between 8th-grade girls in schools assigned to INT or CON groups. However, significant differences were found between INT and CON groups after an additional year of program champion delivered intervention (INT group 10.9 min more MET-weighted MVPA, $p < 0.05$).
20. Young et al. (2006)	RCT, USA	Social action theory	n = 221; mean age = 13.8 (0.5) Girls only study must be enrolled in 2 semesters of PE	8-month intervention & 9-month study from baseline to follow-up (month 8 or 9).	Self-report & estimated total energy expenditure	INT classes spent 46.9% of PE class time in MVPA compared with 30.5% of the time for control classes ($p < 0.001$). However, while the INT was successful in increasing MVPA in PE class, no changes were observed in overall, daily, moderate, or hard to very hard mean energy expenditure in either the INT or the CON group.

Notes. CON = control, INT = intervention, PA = physical activity, SB = sedentary behaviour, MVPA = moderate to vigorous physical activity, VPA = vigorous physical activity, LPA = light physical activity, MPA = moderate physical activity, min = minutes, PE = physical education, TTM = trans-theoretical model, HP = health promotion, RCT = randomised control trial.
For mixed sex studies participant characteristics are shown for girls only.

studies published from 2005 to 2010 used self-reported measures of PA (Haerens et al., 2006; Jones et al., 2008; Pate et al., 2005; Robbins et al., 2006; Schofield et al., 2005; Spruijt-Metz et al., 2008; Taymoori et al., 2008; Young et al., 2006).

Thirteen different units of measurement were used to report a change in PA levels (Table 2). Studies reported percentage of lesson time in MVPA (Fairclough and Stratton, 2005; How et al., 2013; Martin and Fairclough, 2008), weekday MVPA minutes (Jago et al., 2015; Jago et al., 2012), total week MVPA minutes per day (Huberty et al., 2014; Jones et al., 2008), self-reported 30 min blocks of activity (Pate et al., 2005; Spruijt-Metz et al., 2008), minutes of total PA per day (Haerens et al., 2006; Taymoori et al., 2008), MVPA per hour (Robbins et al., 2012), total MVPA percentage per valid day (Dewar et al., 2014), average daily minutes of MET-weighted minutes of MVPA (Webber et al., 2008), minutes in MPA plus VPA (Robbins et al., 2006), estimated total energy expenditure (Young et al., 2006), accelerometer counts (Dudley et al., 2010), self-reported frequency of weekly leisure-time PA (Bronikowski and Bronikowska, 2011), and step counts (Loucaides et al., 2009; Schofield et al., 2005). As thirteen different units of measurement were used to assess PA, from this point onwards, changes in PA across groups of studies with different units of measurement, will be referred to as 'activity'.

3.7. Behaviour change theories

Thirteen studies explicitly reported that the interventions incorporated one or more behaviour change theories. These were Social Cognitive Theory (Dewar et al., 2014; Dudley et al., 2010; Jones et al., 2008; Pate et al., 2005; Webber et al., 2008), The Theory of Planned Behaviour (Haerens et al., 2006), Trans-theoretical Model (Haerens et al., 2006; Jones et al., 2008; Robbins et al., 2006; Taymoori et al., 2008), Self-Determination Theory (How et al., 2013; Jago et al., 2015; Spruijt-Metz et al., 2008), Pender's Health Promotion Model (Robbins et al., 2006; Robbins et al., 2012; Taymoori et al., 2008), Theory of Meanings Behaviour (Spruijt-Metz et al., 2008), and The Social Action Theory (Young et al., 2006). The largest study (Webber et al., 2008) incorporated numerous theories within a Socio-ecologic Framework, including Operant Learning Theory, Social Cognitive Theory, Organisational Change Theory, and The Diffusion of Innovation Model. One study (Bronikowski and Bronikowska, 2011) used Hellison's Model of Teaching Responsibility through PA. Only five of the studies underpinned by behaviour change theory lasted 12 months or longer. The remaining six studies, which used relatively modest sample sizes ($n \leq 203$) did not specify the use of a behaviour change model or theory (Fairclough and Stratton, 2005; Huberty et al., 2014; Jago et al., 2012; Loucaides et al., 2009; Martin and Fairclough, 2008; Schofield et al., 2005).

3.8. Risk of Bias (Table 3)

Fifteen studies provided outcome data with < 20% dropout/withdrawal rates. Thirteen studies employed objective measures of PA, either for the complete sample size or for a sub-sample. Only seven of the included studies described the randomisation processes. Although eleven studies stated a randomisation procedure, the majority ($n = 10$) did not provide an explicit explanation of the randomisation process (Bronikowski and Bronikowska, 2011; Haerens et al., 2006; How et al., 2013; Loucaides et al., 2009; Martin and Fairclough, 2008; Pate et al., 2005; Robbins et al., 2012; Schofield et al., 2005; Spruijt-Metz et al., 2008; Webber et al., 2008; Young et al., 2006), which led to their poor randomisation scores. All studies scored weakly for allocation of concealment and/or blinding, with just two studies attempting to blind intervention staff (Jago et al., 2012; Webber et al., 2008). Only one study received a 'very strong' risk of bias score (Jago et al., 2012); three studies received a 'strong' risk of bias score (Fairclough and Stratton, 2005; Jago et al., 2015; Webber et al., 2008); ten studies received a 'moderate' risk of bias score (Dewar et al., 2014; Dudley et al., 2010; How et al., 2013; Huberty et al., 2014; Jones et al., 2008; Loucaides et al., 2009; Martin and Fairclough, 2008; Robbins et al., 2006; Robbins et al., 2012; Taymoori et al., 2008), six studies received a 'weak' risk of bias score (Bronikowski and Bronikowska, 2011; Haerens et al., 2006; Pate et al., 2005; Schofield et al., 2005; Spruijt-Metz et al., 2008; Young

Table 3
Risk of bias assessment.

