School-based interventions modestly increase physical activity and cardiorespiratory fitness but are least effective for youth who need them most: An individual participant pooled analysis of 20 controlled trials

Timothy B Hartwig, Taren Sanders, Diego Vasconcellos, Michael Noetel, Philip Parker, David R. Lubans, Susana Andrade, Manuel Ávila-García, John B Bartholomew, Sarahjane Belton, Naomi E Brooks, Anna Bugge, Iván Cavero-Redondo, Lars B Christiansen, Kristen E Cohen, Tara Coppinger, Sindre M Dyrstad, Vanessa L Errisuriz, Stuart J Fairclough, Trish Gorely, Francisco Javier Huertas-Delgado, Johann Issartel, Susi Kriemler, Silje Eikanger Kvalø, Pedro Marques-Vidal, Vicente Martínez-Vizcaíno, Niels Christian Møller, Colin N Moran, John G Morris, Mary E Nevill, Angélica Ochoa-Avilés, Mai O'Leary, Louisa R Peralta, Karin A Pfeiffer, Jardena J Puder, Andrés Redondo-Tébar, Lorraine B Robbins, Mairena Sánchez-López, Jakob Tarp, Sarah L Taylor, Pablo Tercedor, Mette Toftager, Emilio Villa-González, Niels Wedderkopp, Kathryn L Weston, Zenong Yin, Zhixiong Zhou, Chris Lonsdale, Borja del Pozo Cruz.

Corresponding author:

Timothy B Hartwig

timothy.hartwig@acu.edu.au

+612 9701 4355

Sydney, Australia

Level 2, 163-167 Albert Rd, Locked Bag 2002, Strathfield, NSW 2135

Registration Study <u>protocol</u> was pre-registered on Open Science Framework.[1]

ABSTRACT

Objectives To determine if subpopulations of students benefit equally from school-based physical activity interventions in terms of cardiorespiratory fitness and physical activity. To examine if physical activity intensity mediates improvements in cardiorespiratory fitness.

Design Pooled analysis of individual participant data from controlled trials that assessed the impact of school-based physical activity interventions on cardiorespiratory fitness and device-measured physical activity.

Participants Data for 6621 children and adolescents aged 4-18 years from 20 trials were included.

Main outcome measures Peak oxygen consumption (VO_{2Peak} mL/kg/min) and minutes of moderate and vigorous physical activity.

Results Interventions modestly improved students' cardiorespiratory fitness by 0.47 mL/kg/min (95% confidence interval 0.33 to 0.61), but the effects were not distributed equally across subpopulations. Girls and older students benefited less than boys and younger students, respectively. Students with lower levels of initial fitness, and those with higher levels of baseline physical activity benefitted more than those who were initially fitter and less active, respectively. Interventions had a modest positive effect on physical activity with approximately one additional minute per day of both moderate and vigorous physical activity. Changes in vigorous, but not moderate intensity, physical activity explained a small amount (~5%) of the intervention effect on cardiorespiratory fitness.

Conclusions Future interventions should include targeted strategies to address the needs of girls and older students. Interventions may also be improved by promoting more vigorous intensity physical activity. Interventions could mitigate declining youth cardiorespiratory fitness, increase physical activity, and promote cardiovascular health if they can be delivered equitably and their effects sustained at the population level.

SUMMARY BOX

What is already known

- School-based physical activity interventions can reduce risk of cardiovascular diseases and all-cause mortality by improving children's and adolescents' cardiorespiratory fitness. Intervention effects on children's and adolescents' physical activity are unclear.

- Previous studies have been underpowered to determine if interventions differentially benefit subpopulations of students.

- Higher intensity physical activity may be more impactful in youth than lower intensity physical activity but the role of intensity in interventions for improving cardiorespiratory fitness is unclear.

What are the new findings

- Interventions modestly improved cardiorespiratory fitness and physical activity, but benefits were not equally distributed among subpopulations of students. Even in the school-setting, equity in intervention effectiveness is not guaranteed and inequalities in subpopulations of youth persist.

- Intervention designers and policymakers should ensure future interventions include targeted strategies to address the needs of girls and older students. Future interventions may also benefit from focusing on ways to promote vigorous intensity physical activity.

INTRODUCTION

Low cardiorespiratory fitness (CRF) is an independent modifiable risk factor for cardiovascular diseases and all-cause mortality.[2-4] Greater CRF is associated with better health outcomes in people of all ages, including children and adolescents.[5] Poor CRF when a person is young is also linked with all-cause mortality in adulthood.[6-9] Physical activity levels are also associated with all-cause mortality.[10] Worryingly, youth CRF has declined over the last three decades[11] and children's physical activity declines from as young as age seven years.[12] Schools are ideally positioned for health promoting interventions that aim to increase physical activity and CRF.[13] Meta-analyses show that school-based physical activity interventions can improve children's and adolescents' CRF[14-16] Paradoxically, recent meta-analyses of school-based physical activity interventions have found minimal intervention effects on device-measured moderate to vigorous physical activity.[17,18]

In addition to equivocal findings for the effectiveness of school-based interventions to improve CRF and physical activity, it is not known if interventions benefit subpopulations equally. Youth from lower socioeconomic backgrounds, those with overweight or obesity, and girls typically have poorer CRF and are less physically active than their peers who are from higher socioeconomic backgrounds, normal weight, and boys.[19-22] Without knowledge of intervention effects on subpopulations, policymakers cannot determine if disseminating school-based programs is likely to improve health among those who need it most or if they will disproportionately benefit the most advantaged students.[23] Evidence regarding the role of physical activity of varying intensities in improving CRF in school-based physical activity interventions has also not previously been synthesised. Higher intensity activity likely confers physical and mental health benefits for youth over and above those related to moderate-intensity physical activity.[24-26] If physical activity intensity mediates intervention effects on CRF, such a finding could be instructive for future interventions and policies. Individual studies are typically underpowered to determine if interventions differentially benefit subpopulations.[27] Meta-analyses of study-level results are limited because they cannot fully assess important moderators e.g., age, gender, baseline fitness levels, body mass index (BMI), without access to individual or subpopulation level data (i.e., where investigators have not themselves conducted and reported subgroup analysis).[28] Individual studies and meta-analyses might also be underpowered to effectively examine mediating pathways to understand causal mechanisms such as the role of physical activity intensity in improving CRF in school-based physical activity interventions.[29] We aimed to harmonise and pool all available individual participant data (IPD) from controlled trials of school-based physical activity interventions that assessed CRF and device-measured physical activity. We sought to determine the effect of interventions on CRF and physical

activity for subpopulations of students and whether changes in physical activity of differing intensity (i.e., moderate versus vigorous) explain the effects of interventions on CRF.

METHOD

Study population

Eligible interventions targeted school-aged youth (4-18 years), or school staff who delivered the intervention. The interventions' major focus needed to be the promotion of moderate physical activity (MPA) or vigorous physical activity (VPA) or health-related fitness components (i.e., CRF, muscular fitness and body composition). Interventions needed to be at least 8 weeks in duration and primarily implemented in the school setting. Appendix **Table 1** and **Table 2** in the supplementary file provide intervention design details.

To be included in the pooled IPD analysis, studies needed to use a randomised or quasi-randomised controlled trial design. Studies needed baseline and follow-up or post-intervention assessments of CRF and accelerometer-derived physical activity from which MPA and VPA and daily wear time data could be derived.[30-32] Two authors independently assessed risk of bias in the studies included in the analyses and resolved disagreements by consensus.

We identified eligible trials from 5 systematic reviews on school-based interventions.[14-17,33] We then combined the search terms used by these systematic reviews to create a new search strategy, and repeated searches to identify studies published subsequent to these reviews and up until April 2019 when our search was completed. We also sent emails or electronic posts to members of three relevant organisations and searched three trial registries for unpublished trials. The full details of the search strategy are available in the supplementary file.

Two authors independently screened titles and abstracts and read full-text versions of potentially relevant articles. We invited lead investigators of eligible studies to contribute raw IPD for the purposes of data harmonising for pooled analyses. Contributing investigators signed a data transfer agreement. The lead university's Research Ethics Committee granted permission for each study's data to be obtained and investigators transferred their de-identified data via a secure file-sharing system.