	Study	Appropriate sequence generation and/or randomisation	Allocation concealment and/or blinding	Complete outcome data and/or low withdrawal/drop-out	Appropriate outcome measure (PA)	Risk of bias score
1.	Bronikowski and Bronikowska (2011)			X		*
2.	Dewar et al. (2014)	X			X	**
3.	Dudley et al. (2010)	X			X	**
4.	Fairclough and Stratton (2005)	X		X	X	***
5.	Haerens et al. (2006)				X	*
6.	How et al. (2013)			X	X	**
7.	Huberty et al. (2014)			X	X	**
8.	Jago et al. (2012)	X		X	X	***
9.	Jago et al. (2015)	X	X	X	X	****
10.	Jones et al, (2008)	X		X		**
11.	Loucaides et al. (2009)			X	X	**
12.	Martin and Fairclough (2008)			X	X	**
13.	Pate et al. (2005)			X		*
14.	Robbins et al. (2006)	X		X		**
15.	Robbins et al. (2012)			X	X	**
16.	Schofield et al. (2005)				X	*
17.	Spruijt-Metz et al. (2008)			X		*
18.	Taymoori et al. (2008)	X		X		**
19.	Webber et al. (2008)		X	X	X	***
20.	Young et al. (2006)			X		*

* = Weak.
** = Moderate.
*** = Strong.
**** = Very strong.

Table 4
Sub-group analyses.

Subgroup variables	Effect size statistics				Null test	Heterogeneity statistics		Publication bias
	k	*g*	SE	95% CI	Z	Q	I^2	Eggers' z
Pooled effect	17	0.37	0.19	0.0008, 0.73	1.96*	80.12**	94.91%	2.05*
Pooled effect[a]	16	0.07	0.04	−0.002, 0.14	1.92=	23.98	0.01%	2.07*
Measurement method								
Objective	10	0.16	0.14	−0.11, 0.43	1.14	17.92*	55.62%	1.71
Self-report[a]	6	0.08	0.04	−0.002, 0.16	1.92=	5.81	0.04%	1.57
Study duration								
Short (< 6 months)	8	0.22	0.14	−0.06, 0.50	1.53	15.01*	56.92%	1.75
Long[a] (> 6 months)	8	0.06	0.04	−0.02, 0.14	1.51	8.84	0.00%	0.76
Risk of bias								
* or **[a]	13	0.09	0.04	0.02, 0.17	2.37*	16.67	0.00%	1.93=
*** or ****	3	0.01	0.23	−0.44, 0.46	0.05	4.99	65.95%	1.30
Study design								
Single	9	0.02	0.06	−0.09, 0.14	0.41	11.83	0.00%	2.13*
Multi[a]	7	0.09	0.04	0.006, 0.18	2.09*	11.30	0.02%	1.47
Theory included								
Yes[a]	12	0.07	0.04	0.0009, 0.15	1.98*	18.35	0.01%	2.11*
No	4	0.06	0.20	−0.33, 0.45	0.31	5.38	38.75%	2.25*
INT gender target								
Girls only[a]	13	0.06	0.04	−0.02, 0.13	1.53	19.35	0.03%	1.61
Mixed	3	0.28	0.17	−0.05, 0.61	1.65	2.64	20.27%	0.92

Note. k = number of effect sizes. *g* = effect size (Hedges' *g*). SE = standard error. 95% CI = confidence intervals (lower limit, upper limit). Z = test of null hypothesis. Q = test of variance between effects sizes. I^2 = total variance unexplained by moderator. Eggers' z = test of publication bias.

[a] Outlier removed from subgroup.

= p equal to 0.05.

* $p < 0.05$.

** $p < 0.001$.

et al., 2006).

3.9. Meta-analysis

Of the 20 studies included in the narrative synthesis 17 provided sufficient data for inclusion in the meta-analysis. Huberty et al. (2014) was excluded for not reporting sample size, Martin and Fairclough (2008) did not use a control group and Webber et al. (2008) did not report variance of data. Cohen's (1988) effect size criteria were used to interpret the overall treatment effect for the main analysis and subgroup analyses. Of the 17 included studies, 12 reported a small effects ($g = -0.29$ to 0.26), four studies reported moderate to strong effects ($g = 0.65$ to 1.04) and one reported a very strong effect size ($g = 3.43$) (Taymoori et al., 2008). The meta-analysis revealed a significant small positive treatment effect (k = 17, $g = 0.37$, $p < 0.05$,) for school-based PA interventions for adolescent girls (Table 4). Heterogeneity analysis indicated significant between-study variance ($Q = 80.12$, $p < 0.001$; $I^2 = 94.91\%$). The Taymoori et al. (2008) intervention was identified as an outlier due to large residual effects (z = 7.61). Once this study was removed the average treatment effect was significantly reduced by 0.30, indicating a very small positive effect which approached significance (k = 16, $g = 0.07$, $p = 0.05$) (Fig. 2). Heterogeneity was also substantially reduced when the outlier was removed ($Q = 23.98$, $p = 0.05$; $I^2 = 0.01\%$).

Inspection of the funnel plot for publication bias indicated asymmetry. The trim and fill procedure added 3 studies to the left side of the plot which reduced the overall treatment effect by 0.01. Orwin's fail-safe N calculation suggested that there would need to be 16 unpublished studies to reduce the treatment effect to a target effect size of $g = 0.11$, and Egger's regression test was significant (z = 2.07, $p < 0.05$). Collectively, these results indicated a high probability of publication bias.

Although heterogeneity from the pooled analysis was low, the individual effects from the included studies were extremely inconsistent, ranging from $g = -0.29$ to 1.04. Thus, subgroup analyses were performed as planned to explore whether the identified subgroups moderated the average intervention effect. The identified outlier study was removed from the relevant subgroups in all analyses. Significant effects were observed for studies with * or ** bias ratings (k = 13, $g = 0.09$, $p < 0.05$), for multi-component interventions (k = 7, $g = 0.09$, $p < 0.05$), and for interventions underpinned by theory (k = 12, $g = 0.07$, $p < 0.05$) but the magnitudes of these were small (Table 4). Subgroup analyses also revealed no effect for whether the interventions were targeted at girls only or mixed-sex, although only 3 mixed sex studies were included.

The three studies excluded from the meta-analysis all indicated positive results. Huberty et al. (2014) found that on the days the after school club was delivered the intervention group significantly increased MVPA by 1.5 min compared to the control group (non-afterschool club). Martin and Fairclough (2008) found that girls increased their percentage of lesson time MVPA by 5.2% (2.7 min) from non-intervention lessons to intervention lessons. Webber et al. (2008) found no significant differences after 2 years of the staff directed intervention. However, after a further year of program champion delivered intervention, girls had significantly more MET-weighted minutes of MVPA (10.9) compared to girls in the control school.

4. Discussion

This systematic review and meta-analysis examined the effect of school-based PA interventions on PA outcomes among adolescent girls. The meta-analysis results indicate that school-based PA interventions have only a very small effect on adolescent girls' PA levels. Some individual studies showed positive results and the subgroup analyses revealed promise for approaches underpinned by theory and multi-component interventions. Although school-based interventions have been suggested as being the most promising setting to intervene with adolescent girls (Camacho-Minano et al., 2011; Pearson et al., 2015), the observed small effect illustrates the difficulties and challenges of positively impacting adolescent girls' PA behaviors through the school setting. These difficulties may in part be due to a number of factors such as, social or cultural norms, ability to provide a wide range of PA