Outcomes

School-based interventions used a variety of field-based tests of physical performance to measure CRF.[34] Individual studies reported researchers collected these data using standardised testing procedures. To harmonise these data, we converted all students' scores to an estimate of peak oxygen consumption (VO_{2Peak}) using established equations.[35-38] Appendix **Table 3** in the supplementary file provides VO_{2Peak} estimation details.

We determined physical activity from accelerometer data. There were no exclusion criteria for accelerometer device type, wear location, or epoch length. We harmonised these data by reprocessing raw accelerometer files with a standardised protocol, with wear time criteria of 600 minutes per day on at least 3 weekdays and 1 weekend day.[39-41] Hip-worn accelerometers were reprocessed to the lowest common resolution using a 60s epoch (i.e., the period of time over which accelerations are aggregated) with Evenson cut-points.[39] Wrist-worn accelerometers were reprocessed using 10s epochs with Chandler cut-points applied to the vector magnitude,[40] or reprocessed in R to calculate the average magnitude of dynamic acceleration using the Euclidean norm minus one method.[41]

Socio-economic status (SES) was measured using a variety of instruments and included student-reported parental income, parental education, parental occupation, family wealth, access to literature, or indices based on zip/postcode. Some authors reported SES dichotomously while others used an ordinal scale. To harmonise these data to allow for SES to be assessed across studies with differences in scales, we converted all measures to a dichotomous outcome of 'low' or 'high' SES. Appendix

Table 4 in the supplementary file provides these details. All included studies were able to provide individual participant height and weight measured by researchers and we used these to calculate BMI and BMI z-scores.

Statistical analysis

From the initial pooled dataset, we excluded participants who did not provide accelerometer and CRF data at baseline. In all analyses, we used full information maximum likelihood estimation to account for missing data. This method ensured that we used all available information from participants who met the inclusion criteria.[42] For all analyses, we examined 95% confidence intervals and set our P-value at 0.05. We used R and Mplus for data cleaning and analysis.[43,44]

To examine intervention effects on students' physical activity and CRF we used structural equation modelling. As suggested by Curran et al.,[45] we used study fixed-effects to account for the clustering of participants in studies. Dependent variables for CRF and physical activity are the outcomes at follow-up adjusted for baseline. Intervention effects were estimated by the total effect in a mediation model that included VPA and MPA as mediators. We examined both main effects and student-level subpopulation effects, including the following pre-specified student-level moderators: age (chronological), BMI, and baseline CRF as continuous variables and gender as categorical. For moderator analyses, continuous variables were stratified into bottom, middle, and upper tertiles. At the suggestion of investigators who contributed datasets, we also explored possible moderating effects of baseline MPA, VPA, and SES. We then tested pre-specified study-level moderators,

including type of CRF test and accelerometer wear location. A post-hoc study-level moderation analysis for length of intervention was also included. We used fixed-effects models for student-level moderators and random-effects models for study-level moderators.

To test the possible mediating effect of intervention-related changes in physical activity intensity on CRF improvements, we conducted mediation analysis in a structural equation framework using study fixed-effects models and the lavaan package.[46,47] We included multiple mediators (i.e., follow-up MPA and VPA adjusted for baseline separately) in all mediation models.

For physical activity data we used the proportion of accelerometer wear time spent in MPA and VPA, rather than minutes of physical activity (which is influenced by daily wear time). To aid the interpretation of these data, we reported changes in proportion of MPA and VPA in minutes. We included accelerometer-derived variables even if the raw data could not be harmonised into 60s epochs and differed by wear location.

Studies included in our IPD pooled analysis differed in accelerometer wear location (wrist and hip). To harmonise these data we used cut-points derived from proprietary count measures and raw accelerations depending on what raw data was provided. To test the potential impact of harmonising physical activity data from different accelerometer wear locations we conducted additional analyses. We used a between-study level moderator analysis and a multigroup analysis to test whether treatment effects were different across studies using wrist or hip wear locations. For the multigroup analysis, we first ran our main treatment effects models separately for each group (i.e., for studies using hip and wrist location, respectively) in which estimates for MPA and VPA were freely estimated. Second, we ran the same models but with the estimates constrained to be equal. We then compared the model fit using Anova.

Deviations from protocol

Our pre-registered protocol[1] stated we would use regression models with cluster robust standard errors for study clustering. This was driven by a belief that study intraclass correlations (ICCs) would be small. However, ICCs from a random-effects model for treatment effects were substantial for some outcomes: 0.06 for CRF, 0.26 for VPA, and 0.49 for MPA. A cluster robust model would provide poor control for latent study-to-study variation. We chose not to use a random-effects approach to account for study clustering because participants in studies were not random draws from a single homogenous population. Hence, we used a study-fixed effects approach[45] except when testing study-level moderators, where we used random-effects models because fixed-effects are not appropriate for study-level data.[45]

Patient and public involvement

Patients and members of the public were not involved in the design, analysis or interpretation of this pooled IPD study. The results of this research are of broad public health interest, particularly for students and parents or anyone involved with school aged children and adolescents. The results of this study will be disseminated through institutional websites, press releases, and tailored messages to schools, educational organisations and governing bodies.

RESULTS

Study and participant characteristics

The flow of studies and participants is available in the hybrid PRISMA and CONSORT diagram (**Figure 1**). Of the 33 eligible studies, investigators provided data from 24 studies,[48-70] with the remaining 9 investigators unable or unwilling to provide data. Appendix **Table 5** in the supplementary file provides these details. The initial sample included 13,155 students, of which we excluded 6,534 due to missing baseline physical activity or CRF data, or insufficient accelerometer wear time. Four studies had no remaining participants after these exclusions, meaning we included 20 of the 24 studies with a total of 6,621 students included. Excluded students did not appear to be systematically different from those included. Appendix **Table 6** in the supplementary file provides these details. **Table 1** summarises participant characteristics. Study sample sizes after exclusion ranged from N = 16 to 1,297. Interventions' lengths varied from 2 to 28 months (mean (SD) = 8.3 (5.2) months).

Table 1. Included Sample Characteristics

Variable	Total Sample			Control Group			Treatment Group		
	Mean	SD	Missing %	Mean	SD	Missing %	Mean	SD	Missing %
Percentage Girls	60.90		0.00	61.90		0.00	60.00		0.00
Age (in years)	10.06	2.29	0.27	10.20	2.26	0.36	9.94	2.31	0.20
BMI	18.78	4.36	0.33	18.96	4.46	0.42	18.62	4.26	0.26
CRF T1	46.84	8.96	0.00	46.70	9.18	0.00	46.97	8.77	0.00
CRF T2	47.94	9.38	11.31	47.41	9.59	10.67	48.41	9.16	11.88
MPA Min T1	45.06	30.08	0.00	42.94	29.07	0.00	46.93	30.82	0.00
MPA Min T2	47.09	32.18	37.61	42.93	29.31	38.99	50.61	34.03	36.39
MPA Proportion T1	4.54	2.52	0.00	4.40	2.48	0.00	4.67	2.54	0.00
MPA Proportion T2	4.64	2.64	37.61	4.32	2.50	38.99	4.91	2.73	36.39

VPA Min T1	14.29	20.08	0.00		14.18	19.24	0.00		14.38	20.80	0.00
VPA Min T2	13.81	17.99	37.61		12.67	16.08	38.99		14.78	19.40	36.39
VPA Proportion T1	1.51	1.96	0.00		1.51	1.87	0.00		1.51	2.04	0.00
VPA Proportion T2	1.45	1.75	37.61		1.34	1.54	38.99		1.54	1.91	36.39
Wear Time Min T1	964.28	251.03	0.00		952.84	251.01	0.00		974.36	250.66	0.00
Wear Time Min T2	985.73	253.93	37.61		971.08	253.83	38.99		998.11	253.41	36.39

T1 = baseline, T2 = follow-up, CRF = Cardiorespiratory fitness (VO$_{2Peak}$ mL/kg/min), Wear time = average daily wear time, VPA = Vigorous physical activity, MPA = Moderate physical activity, Min = Mins per day, Proportion = proportion of daily wear time. Missing data on CRF, VPA, and MPA is zero due to inclusion criteria.

Risk of bias assessment

Appendix **Table 7** in the supplementary file provides risk of bias assessment. Of the 20 studies included in the main analysis, 9 studies rated as low risk for sequence generation, 10 studies for allocation concealment, 2 studies for blinding of participants and assessors for all outcomes, 13 studies for incomplete outcome, and 8 studies for other biases. All studies rated low risk for selective reporting and no studies blinded all personnel for all outcomes.