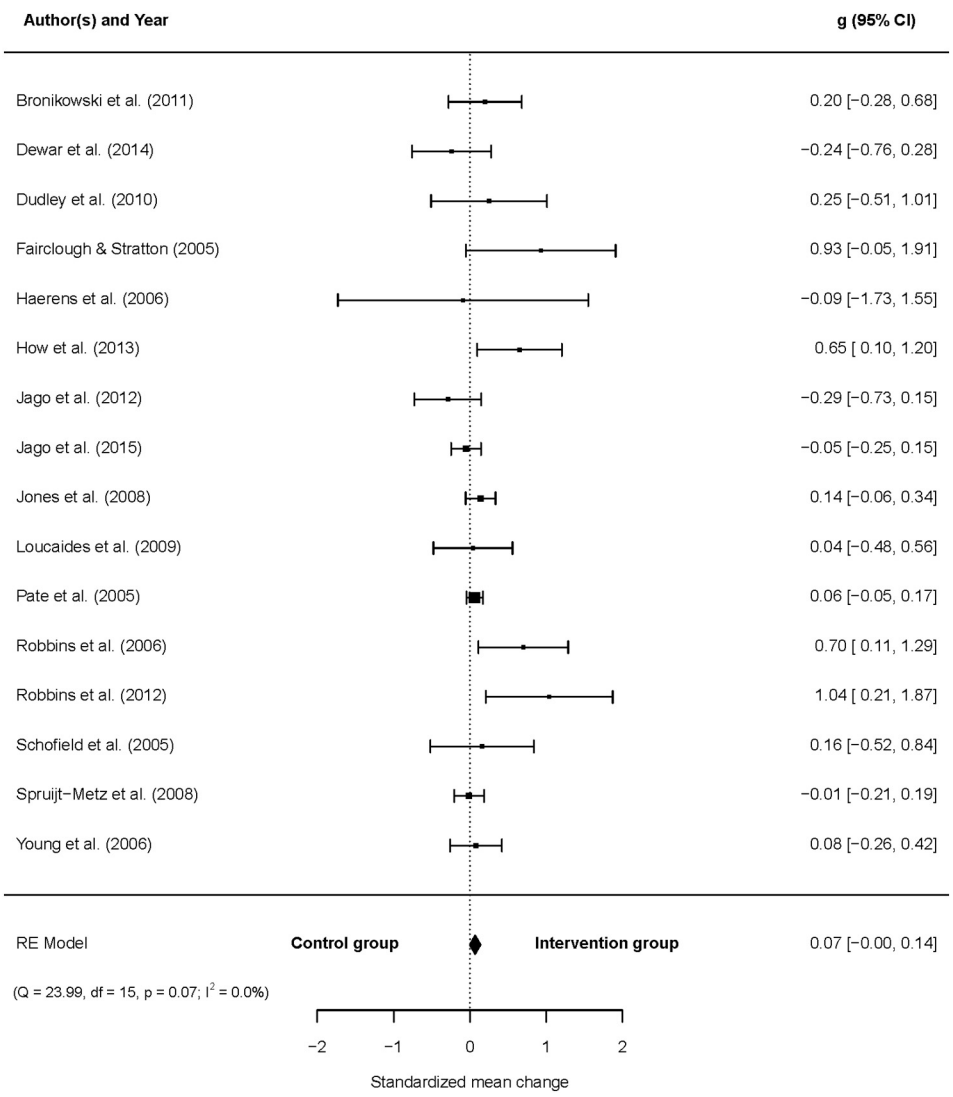

Fig. 2. Forest plot with outlier removed (k = 16). Graph depicts effect size and 95% CI for individual studies and the pooled estimate.

opportunities, short-term intervention periods, PA measurement methods, and small sample sizes which precluded the detection of significance.

Although subgroup analysis inferred a significant effect for interventions underpinned by behaviour change theory, this was a very small effect. This is consistent with findings from a recent review investigating the effectiveness of after-school PA interventions to increase MVPA (Mears and Jago, 2016). It was reported that a lack of convincing evidence exists that interventions underpinned by theory were more effective than those with no specified theory (Mears and Jago, 2016). The lack of a clear link between reported theoretical design and effectiveness could also be due to the implementation of the theories within the interventions. Few studies reported theoretical fidelity, which precludes direct inferences being made between intervention effectiveness and underpinning theory. To address this, future studies need to illustrate the direct links from theory to implementation as poor implementation of the theory could be contributing to the lack of success in some interventions (Naylor et al., 2015). The recently proposed Theory of Expanded, Extended, and Enhanced Opportunities (TEO) could provide a more practical and PA-specific theory to implement in school-based PA interventions, which is not clearly present in any of the reviewed interventions, and warrants further exploration (Beets et al., 2016). This theory can be used in conjunction with other more traditional behaviour change theories but helps provide a more PA-specific framework, to increase PA opportunities within the school setting.

Multi-component interventions were also found to have small significant effects. School-based multi-component interventions are well supported as effective approaches to impact adolescent PA levels (Kriemler et al., 2011; Pearson et al., 2015; Van Sluijs et al., 2007). Multicomponent intervention designs are consistent with the concept of Comprehensive School PA Programmes (CSPAPs), which are recommended as effective strategies to increase young people's PA (Centers for Disease Control and Prevention, 2013; World Health Organisation, 2010). CSPAPs are multicomponent in nature, aiming to intervene through PE, before and after school PA, during school PA, staff involvement, and family and community engagement. Using the CSPAP model as a form of comprehensive multicomponent intervention to target adolescent girls, integrated with an appropriate research design, may be a promising approach for future intervention efforts (Carson et al., 2014; McMullen et al., 2015).

Modified PE lessons were commonly used as single component interventions or as part of multicomponent interventions, and were effective in significantly increasing lesson time PA (Bronikowski and Bronikowska, 2011; Fairclough and Stratton, 2005; How et al., 2013; Martin and Fairclough, 2008). This supports previous research which has shown the impact of modified PE lessons designed to increase MVPA, with students engaging in 24% more MVPA during modified PE compared with students in usual PE practice conditions (Lonsdale et al., 2013). Similarly, Camacho-Minano et al. (2011) suggested that school-based interventions are more effective when enjoyment of PE is

prioritised and girls are given freedom of choice of activities. Enjoyment has been found to partially mediate the positive effect of modified PE interventions (Dishman et al., 2005), which further emphasises the importance of choice and enjoyment within school-based interventions for adolescent girls. This reinforces the importance of autonomy-supportive teaching principles such as, the Supportive, Active, Autonomous, Fair, Enjoyable (SAAFE) framework (Lubans et al., 2017). This evidence based framework encourages teachers to provide students with opportunities for autonomy during PA sessions to support the promotion of more activity during sessions (Lubans et al., 2017). However, PE occurs infrequently within schools (usually 1–2 h per week) and accounts for only a very small percentage of weekly waking hours, therefore its impact on total daily MVPA is limited.

The current review reveals a shift in the last seven years in school-based PA interventions for adolescent girls towards objective measurements of PA rather than subjective measures. Specifically, accelerometers were the preferred method of measurement, in 7 out of 8 studies conducted since 2010. The use of accelerometer-based measures allows for a more accurate assessment of PA intensity (Butte et al., 2012; Cain et al., 2013; De Vries et al., 2009). However, accelerometers provide no contextual information such as, who girls are doing activity with and what activity they are doing, which is valuable in social and fluid environments like schools. Moreover, issues such as waterproofing and wear site preclude adequate assessments of some movement modes such as, swimming or cycling (Dollman et al., 2009). Additionally, accelerometers have been found to have poor wear compliance in PA studies with adolescents (Borde et al., 2017). Few included studies utilised focus groups or interviews with participants post-intervention. Understanding the context for PA through these measurement methods may help researchers and practitioners to truly assess the effectiveness of interventions and refine and amend interventions.

Risk of bias scores did not appear to be associated with intervention effectiveness. Studies that scored poorly (* or **) for risk of bias showed a small significant effect in subgroup analyses. Risk of bias scores were low across the included interventions mainly due to the need for a greater explanation of the randomisation process which is consistent with a previous systematic review of adolescent girls (Camacho-Minano et al., 2011). Thus, poor scores may have been due to poor reporting rather than poor methodological design. Without a detailed explanation of the randomisation process, it could not be confirmed that the groups were truly distributed randomly (Higgins and Green, 2011). As found in previous reviews, both for PA interventions for adolescents (Camacho-Minano et al., 2011) and school-based behavioural interventions (Khambalia et al., 2012), allocation concealment and blinding were usually absent, and this negatively affected the risk of bias scores for the majority of included studies. The majority of studies showed low withdrawal and dropout rates (< 20%) which is positive considering the range of participant numbers and measurement methods reported. This could be due to the structure a school environment provides and the influence schools have on girls of this age (Kohl and Cook, 2013).

4.1. Strengths and limitations

To our knowledge, this is the first systematic review to combine girls-only and mixed-sex school-based PA interventions (2005 onwards) to assess their effectiveness for adolescent girls. Twenty one studies were excluded from the final synthesis because the authors did not respond to requests to provide PA data by gender within the 7-day timescale allowed. This limited response time is a limitation as this data potentially could have doubled the number of included studies, and interaction by sex tests were not explored for these studies. The inclusion of all study types, including feasibility and pilot studies, may have impacted the overall findings of the review as these tended to be smaller scale projects with small sample sizes. Where multiple primary PA outcomes were reported we used MVPA or MPA wherever possible to maintain relevance to PA guidelines. However, there were instances were alternative PA outcomes were also included (e.g., steps, accelerometer counts).

4.2. Conclusion

The meta-analysis indicated a small but significant positive effect of school-based interventions on adolescent girls' PA. Sub-group analyses indicated small but significant effects for multicomponent interventions and interventions underpinned by theory. The recent trend towards the objective measurement of PA within the school setting with accelerometry data should continue. It is important that future research and policy makers continue to recognise the school environment as a vehicle for changing girls' PA levels with an emphasis on multicomponent interventions underpinned by theory.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

None.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgments

With thanks to Professor David Lubans for manuscript advice and feedback.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ypmed.2017.09.018.

References

Ar-yuwat, S., Clark, M.J., Hunter, A., James, K.S., 2013. Determinants of physical activity in primary school students using the health belief model. J. Multidiscip. Healthc. 6, 119–126. http://dx.doi.org/10.2147/JMDH.S40876.

Barr-Anderson, D.J., Neumark-Sztainer, D., Lytle, L., Schmitz, K.H., Ward, D.S., Conway, T.L., ... Pate, R.R., 2008. But I like PE: factors associated with enjoyment of physical education class in middle school girls. Res. Q. Exerc. Sport 79 (1), 18–27. http://dx.doi.org/10.1080/02701367.2008.10599456.

Beets, M.W., Okely, A., Weaver, R.G., Webster, C., Lubans, D., Brusseau, T., ... Cliff, D.P., 2016. The theory of expanded, extended, and enhanced opportunities for youth physical activity promotion. Int. J. Behav. Nutr. Phys. Act. 13 (1), 120.

Biddle, S.J.H., Asare, M., 2011. Physical activity and mental health in children and adolescents: a review of reviews. Br. J. Sports Med. 45, 886–895. http://dx.doi.org/10.1136/bjsports-2011-090185.

Borde, R., Smith, J., Sutherland, R., Nathan, N., Lubans, D., 2017. Methodological considerations and impact of school-based interventions on objectively measured physical activity in adolescents: a systematic review and meta-analysis. Obes. Rev. 18 (4), 476–490.

Borenstein, M., Hedges, L.V., Higgins, J., Rothstein, H.R., 2010. A basic introduction to fixed-effect and random-effects models for meta-analysis. Research Synthesis Methods 1 (2), 97–111.

Bronikowski, M., Bronikowska, M., 2011. Will they stay fit and healthy? A three-year follow-up evaluation of a physical activity and health intervention in Polish youth. Scand. J. Public Health 39 (7), 704–713. http://dx.doi.org/10.1177/1403494811421059.

Butte, N.F., Ekelund, U., Westerterp, K.R., 2012. Assessing physical activity using wearable monitors: measures of physical activity. Med. Sci. Sports Exerc. 44 (1 Suppl 1), S5–12.