Intervention effectiveness

The overall effect of interventions on CRF was 0.47 mL/kg/min (95% confidence interval 0.33 to 0.61) (**Table 2**). There were also small but significant effects on physical activity. Compared with controls, interventions led to a 0.15 (95% confidence interval 0.08 to 0.22) percentage point increase in time spent in VPA, and a 0.13 (95% confidence interval 0.03 to 0.24) percentage point increase in time spent in MPA. Using mean baseline wear time (964.28 min), these translate to 1.45 and 1.25 min/day of additional VPA and MPA, respectively.

Effect of accelerometer wear location

A between-study level moderator analysis showed that accelerometer wear location did not moderate treatment effects on VPA (estimated difference = 0.209 min, S.E. = 0.49 [$P = 0.673$]) or MPA (Estimated difference= 0.551 min, S.E. = 0.803 [$P = 0.492$]). Multi-group analyses showed there were no differences in physical activity intervention effects by wear location, which supports the observation at between-study level. Appendix **Table 8** and **Table 9** in the supplementary file provides accelerometer wear location multigroup sensitivity analysis results.

Table 2. Mediation Model for CRF

Estimate	Effect	95% CI Lower	95% CI Upper	P
Direct Effect	0.44	0.30	0.58	0.00
Indirect via VPA	0.02	0.01	0.04	0.01
Indirect via MPA	0.00	0.00	0.01	0.26
Total Effect	0.47	0.33	0.61	0.00
Difference between Indirect Effects	0.02	0.00	0.04	0.06
Proportion of Treatment Effect mediated by VPA	0.05	0.01	0.10	0.01
Proportion of Treatment Effect mediated by MPA	0.01	-0.01	0.03	0.27

CRF = Cardiorespiratory fitness (VO_{2Peak} mL/kg/min), VPA = Vigorous physical activity (per day percentage units), MPA = Moderate physical activity (per day percentage units), CI = confidence interval.

Student-level moderators of intervention effectiveness

Effects on cardiorespiratory fitness

Several factors moderated intervention effects on CRF (**Figure 2**). Younger students (bottom tertile: mean age = 7.77) received approximately twice the intervention effect of older students (upper tertile: mean age = 12.35). Gender was also a significant moderator; boys received a much larger effect than girls. Students' SES was not associated with intervention effectiveness for CRF. We also examined baseline fitness, physical activity, and BMI as moderators. Baseline CRF was negatively associated with intervention effectiveness. A student with low baseline CRF (bottom tertile: mean baseline VO_{2Peak} = 37.98 mL/kg/min) typically received greater benefit than a student with high baseline CRF (upper tertile: mean baseline VO_{2Peak} = 55.90 mL/kg/min. Baseline physical activity levels were positively associated with intervention effectiveness. Students with low levels of baseline VPA (bottom tertile: mean = 0.49% of wear time or 4.7 min) had a smaller increase in CRF compared to those with high baseline VPA (upper tertile: mean = 1.78% of wear time or 17.2 min). Finally, BMI-z score did not moderate the effect of the interventions on CRF.

Effects on physical activity

Compared with results related to CRF, there were fewer significant moderators for the effect of interventions on students' physical activity intensity. Students' age increased the intervention effect on both MPA and VPA. Older students showed a larger intervention effect on VPA and MPA while younger students' effect sizes were not significantly different from zero. Gender also moderated the

intervention effect on VPA, but not MPA. Students' SES was not associated with intervention effectiveness for MPA or VPA.

Baseline CRF was not associated with intervention effects on students' MPA or VPA (**Figure 2**). Students with high levels of baseline VPA showed a larger intervention effect on their physical activity (VPA and MPA), compared to those with low levels of baseline VPA (both non-significant). Results were similar for baseline MPA where high active students at baseline showed beneficial effects (VPA and MPA), while intervention effects for students with low baseline MPA were not different from zero.

Study-level moderators of intervention effectiveness

Regarding study-level moderators, CRF test type varied between studies and we tested whether CRF test type moderated treatment effects. CRF test type did not moderate the effect of interventions on CRF (CRF: -0.36 [$P = 0.54$]) or physical activity (VPA: 0.40 [$P = 0.07$]; MPA: 0.67 [$P = 0.48$]). A post-hoc moderation analysis also showed no moderation effects for intervention duration (time to follow-up) for CRF (-0.014 S.E. 0.018 [$P = 0.446$]) or VPA (-0.016 S.E. 0.011 [$P = 0.149$]) but intervention length did appear to influence intervention effects on MPA (-0.072 S.E. 0.034 [$P = 0.033$]) with shorter interventions having a greater impact on MPA.

Mediators of intervention effects

We examined mediation pathways through MPA and VPA (**Figure 3**). Vigorous, but not moderate, intensity physical activity significantly mediated the intervention effect on CRF; however, these mediation effects were small. Changes in VPA explained 0.02 mL/kg/min (or 5%) of the 0.47 mL/kg/min total effect (**Table 2**).

DISCUSSION

Principal findings

Our pooled IPD analysis found that school-based physical activity interventions produced small improvements in students' CRF by 0.47 mL/kg/min and modestly increased physical activity with approximately one additional minute per day of both MPA and VPA. Only VPA mediated the effects of interventions on CRF but these mediation effects were small. We also found that not all students benefited equally. Boys accumulated more VPA and had approximately double the increase in CRF compared with girls. Barriers and facilitators of physical activity are different for boys and girls[71], it is possible that interventions were better suited to boys than girls or failed to meet the specific needs of girls.[72] Adiposity and lean mass also differentially affect fitness test performance and estimates of CRF expressed as a ratio scaled to body weight (i.e., VO_{2Peak} mL/kg/min), and gender

differences in growth and pubertal status may have exacerbated this effect. Despite older students accumulating more VPA from interventions, younger students' CRF benefited more than older students. Physical activity in children declines with age[12] and, while the ability of interventions to increase VPA in older students is encouraging, it is worrying that interventions were less effective at improving CRF in older students. As expected[73], low baseline CRF was associated with greater intervention effects on CRF. It is reassuring that students with the lowest CRF were most likely to benefit from interventions. Paradoxically, high baseline VPA was associated with greater intervention effects on CRF and VPA and MPA. If we assume that students with higher baseline VPA are distinct, for example in motivation, school-based interventions may provide opportunities for behaviour change in this subpopulation of students, whereas students with lower motivation may not take advantage of these same opportunities.[74,75] It is also plausible that students with high levels of VPA at baseline lived in an environment conducive for physical activity, and benefited more from interventions than students with limited support and resources in their schools and communities.[76] In sum, it seems that even in the school-setting, equity in intervention effectiveness is not guaranteed and inequalities in subpopulations of youth persist despite the efforts of interventions to date.[77]

Previous studies support that school-based physical activity interventions can, at least on the whole, improve children's and adolescents' CRF[14-16] but have shown minimal intervention effects on device-measured physical activity.[17,18] Trials included in our study differ from previous meta-analyses of physical activity in school-based interventions by our additional requirement of data on CRF which could partly explain why our IPD analysis showed a moderate increase in physical activity. Modulating physical activity and understanding determinants has proven difficult[78] and doubts remain regarding the ability of school-based interventions to meaningfully increase youth physical activity. However, even modest increases in physical activity, especially vigorous activity, could have important health impacts if achieved at the population level and maintained throughout childhood. For example, independent associations between vigorous intensity physical activity and cardiometabolic risk factors that included lower levels of waist circumference, BMI z-score, systolic blood pressure, and increased CRF in children and adolescents have been shown across a narrow range of VPA (approximately 7 minutes of VPA per day).[79]

In our analysis, only VPA mediated the effects of interventions on CRF. VPA, however, explained only about 5% of the change in CRF. Physical activity is the principal pathway through which CRF can be increased.[80] We expected a strong mediating effect through physical activity and it was surprising that only limited mediation through VPA was shown in a large pooled sample. It is worth noting that the accelerometer-measured physical activity included in our pooled IPD study provided an estimate of absolute intensity. It was thus not possible to explore whether intensity relative to

individuals' maximum aerobic capacity also contributed to intervention effects on CRF and influenced the mediation results. It is plausible that individuals with a low aerobic capacity, those with a higher BMI, and those who were younger could have experienced relatively high intensities of physical activity even if absolute measures appeared to be low. This might also partly explain why older students accumulated more VPA from interventions but attained a lower intervention effect than younger students for CRF. It is unclear what other factors could explain the changes in CRF but limitations in the measurement of physical activity in the studies included in our analysis might also have contributed to this finding. Specifically, although accelerometers provide an optimal balance between validity and feasibility in field-based physical activity assessment,[81] they are not accurate measures of some types of physical activity (e.g., cycling) and are typically not worn during water-based activities or contact sports (e.g., football, basketball). Since children frequently participate in organised sport (e.g., ~50% of children and adolescents globally),[82] these limitations in measurement, which appear to be greater when a 24 hour accelerometer wear protocol is not used, could obfuscate the relationship between physical activity and CRF in our study.