Cain, K.L., Sallis, J.F., Conway, T.L., Van Dyck, D., Calhoon, L., 2013. Using accelerometers in youth physical activity studies: a review of methods. J. Phys. Act. Health 10 (3), 437–450.

Camacho-Minano, M.J., LaVoi, N.M., Barr-Anderson, D.J., 2011. Interventions to promote physical activity among young and adolescent girls: a systematic review. Health

Educ. Res. 26 (6), 1025–1049. http://dx.doi.org/10.1093/her/cyr040.
Carson, R.L., Castelli, D.M., Beighle, A., Erwin, H., 2014. School-based physical activity promotion: a conceptual framework for research and practice. Childhood Obesity 10 (2), 100–106.
Centers for Disease Control and Prevention, 2013. Comprehensive School Physical Activity Programs: A Guide for Schools. (Atlanta, GA).
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Earlbaum Associates, Hilsdale, NJ, pp. 2.
Cooper, A.R., Goodman, A., Page, A.S., Sherar, L.B., Esliger, D.W., van Sluijs, E.M.F., ... Ekelund, U., 2015. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). Int. J. Behav. Nutr. Phys. Act. 12, 113.
De Vries, S.I., Van Hirtum, H.W., Bakker, I., Hopman-Rock, M., Hirasing, R.A., Van Mechelen, W., 2009. Validity and reproducibility of motion sensors in youth: a systematic update. Med. Sci. Sports Exerc. 41 (4), 818–827.
Dewar, D.L., Morgan, P.J., Plotnikoff, R.C., Okely, A.D., Batterham, M., Lubans, D.R., 2014. Exploring changes in physical activity, sedentary behaviors and hypothesized mediators in the NEAT girls group randomized controlled trial. J. Sci. Med. Sport 17 (1), 39–46. http://dx.doi.org/10.1016/j.jsams.2013.02.003.
Dishman, R.K., Motl, R.W., Saunders, R., Felton, G., Ward, D.S., Dowda, M., Pate, R.R., 2005. Enjoyment mediates effects of a school-based physical-activity intervention. Med. Sci. Sports Exerc. 37 (3), 478–487. http://dx.doi.org/10.1249/01.mss.0000155391.62733.a7.
Dollman, J., Okely, A.D., Hardy, L., Timperio, A., Salmon, J., Hills, A.P., 2009. A hitchhiker's guide to assessing young people's physical activity: deciding what method to use. J. Sci. Med. Sport 12 (5), 518–525. http://dx.doi.org/10.1016/j.jsams.2008.09.007.
Dudley, D.A., Okely, A.D., Pearson, P., Peat, J., 2010. Engaging adolescent girls from linguistically diverse and low income backgrounds in school sport: a pilot randomised controlled trial. J. Sci. Med. Sport 13 (2), 217–224. http://dx.doi.org/10.1016/j.jsams.2009.04.008.
Dumith, S.C., Gigante, D.P., Domingues, M.R., Kohl, H.W., 2011. Physical activity change during adolescence: a systematic review and a pooled analysis. Int. J. Epidemiol. 40 (3), 685–698. http://dx.doi.org/10.1093/ije/dyq272.
Duval, S., Tweedie, R., 2000a. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. J. Am. Stat. Assoc. 95 (449), 89–98.
Duval, S., Tweedie, R., 2000b. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56 (2), 455–463.
Egger, M., Smith, G.D., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315 (7109), 629–634.
Fairclough, S., Stratton, G., 2005. Improving health-enhancing physical activity in girls' physical education. Health Educ. Res. 20 (4), 448–457. http://dx.doi.org/10.1093/her/cyg137.
Haerens, L., Deforche, B., Maes, L., Cardon, G., Stevens, V., De Bourdeaudhuij, I., 2006. Evaluation of a 2-year physical activity and healthy eating intervention in middle school children. Health Educ. Res. 21 (6), 911–921. http://dx.doi.org/10.1093/her/cyl115.
Hallal, P.C., Andersen, L.B., Bull, F.C., Guthold, R., Haskell, W., Ekelund, U., 2012. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 380 (9838), 247–257. http://dx.doi.org/10.1016/s0140-6736(12)60646-1.
Heitzler, C., Lytle, L., Erickson, D., Sirard, J., Barr-Anderson, D., Story, M., 2011. Physical activity and sedentary activity patterns among children and adolescents: a latent class analysis approach. J. Phys. Act. Health 8 (4), 457–467.
Higgins, J.P., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions. Vol. 4 John Wiley & Sons.
Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. BMJ [Br. Med. J.] 327 (7414), 557.
How, Y.M., Whipp, P., Dimmock, J., Jackson, B., 2013. The effects of choice on autonomous motivation, perceived autonomy support, and physical activity levels in high school physical education. J. Teach. Phys. Educ. 32 (2), 131–148. http://dx.doi.org/10.1123/jtpe.32.2.131.
Huberty, J.L., Dinkel, D.M., Beets, M.W., 2014. Evaluation of GoGirlGo!; a practitioner based program to improve physical activity. BMC Public Health 14 (1). http://dx.doi.org/10.1186/1471-2458-14-118.
Jago, R., Sebire, S.J., Cooper, A.R., Haase, A.M., Powell, J., Davis, L., ... Montgomery, A.A., 2012. Bristol girls dance project feasibility trial: outcome and process evaluation results. Int. J. Behav. Nutr. Phys. Act. 9 (1), 83. http://dx.doi.org/10.1186/1479-5868-9-83.
Jago, R., Edwards, M.J., Sebire, S.J., Tomkinson, K., Bird, E.L., Banfield, K., ... Blair, P.S., 2015. Effect and cost of an after-school dance programme on the physical activity of 11–12 year old girls: the Bristol girls dance project, a school-based cluster randomised controlled trial. Int. J. Behav. Nutr. Phys. Act. 12 (1). http://dx.doi.org/10.1186/s12966-015-0289-y.
Janssen, I., Leblanc, A.G., 2010. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int. J. Behav. Nutr. Phys. Act. 7, 40–55. http://dx.doi.org/10.1186/1479-5868-7-40.
Jones, D., Hoelscher, D.M., Kelder, S.H., Hergenroeder, A., Sharma, S.V., 2008. Increasing physical activity and decreasing sedentary activity in adolescent girls – the incorporating more physical activity and calcium in teens (IMPACT) study. Int. J. Behav. Nutr. Phys. Act. 5 (1), 42. http://dx.doi.org/10.1186/1479-5868-5-42.
Khambalia, A.Z., Dickinson, S., Hardy, L.L., Gill, T., Baur, L.A., 2012. A synthesis of existing systematic reviews and meta-analyses of school-based behavioural interventions for controlling and preventing obesity. Obes. Rev. 13 (3), 214–233. http://dx.doi.org/10.1111/j.1467-789x.2011.00947.x.
Khunti, K., Stone, M.A., Bankart, J., Sinfield, P.K., Talbot, D., Farooqi, A., Davies, M.J., 2007. Physical activity and sedentary behaviours of South Asian and white European children in inner city secondary schools in the UK. Fam. Pract. 24 (3), 237–244. http://dx.doi.org/10.1093/fampra/cmm013.
Kohl, H.W.I., Cook, H.D., 2013. Educating the Student Body: Taking Physical Activity and Physical Education to School. National Academies Press, USA: Washington, DC.
Kriemler, S., Meyer, U., Martin, E., van Sluijs, E.M., Andersen, L.B., Martin, B.W., 2011. Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update. Br. J. Sports Med. 45 (11), 923–930.
Lonsdale, C., Rosenkranz, R.R., Peralta, L.R., Bennie, A., Fahey, P., Lubans, D.R., 2013. A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical education lessons. Prev. Med. 56 (2), 152–161. http://dx.doi.org/10.1016/j.ypmed.2012.12.004.
Loucaides, C.A., Jago, R., Charalambous, I., 2009. Promoting physical activity during school break times: piloting a simple, low cost intervention. Prev. Med. 48 (4), 332–334. http://dx.doi.org/10.1016/j.ypmed.2009.02.005.
Lubans, D.R., Lonsdale, C., Cohen, K., Eather, N., Beauchamp, M.R., Morgan, P.J., ... Smith, J.J., 2017. Framework for the design and delivery of organized physical activity sessions for children and adolescents: rationale and description of the 'SAAFE' teaching principles. Int. J. Behav. Nutr. Phys. Act. 14 (1), 24. http://dx.doi.org/10.1186/s12966-017-0479-x.
Martin, S., Fairclough, S., 2008. Dance as a health enhancing physical activity: an extra-curricular teaching intervention to increase physical activity in girls dance. Physical Education Matters 3 (1), 37–41.
McMahon, E.M., Corcoran, P., O'Regan, G., Keeley, H., Cannon, M., Carli, V., ... Wasserman, D., 2017. Physical activity in European adolescents and associations with anxiety, depression and well-being. Eur. Child Adolesc. Psychiatry 26 (1), 111–122. http://dx.doi.org/10.1007/s00787-016-0875-9.
McMullen, J., Ní Chróinín, D., Tammelin, T., Pogorzelska, M., van der Mars, H., 2015. International approaches to whole-of-school physical activity promotion. Quest 67 (4), 384–399.
Mears, R., Jago, R., 2016. Effectiveness of after-school interventions at increasing moderate-to-vigorous physical activity levels in 5- to 18-year olds: a systematic review and meta-analysis. Br. J. Sports Med. 50 (21), 1315–1324. http://dx.doi.org/10.1136/bjsports-2015-094976.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. 151 (4), 264. http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00135.
Morton, K.L., Atkin, A.J., Corder, K., Suhrcke, M., van Sluijs, E.M.F., 2016. The school environment and adolescent physical activity and sedentary behaviour: a mixed-studies systematic review. Obes. Rev. 17 (2), 142–158. http://dx.doi.org/10.1111/obr.12352.
Naylor, P.-J., Nettlefold, L., Race, D., Hoy, C., Ashe, M.C., Higgins, J.W., McKay, H.A., 2015. Implementation of school based physical activity interventions: a systematic review. Prev. Med. 72, 95–115.
Officers, Chief Medical, 2011. Start active, stay active: a report on physical activity for health from the four home countries. Retrieved from. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216370/dh_128210.pdf.
Orwin, R.G., 1983. A fail-safe N for effect size in meta-analysis. J. Educ. Stat. 8 (2), 157–159.
Pate, R.R., Ward, D.S., Saunders, R.P., Felton, G., Dishman, R.K., Dowda, M., 2005. Promotion of physical activity among high-school girls: a randomized controlled trial. Am. J. Public Health 95 (9), 1582–1587. http://dx.doi.org/10.2105/AJPH.2004.045807.
Pearson, N., Braithwaite, R., Biddle, S.J.H., 2015. The effectiveness of interventions to increase physical activity among adolescent girls: a meta-analysis. Acad. Pediatr. 15 (1), 9–18. http://dx.doi.org/10.1016/j.acap.2014.08.009.
Pluye, P., Gagnon, M.-P., Griffiths, F., Johnson-Lafleur, J., 2009. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. Int. J. Nurs. Stud. 46 (4), 529–546. http://dx.doi.org/10.1016/j.ijnurstu.2009.01.009.
Robbins, L.B., Gretebeck, K.A., Kazanis, A.S., Pender, N.J., 2006. Girls on the move program to increase physical activity participation. Nurs. Res. 55 (3), 206–216. http://dx.doi.org/10.1097/00006199-200605000-00007.
Robbins, L.B., Pfeiffer, K.A., Maier, K.S., Lo, Y.J., Wesolek, S.M., 2012. Pilot intervention to increase physical activity among sedentary urban middle school girls: a two-group pretest-posttest quasi-experimental design. J. Sch. Nurs. 28 (4), 302–315. http://dx.doi.org/10.1177/1059840512438777.
Schofield, L., Mummery, W.K., Schofield, G., 2005. Effects of a controlled pedometer-intervention trial for low-active adolescent girls. Med. Sci. Sports Exerc. 37 (8), 1414–1420. http://dx.doi.org/10.1249/01.mss.0000174889.89600.e3.
Sisson, S.B., Broyles, S.T., Baker, B.L., Katzmarzyk, P.T., 2010. Screen time, physical activity, and overweight in U.S. Youth: national survey of children's health 2003. J. Adolesc. Health 47 (3), 309–311. http://dx.doi.org/10.1016/j.jadohealth.2010.02.016.
Spruijt-Metz, D., Nguyen-Michel, S.T., Goran, M.I., Chou, C.-P., Huang, T.T.K., 2008. Reducing sedentary behavior in minority girls via a theory-based, tailored classroom media intervention. Int. J. Pediatr. Obes. 3 (4), 240–248. http://dx.doi.org/10.1080/17477160802113415.
Sterne, J.A., Egger, M., 2001. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J. Clin. Epidemiol. 54 (10), 1046–1055.
Taymoori, P., Niknami, S., Berry, T., Lubans, D., Ghofranipour, F., Kazemnejad, A., 2008. A school-based randomized controlled trial to improve physical activity among Iranian high school girls. Int. J. Behav. Nutr. Phys. Act. 5 (1), 18. http://dx.doi.org/10.1186/1479-5868-5-18.
Van Sluijs, E.M., McMinn, A.M., Griffin, S.J., 2007. Effectiveness of interventions to