Study strengths and limitations

Strengths of this pooled IPD analysis include the large sample of participants with student-level confounders and covariates and the ability to harmonise device-measured physical activity and CRF measures. This meant we were uniquely positioned to explore moderating and mediating effects on CRF and physical activity within interventions. Limitations include that we delimited our sample to trials that included measures of CRF and device-measured physical activity. Excluding trials that measured CRF or physical activity but not both, reduced the number of studies eligible to provide data for analysis. We also excluded trials that only measured physical activity during school hours because interventions can affect physical activity behaviours across the whole day. We used a systematic approach to identify eligible studies and a large proportion of invited studies contributed individual data, but our study inclusion criteria might have increased the risk of selection bias. Some factors varied across interventions included in our analysis, including intervention design, doses in intensity, length of each session, and duration of the study and these could influence the effectiveness of individual interventions and thus impact the overall findings of the pooled IPD analysis. We also did not exclude studies at high risk of biases. Many school-based interventions are at high risk for bias because implementing controlled trials in a school setting poses unique challenges. For example, it is not feasible to blind school students and assessors to all outcomes. The studies we included objectively measured the primary outcomes which can limit the impact of some types of bias, but we cannot rule out that risk of bias could affect our conclusions. We also did not assess the fidelity of individual interventions included in our study. We included data from trials that

were not randomised which could increase the risk of bias and confounding. However, most included data (>70%) came from randomised controlled study designs. An additional limitation is the strict wear time inclusion criteria we applied which may have biased the sample towards more conscientious, more active, or fitter youth. This bias is unlikely to have impacted internal validity but may limit the extent to which we can generalise findings to all youth.

We harmonised physical activity data from hip and wrist wear locations and used cut-points derived from proprietary count measures and raw accelerations depending on what raw data was provided for the pooled IPD analysis. This was a pragmatic approach to harmonising accelerometer data. In some respects, excluding data from studies that used wrist wear location would have improved our data harmonisation but removing about 25% of the overall sample would have reduced the overall generalisability of our results.[83] Further, because additional analyses showed that wear location did not appear to meaningfully impact main outcomes in the present study, we chose to pool all valid data from both hip and wrist wear locations. This, however, may have affected the magnitude of the point estimates for MPA and VPA and our physical activity results should be interpreted cautiously. Regarding our analysis, we were unable to account for school or class clustering due to the diversity of designs and the relative lack of clear information on clustering across studies. It was not possible to provide a clear data structure that could be applied consistently across the studies that adequately accounted for the variety of research designs. Clustering in school-based interventions is typically low and unlikely to have a large effect on our results. For example, in one study dataset included in our pooled sample for which school-level data were available ICCs ranged from 0.03-0.07. In the mediation analysis, confounding can bias the mediation effect because only the exposure and not the mediator is being randomised.[84] This can lead to the extent of mediation being overestimated. Trials included in our pooled sample originated from 11 different countries, but these were largely middle to high-income countries, which could also affect generalisability. Further, the finding of no moderation by SES could be the result of large heterogeneity in the different measures of SES which we were also only able to harmonise to a dichotomous outcome. Available data also could not account for differences in maturational status which could be an important moderator. Finally, we relied on field-based measures of CRF and regression equations to estimate peak oxygen uptake. These equations are less accurate at the individual level and have not been formally validated in children younger than 8 years old. These tests are however, widely used and show reasonable validity compared with lab-based measures of peak oxygen uptake .[85] Field-based tests can be influenced by motivation to exercise at, or near, a maximum level of exertion and can penalise heavier students who perform more work to complete the test and change direction with greater difficulty as a result of their extra weight. If school-based interventions increased student motivation to excel in these tests, trial results may be biased toward the intervention condition.

Implications and future directions

There is no established criterion for determining a clinically meaningful change in children's cardiorespiratory fitness or physical activity levels. Interventions included in our study, on average, produced only small changes in physical activity and CRF. Indeed, the intervention effect for CRF was smaller than the standard error of estimate of the peak oxygen consumption prediction equations used to derive the VO_{2Peak}.[35-37] Effects were larger for some subpopulations of students, but the meaningfulness of these changes may need to be determined in the context of their potential impact over time. For example, if improvements in CRF and physical activity can accumulate over many years of schooling or at a population level, the impacts on cardiovascular health may be substantial. Variance in CRF is largely heritable,[86] but children's and adolescent's CRF can improve or worsen through changes in physical activity.[87] A recent pooled analysis showed that youth (9-17 year) fitness levels appear to have declined over a 33-year period with a reduction in peak oxygen uptake of 3.3 mL/kg/min (95% confidence interval -3.5 to -3.1).[11] Even though interventions produced only small improvements in CRF, our results suggest that relatively brief school-based interventions (mean duration = 8.3 months) can reduce this decline by about 15% (i.e., 0.47 mL/kg/min). Cardiovascular risk factors are already present in a high proportion of children and adolescents.[88] Low and declining CRF is thus a significant public health issue. School-based physical activity interventions can be wide-reaching and there may be a role for interventions to help mitigate declining youth CRF and increase physical activity. Small improvements in CRF and physical activity could help promote cardiovascular health but intervention designers and policymakers need to ensure future interventions are delivered equitably, and their effects sustained at the population level. It is also worth considering what future interventions must do differently to potentially increase their effects on CRF and physical activity.

Finally, pooling individual participant data offers important advantages over study-level meta-analyses but requires more effort to obtain, harmonise and analyse raw data[27, 83]. There may also be additional obstacles. In our study, we encountered different principles and practices regarding the sharing of raw deidentified data. These varied by country and data sharing laws were not always clear or accessible to researchers. This might contribute to some data custodians declining to share their raw data. Nevertheless, researchers considering undertaking IPD studies should be encouraged by the progress being made to facilitate data sharing practices.[89, 90] Future IPD studies, especially those using accelerometer data, should consider providing data processing scripts so that data custodians can reprocess their data. This could streamline data harmonization and reduce the challenges of transferring large quantities of raw unprocessed data.

Conclusions

Benefits of school-based physical activity interventions were not equally distributed among subpopulations of students. Future school-based interventions should include targeted strategies to address the needs of girls and older students. Interventions may also be improved by focusing on ways to promote more vigorous intensity physical activity.

specifically mentioned contributed raw data (data acquisition) for the pooled analysis and made important intellectual contributions by critically revising the study protocol, manuscript drafts and the final submitted manuscript. All authors agree to being accountable for all aspects of the work related to the accuracy or integrity of any part of the work. The corresponding author (TBH) attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Patient and public involvement statement See methods section in manuscript

REFERENCES

1.   Influence of physical activity intensity on cardiorespiratory fitness in school-based interventions: a pooled mediation analysis of controlled trials. https://osf.io/k59ch/?view_only=06d886e0b72c4bd283535be5ab5fea7c. Published October 17, 2019.

2.   Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation* 2016;134(24):e653-e699. doi:10.1161/CIR.0000000000000461.

3.   Bradley SM, Michos ED, Miedema MD. Physical activity, fitness, and cardiovascular health: insights from publications in JAMA Network Open. *JAMA Netw Open* 2019;2(8):e198343. doi:10.1001/jamanetworkopen.2019.8343.

4.   Mandsager K, Harb S, Cremer P, et al. Association of cardiorespiratory fitness with long-term mortality among adults undergoing exercise treadmill testing. *JAMA Netw Open* 2018;1(6):e183605. doi:10.1001/jamanetworkopen.2018.3605.

5.   Ortega FB, Ruiz JR, Castillo MJ, et al. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes* 2008;32(1):1-11. doi:10.1038/sj.ijo.0803774.

6.   Högström G, Nordström A, Nordström P. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. *Int J Epidemiol* 2016;45(4):1159-1168. doi:10.1093/ije/dyv321.

7.   Sato M, Kodama S, Sugawara A, et al. Physical fitness during adolescence and adult mortality. *Epidemiology* 2009;20(3):463-464. doi:10.1097/EDE.0b013e31819ed09f.

8.   Andersen K, Rasmussen F, Held C, et al. Exercise capacity and muscle strength and risk of vascular disease and arrhythmia in 1.1 million young Swedish men: cohort study. *BMJ* 2015;351:h4543. doi:10.1136/bmj.h4543.

9.   García-Hermoso A, Ramírez-Vélez R, García-Alonso Y, et al. Association of cardiorespiratory fitness levels during youth with health risk later in life: a systematic review and meta-analysis. *JAMA Pediatr* 2020. doi:10.1001/jamapediatrics.2020.2400.

10.   Ekelund U, Tarp J, Steene-Johannessen J, et al. Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ* 2019;366:l4570. doi.org/10.1136/bmj.l4570.

11.   Tomkinson GR, Lang JJ, Tremblay MS. Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. *Br J Sports Med* 2019;53(8):478-486. doi:10.1136/bjsports-2017-097982.

12.   Farooq MA, Parkinson KN, Adamson AJ, et al. Timing of the decline in physical activity in childhood and adolescence: Gateshead Millennium Cohort Study. *Br J Sports Med* 2018;52(15):1002-1006. doi:10.1136/bjsports-2016-096933.

13.   Physical Activity Guidelines Advisory Committee. 2018 Physical Activity Guidelines Advisory Committee Scientific Report. Washington, DC: U.S. Department of Health and Human Services; 2018. https://health.gov/sites/default/files/2019-09/PAG_Advisory_Committee_Report.pdf

14.   Pozuelo-Carrascosa DP, García-Hermoso A, Álvarez-Bueno C, et al. Effectiveness of school-based physical activity programmes on cardiorespiratory fitness in children: a meta-analysis of randomised controlled trials. *Br J Sports Med* 2018;52(19):1234-1240. doi:10.1136/bjsports-2017-097600.

15.   Minatto G, Barbosa Filho VC, Berria J, et al. School-based interventions to improve cardiorespiratory fitness in adolescents: systematic review with meta-analysis. *Sports Med* 2016;46(9):1273-1292. doi:10.1007/s40279-016-0480-6.

16.   Dobbins M, Husson H, DeCorby K, et al. 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). doi:10.1002/14651858.CD007651.pub2.

17.   Love R, Adams J, van Sluijs EMF. Are school-based physical activity interventions effective and equitable? A meta-analysis of cluster randomized controlled trials with accelerometer-assessed activity. *Obes Rev* 2019;20(6):859-870. doi:10.1111/obr.12823.

18.   Borde R, Smith JJ, Sutherland R, et al. Methodological considerations and impact of school-based interventions on objectively measured physical activity in adolescents: a systematic review and meta-analysis. *Obes Rev* 2017;18(4):476-490. doi:10.1111/obr.12517.

19.   Hallal PC, Andersen LB, Bull FC, et al. Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet* 2012;380(9838):247-257. doi:10.1016/S0140-6736(12)60646-1.

20.   Trost SG, Pate RR, Sallis JF, et al. Age and gender differences in objectively measured physical activity in youth. *Med Sci Sports Exerc* 2002;34(2):350-355. doi:10.1097/00005768-200202000-00025.

21.   Hardy LL, Mihrshahi S, Drayton BA, et al. NSW Schools Physical Activity and Nutrition Survey (SPANS) 2015: Full Report. NSW Ministry of Health; 2017. https://www.health.nsw.gov.au/heal/Publications/spans-2015-summary-report.PDF.

22.   Donkin A, Marmot M. Poverty and reductions in fitness levels in children and adolescents in upper middle-income countries. *Br J Sports Med* 2019;53(8):462-463. doi:10.1136/bjsports-2017-098992.

23.   Bakermans-Kranenburg MJ, van IJzendoorn MH, Bradley RH. Those who have, receive: The Matthew Effect in early childhood intervention in the home environment. *Rev Educ Res* 2005;75(1):1-26. doi:10.3102/00346543075001001.

24.   Costigan SA, Lubans DR, Lonsdale C, et al. Associations between physical activity intensity and well-being in adolescents. *Prev Med* 2019;125:55-61. doi:10.1016/j.ypmed.2019.05.009.

25.   Ruiz JR, Rizzo NS, Hurtig-Wennlöf A, et al. Relations of total physical activity and intensity to fitness and fatness in children: the European Youth Heart Study. *Am J Clin Nutr* 2006;84(2):299-303. doi:10.1093/ajcn/84.2.299.

26.   Tarp J, Child A, White T, et al. Physical activity intensity, bout-duration, and cardiometabolic risk markers in children and adolescents. *Int J Obes* 2018;42(9):1639-1650. doi:10.1038/s41366-018-0152-8.

27.   Ioannidis J. Next-generation systematic reviews: prospective meta-analysis, individual-level data, networks and umbrella reviews. *Br J Sports Med* 2017;51:1456-1458. doi.org/10.1136/bjsports-2017-097621.

28.   Cooper H, Patall EA. The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychol Methods* 2009;14(2):165-176. doi:10.1037/a0015565.

29.   Lee H, Herbert RD, McAuley JH. Mediation analysis. *JAMA* 2019;321(7):697-698. doi:10.1001/jama.2018.21973.

30.   Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc* 2005;37(11 Suppl):S531-S543. doi:10.1249/01.mss.0000185657.86065.98.

31.   Ward DS, Evenson KR, Vaughn A, et al. Accelerometer use in physical activity: best practices and research recommendations. *Med Sci Sports Exerc* 2005;37(11 Suppl):S582-S588. doi:10.1249/01.mss.0000185292.71933.91.

32.   Rich C, Geraci M, Griffiths L, et al. Quality control methods in accelerometer data processing: defining minimum wear time. *PLoS One* 2013;8(6):e67206. doi:10.1371/journal.pone.0067206.

33.   Metcalf B, Henley W, Wilkin T. Effectiveness of intervention on physical activity of children: systematic review and meta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). *BMJ* 2012;345:e5888. doi:10.1136/bmj.e5888.

34.   Domone S, Mann S, Sandercock G, et al. A method by which to assess the scalability of field-based fitness tests of cardiorespiratory fitness among schoolchildren. *Sports Med* 2016;46(12):1819-1831. doi:10.1007/s40279-016-0553-6.

35.   Mahar MT, Guerieri AM, Hanna MS, et al. Estimation of aerobic fitness from 20-m multistage shuttle run test performance. *Am J Prev Med* 2011;41(4 Suppl 2):S117-S123. doi:10.1016/j.amepre.2011.07.008.

36.   Aadland E, Andersen LB, Lerum Ø, et al. The Andersen aerobic fitness test: new peak oxygen consumption prediction equations in 10 and 16-year olds. *Scand J Med Sci Sports* 2018;28(3):862-872. doi:10.1111/sms.12985.

37.   Cureton KJ, Sloniger MA, O'Bannon JP, et al. A generalized equation for prediction of VO2peak from 1-mile run/walk performance. *Med Sci Sports Exerc* 1995;27(3):445-451. https://www.ncbi.nlm.nih.gov/pubmed/7752874.

38.   McArdle WD, Katch FI, Pechar GS, et al. Reliability and interrelationships between maximal oxygen intake, physical work capacity and step-test scores in college women. *Med Sci Sports* 1972;4(4):182-186. https://www.ncbi.nlm.nih.gov/pubmed/4648576.

39.   Evenson KR, Catellier DJ, Gill K, et al. Calibration of two objective measures of physical activity for children. *J Sports Sci* 2008;26(14):1557-1565. doi:10.1080/02640410802334196.

40.   Chandler JL, Brazendale K, Beets MW, et al. Classification of physical activity intensities using a wrist-worn accelerometer in 8-12-year-old children. *Pediatr Obes* 2016;11(2):120-127. doi:10.1111/ijpo.12033.