promote physical activity in children and adolescents: systematic review of controlled trials. BMJ 335 (7622), 703.
Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36 (3), 1–48.
Voskuil, V.R., Frambes, D.A., Robbins, L.B., 2017. Effect of physical activity interventions for girls on objectively measured outcomes: a systematic review of randomized controlled trials. J. Pediatr. Health Care 31 (1), 75–87. http://dx.doi.org/10.1016/j.pedhc.2016.03.003.
Webber, L.S., Catellier, D.J., Lytle, L.A., Murray, D.M., Pratt, C.A., Young, D.R., ... Pate, R.R., 2008. Promoting physical activity in middle school girls: trial of activity for adolescent girls. Am. J. Prev. Med. 34 (3), 173–184. http://dx.doi.org/10.1016/j.amepre.2007.11.018.
World Health Organisation, 2010. Global Recommendations for Physical Activity for Health. Retrieved from. http://apps.who.int/iris/bitstream/10665/44399/1/9789241599979_eng.pdf.
World Health Organisation, 2014. Global Status Report on NCDs. Retrieved from. http://apps.who.int/iris/bitstream/10665/148114/1/9789241564854_eng.pdf?ua=1.
World Health Organisation, 2015. World Health Statistics 2015. Retrieved from. http://www.who.int/gho/publications/world_health_statistics/EN_WHS2015_Part2.pdf?ua=1.
Young, D., Phillips, J.A., Yu, T., Haythornthwaite, J.A., 2006. Effects of a life skills intervention for increasing physical activity in adolescent girls. Arch. Pediatr. Adolesc. Med. 160 (12), 1255–1261. http://dx.doi.org/10.1001/archpedi.160.12.1255.
Young, D., Saksvig, B.I., Wu, T.T., Zook, K., Li, X., Champaloux, S., ... Treuth, M.S., 2014. Multilevel correlates of physical activity for early, mid, and late adolescent girls. J. Phys. Act. Health 11 (11), 950–960.