41.   Hildebrand M, Van Hees VT, Hansen BH, et al. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc* 2014;46(9):1816-1824. doi:10.1249/MSS.0000000000000289.

42.   Enders CK. Applied Missing Data Analysis. New York: Guilford Press; 2010.

43.   Muthén L, Muthén B. Mplus User's Guide. Muthén & Muthén; 2017.

44.   R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2018. https://www.R-project.org/.

45.   Curran PJ, Hussong AM. Integrative data analysis: the simultaneous analysis of multiple data sets. *Psychol Methods* 2009;14(2):81-100. doi:10.1037/a0015914.

46.   Imai K, Keele L, Tingley D, et al. Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *Am Polit Sci Rev* 2011;105(4):765-789. doi:10.1017/s0003055411000414.

47.   Imai K, Keele L, Tingley D, et al. Causal mediation analysis using R. In: Vinod HD, ed. Advances in Social Science Research Using R. Heidelberg, Germany: Springer Science+Business Media; 2010:129-154. doi:10.1007/978-1-4419-1764-5_8.

48.   Andrade S, Lachat C, Ochoa-Aviles A, et al. A school-based intervention improves physical fitness in Ecuadorian adolescents: a cluster-randomized controlled trial. *Int J Behav Nutr Phys Act* 2014;11:153. doi:10.1186/s12966-014-0153-5.

49.   Bartholomew JB, Jowers EM, Errisuriz VL, et al. A cluster randomized control trial to assess the impact of active learning on child activity, attention control, and academic outcomes: The Texas I-CAN trial. *Contemp Clin Trials* 2017;61:81-86. doi:10.1016/j.cct.2017.07.023.

50.   Chesham RA, Booth JN, Sweeney EL, et al. The Daily Mile makes primary school children more active, less sedentary and improves their fitness and body composition: a quasi-experimental pilot study. *BMC Med* 2018;16(1):64. doi:10.1186/s12916-018-1049-z.

51.   Christiansen LB, Toftager M, Boyle E, et al. Effect of a school environment intervention on adolescent adiposity and physical fitness. *Scand J Med Sci Sports* 2013;23(6):e381-e389. doi:10.1111/sms.12088.

52.   Cohen KE, Morgan PJ, Plotnikoff RC, et al. Physical activity and skills intervention: SCORES cluster randomized controlled trial. *Med Sci Sports Exerc* 2015;47(4):765-774. doi:10.1249/MSS.0000000000000452.

53.   Gorely T, Nevill ME, Morris JG, et al. Effect of a school-based intervention to promote healthy lifestyles in 7-11 year old children. *Int J Behav Nutr Phys Act* 2009;6:5. doi:10.1186/1479-5868-6-5.

54.   Kriemler S, Zahner L, Schindler C, et al. Effect of school based physical activity programme (KISS) on fitness and adiposity in primary schoolchildren: cluster randomised controlled trial. *BMJ* 2010;340:c785. doi:10.1136/bmj.c785.

55.   Lonsdale C, Sanders T, Cohen KE, et al. Scaling-up an efficacious school-based physical activity intervention: Study protocol for the 'Internet-based Professional Learning to help teachers support Activity in Youth' (iPLAY) cluster randomized controlled trial and scale-up implementation evaluation. *BMC Public Health* 2016. doi:10.1186/s12889-016-3243-2.

56.   Martínez-Vizcaíno V, Sánchez-López M, Salcedo-Aguilar F, et al. Protocol of a Randomized Cluster Trial to Assess the Effectiveness of the MOVI-2 Program on Overweight Prevention in Schoolchildren. *Revista Española de Cardiología (English Edition)* 2012;65(5):427-433. doi:10.1016/j.rec.2011.12.009.

57.   McGrane B, Belton S, Fairclough SJ, et al. Outcomes of the Y-PATH Randomized Controlled Trial: Can a School-Based Intervention Improve Fundamental Movement Skill Proficiency in Adolescent Youth? *J Phys Act Health* 2018;15(2):89-98. doi:10.1123/jpah.2016-0474.

58.   Møller NC, Tarp J, Kamelarczyk EF, et al. Do extra compulsory physical education lessons mean more physically active children--findings from the childhood health, activity, and motor performance school study Denmark (The CHAMPS-study DK). *Int J Behav Nutr Phys Act* 2014;11:121. doi:10.1186/s12966-014-0121-0.

59.   Morris JG, Gorely T, Sedgwick MJ, et al. Effect of the Great Activity Programme on healthy lifestyle behaviours in 7-11 year olds. *J Sports Sci* 2013;31(12):1280-1293. doi:10.1080/02640414.2013.781665.

60.   O'Leary M, Rush E, Lacey S, et al. Project Spraoi: two year outcomes of a whole school physical activity and nutrition intervention using the RE-AIM framework. *Irish Educational Studies* 2019;38(2):219-243. doi:10.1080/03323315.2019.1567368.

61.   Peralta LR, Jones RA, Okely AD. Promoting healthy lifestyles among adolescent boys: the Fitness Improvement and Lifestyle Awareness Program RCT. *Prev Med* 2009;48(6):537-542. doi:10.1016/j.ypmed.2009.04.007.

62.   Puder JJ, Marques-Vidal P, Schindler C, et al. Effect of multidimensional lifestyle intervention on fitness and adiposity in predominantly migrant preschool children (Ballabeina): cluster randomised controlled trial. *BMJ* 2011;343:d6195. doi:10.1136/bmj.d6195.

63.   Robbins LB, Ling J, Sharma DB, et al. Intervention Effects of "Girls on the Move" on Increasing Physical Activity: A Group Randomized Trial. *Ann Behav Med* 2019;53(5):493-500. doi:10.1093/abm/kay054.

64.   Sánchez-López M, Pardo-Guijarro MJ, Del Campo DG-D, et al. Physical activity intervention (Movi-Kids) on improving academic achievement and adiposity in preschoolers with or without attention deficit hyperactivity disorder: study protocol for a randomized controlled trial. *Trials* 2015;16:456. doi:10.1186/s13063-015-0992-7.

65.   Seljebotn PH, Skage I, Riskedal A, et al. Physically active academic lessons and effect on physical activity and aerobic fitness. The Active School study: A cluster randomized controlled trial. *Prev Med Rep* 2019;13:183-188. doi:10.1016/j.pmedr.2018.12.009.

66.   Taylor SL, Noonan RJ, Knowles ZR, et al. Evaluation of a Pilot School-Based Physical Activity Clustered Randomised Controlled Trial-Active Schools: Skelmersdale. *Int J Environ Res Public Health* 2018;15(5). doi:10.3390/ijerph15051011.

67.   Tercedor P, Villa-González E, Ávila-García M, et al. A school-based physical activity promotion intervention in children: rationale and study protocol for the PREVIENE Project. *BMC Public Health* 2017;17(1):748. doi:10.1186/s12889-017-4788-4.

68.   Weston KL, Azevedo LB, Bock S, et al. Effect of Novel, School-Based High-Intensity Interval Training (HIT) on Cardiometabolic Health in Adolescents: Project FFAB (Fun Fast Activity Blasts) - An Exploratory Controlled Before-And-After Trial. *PLoS One* 2016;11(8):e0159116. doi:10.1371/journal.pone.0159116.

69.   Zhou Z, Dong S, Yin J, et al. Improving Physical Fitness and Cognitive Functions in Middle School Students: Study Protocol for the Chinese Childhood Health, Activity and Motor Performance Study (Chinese CHAMPS). *Int J Environ Res Public Health* 2018;15(5). doi:10.3390/ijerph15050976.

70.   Tarp J, Domazet SL, Froberg K, et al. Effectiveness of a School-Based Physical Activity Intervention on Cognitive Performance in Danish Adolescents: LCoMotion-Learning, Cognition and Motion - A Cluster Randomized Controlled Trial. *PLoS One* 2016;11(6):e0158087. doi:10.1371/journal.pone.0158087.

71.   Hesketh KR, Lakshman R, van Sluijs EMF. Barriers and facilitators to young children's physical activity and sedentary behaviour: a systematic review and synthesis of qualitative literature. *Obes Rev* 2017;18(9):987-1017. doi:10.1111/obr.12562.

72.   Bailey R, Wellard I, Dismore H. Girls' participation in physical activities and sports: benefits, patterns, influences and ways forward. Centre for Physical Education and Sport Research, Canterbury Christ Church University College, U.K. for the World Health Organisation (WHO); 2005.

73.   Milanović Z, Sporiš G, Weston M. Effectiveness of high-intensity interval training (HIT) and continuous endurance training for VO2max improvements: a systematic review and meta-analysis of controlled trials. *Sports Med* 2015;45(10):1469-1481. doi:10.1007/s40279-015-0365-0.

74.   Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011;6:42. doi:10.1186/1748-5908-6-42.

75.   Höchsmann C, Dorling JL, Apolzan JW, et al. Baseline Habitual Physical Activity Predicts Weight Loss, Weight Compensation, and Energy Intake During Aerobic Exercise. *Obesity* 2020. doi:10.1002/oby.22766.

76.   Morton KL, Atkin AJ, Corder K, et al. The school environment and adolescent physical activity and sedentary behaviour: a mixed-studies systematic review. *Obes Rev* 2016;17(2):142-158. doi:10.1111/obr.12352.

77.   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(1):134. doi:10.1186/s12966-017-0586-8.

78.   Bauman AE, Reis RS, Sallis JF, et al. Correlates of physical activity: why are some people physically active and others not? *Lancet* 2012;380(9838):258-271. doi:10.1016/S0140-6736(12)60735-1.

79.   Hay J, Maximova K, Durksen A, et al. Physical activity intensity and cardiometabolic risk in youth. *Arch Pediatr Adolesc Med* 2012;166(11):1022-1029. doi:10.1001/archpediatrics.2012.1028.

80.   Raghuveer G, Hartz J, Lubans DR, et al. Cardiorespiratory fitness in youth: an important marker of health: a scientific statement from the American Heart Association. *Circulation* 2020;142:e101–e118. doi.org/10.1161/CIR.0000000000000866.

81.   Shephard RJ. The objective monitoring of physical activity. *Prog Prev Med* 2017;2(4):e0007. doi:10.1097/pp9.0000000000000007.

82.   Aubert S, Barnes JD, Abdeta C, et al. Global Matrix 3.0 Physical Activity Report Card Grades for Children and Youth: Results and Analysis From 49 Countries. *J Phys Act Health* 2018;15(S2):S251-S273. doi:10.1123/jpah.2018-0472.

83.   Tierney JF, Vale C, Riley R, et al. Individual participant data (IPD) meta-analyses of randomised controlled trials: guidance on their use. *PLoS Med* 2015;12(7):e1001855. doi.org/10.1371/journal.pmed.1001855.

84.   Bullock JG, Green DP, Ha SE. Yes, but what's the mechanism? (don't expect an easy answer). *J Pers Soc Psychol* 2010;98(4):550-558. doi:10.1037/a0018933.

85.   Lang JJ, Wolfe Phillips E, Orpana HM, et al. Field-based measurement of cardiorespiratory fitness to evaluate physical activity interventions. *Bull World Health Organ* 2018;96(11):794-796. doi:10.2471/BLT.18.213728.

86.   Miyamoto-Mikami E, Zempo H, Fuku N, et al. Heritability estimates of endurance-related phenotypes: A systematic review and meta-analysis. *Scand J Med Sci Sports* 2018;28(3):834-845. doi:10.1111/sms.12958.

87.   Baxter-Jones ADG, Maffulli N. Endurance in young athletes: it can be trained. *Br J Sports Med* 2003;37(2):96-97. doi:10.1136/bjsm.37.2.96.

88.   Ruiz JR, Cavero-Redondo I, Ortega FB, et al. Cardiorespiratory fitness cut points to avoid cardiovascular disease risk in children and adolescents; what level of fitness should raise a red flag? A systematic review and meta-analysis. *Br J Sports Med* 2016;50(23):1451-1458. doi:10.1136/bjsports-2015-095903.

89.   Ross JS. Clinical research data sharing: what an open science world means for researchers involved in evidence synthesis. *Syst Rev* 2016;5:159. doi.org/10.1186/s13643-016-0334-1.

90.   van Middelkoop M, Lohmander S, Bierma-Zeinstra SMA. Sharing data–taming the beast: barriers to meta-analyses of individual patient data (IPD) and solutions. *BJSM* 2020;0:1–2. doi: 10.1136/bjsports-2019-101892.

FIGURE LEGENDS

Figure 1. Hybrid PRISMA and CONSORT diagram, showing the flow of studies and data through the pooled analysis process.

Figure 2 Title. Treatment Effects by Student-level Moderator

Figure 2 Legend. The effect of interventions on CRF, VPA, and MPA by student-level moderators. CRF = Cardiorespiratory fitness (VO_{2Peak} mL/kg/min), VPA = Vigorous physical activity (per day percentage units), MPA = Moderate physical activity (per day percentage units). CRF, VPA, MPA, and age are continuous variables stratified into bottom, middle, and upper tertiles for moderator analysis. Gender (Boy / Girl) is categorical. Effects are unstandardized beta-coefficients.

Figure 3 Title. Structural equation model from primary pooled analysis.

Figure 3 Legend. All coefficients are unstandardised. Estimates drawn from within-study variance-variance explained once the effects of different interventions/studies have been accounted for. Outcomes for CRF and physical activity are follow-up. For all outcomes, age, BMI, and baseline scores for that outcome were included as control variables. Indirect effect of intervention on VO_{2Peak} via MPA is Path AB. Indirect effect of intervention on VO_{2Peak} via VPA is Path CD. Total effect of intervention on VO_{2Peak} = Path AB + Path CD + Path E.

Note. CRF = Cardiorespiratory fitness (VO_{2Peak}), VPA = Vigorous physical activity, MPA = Moderate physical activity.
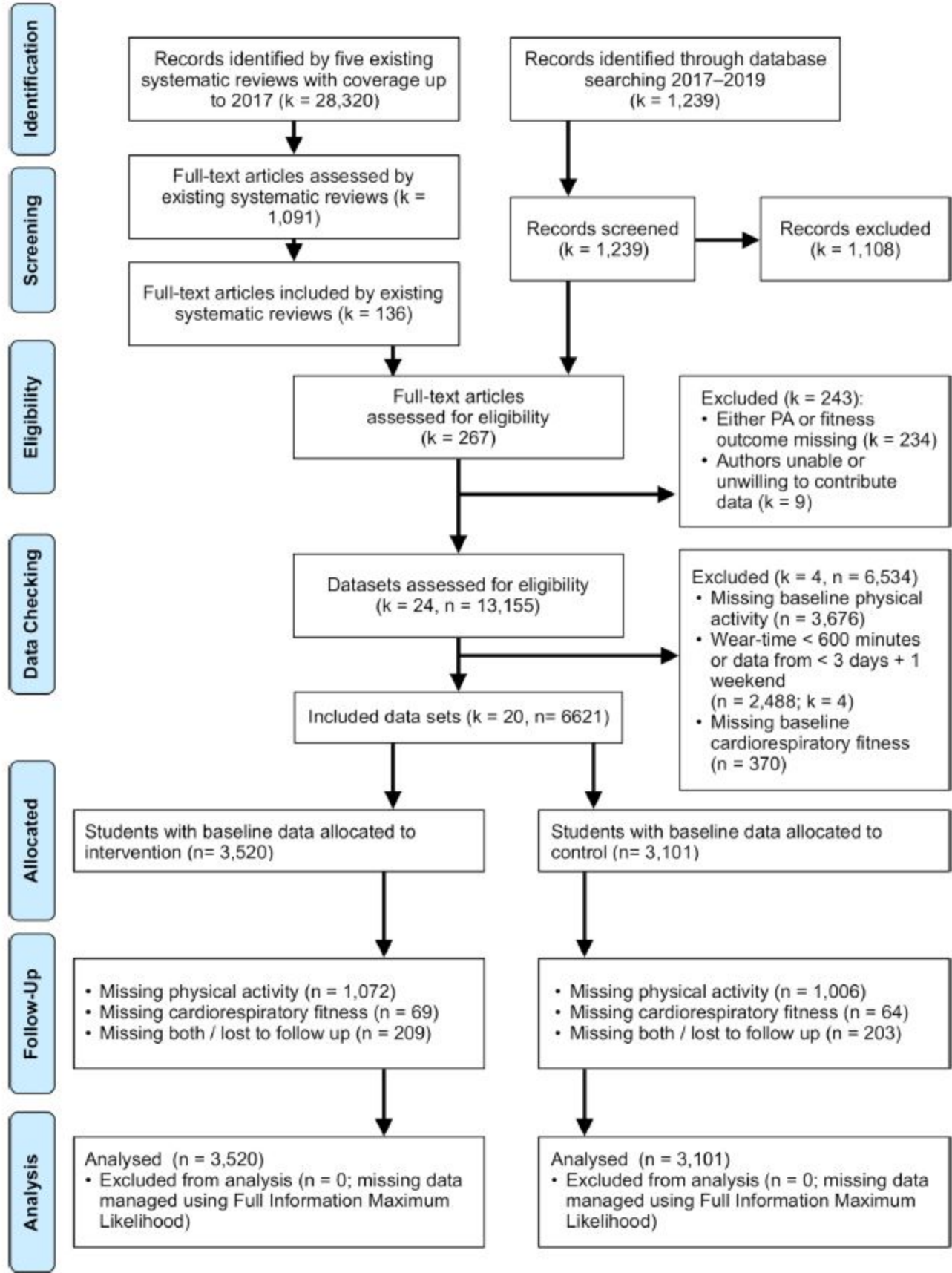
Figure 1 Hybrid Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Consolidated Standards of Reporting Trials (CONSORT) diagram showing the flow of studies and data through the pooled analysis process.
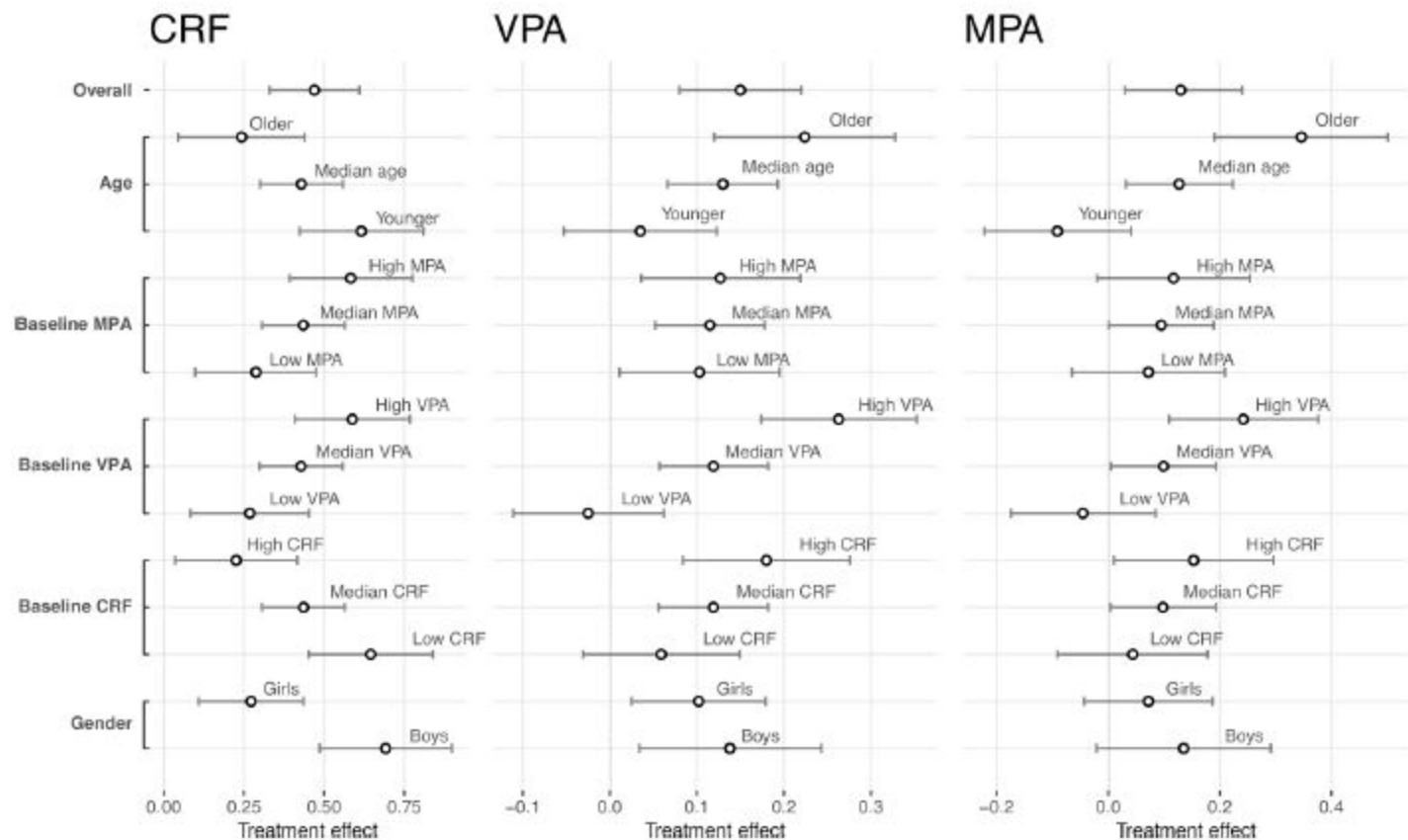
Figure 2 Treatment effects by student-level moderator. The effect of interventions on CRF, VPA and MPA by student-level moderators. CRF, VPA, MPA and age are continuous variables stratified into bottom, middle and upper tertiles for moderator analysis. Gender (boy / girl) is categorical. Effects are unstandardised beta-coefficients. CRF, cardiorespiratory fitness (VO_{2Peak} mL/kg/min); MPA, moderate physical activity (per day percentage units); VPA, vigorous physical activity (per day percentage units).
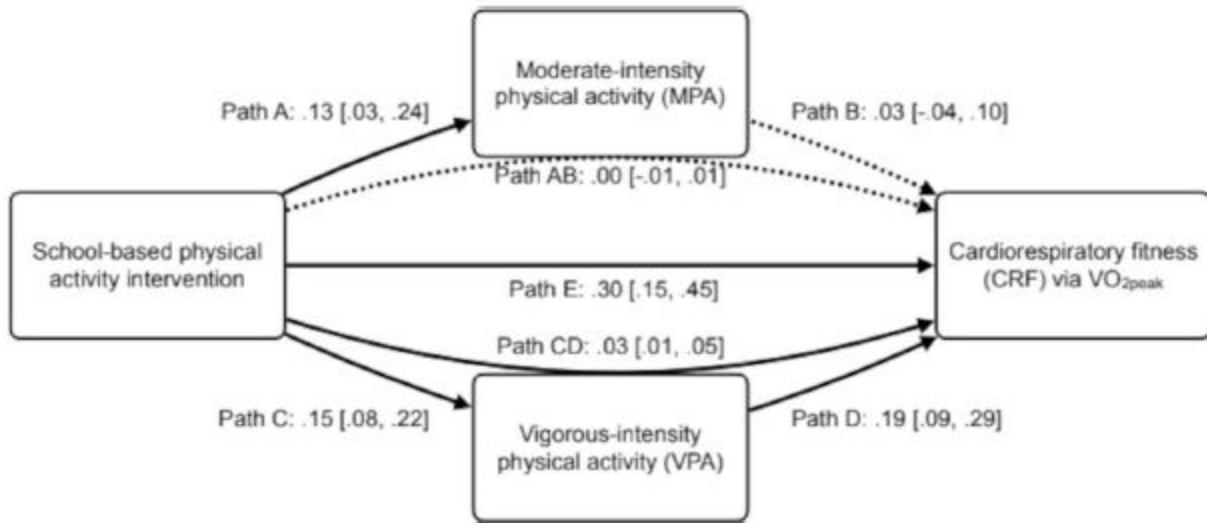
Figure 3 Structural equation model from primary pooled analysis. All coefficients are unstandardised. Estimates drawn from within-study variance–variance explained once the effects of different interventions/studies have been accounted for. Outcomes for CRF and physical activity are follow-up. For all outcomes, age, BMI and baseline scores for that outcome were included as control variables. Indirect effect of intervention on VO_{2Peak} via MPA is path AB. Indirect effect of intervention on VO_{2Peak} via VPA is path CD. Total effect of intervention on VO_{2Peak}=path AB+path CD+path E. BMI, body mass index; CRF, cardiorespiratory fitness (VO_{2Peak}); MPA, moderate physical activity; VPA, vigorous physical activity.