

The impact of healthy lifestyle interventions on inequalities in physical fitness: Evidence from over 4 million SLOfit observations

Antonio Martinko

antonio.martinko@fsp.uni-lj.si

University of Ljubljana

Gregor Jurak

University of Ljubljana

Gregor Starc

University of Ljubljana

Maroje Sorić

University of Zagreb

Research Article

Keywords: childhood obesity, fitness disparities, equitable, school policies, comparative analysis

Posted Date: October 6th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11038732/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

We examined whether nationwide school-based healthy lifestyle policies influenced physical fitness trajectories and heterogeneity among children with normal weight (NW) and excess weight (EW).

Methods

We analysed Slovenian SLOfit surveillance data from nearly 300,000 children aged 6–14 years assessed annually from 2010 to 2019. We classified students by baseline weight status and exposure to Healthy Lifestyle (HLS) and Healthy Schools Network (HSN) policies. Multilevel models estimated changes in cardiorespiratory fitness, muscular fitness, and body composition, including mean trajectories and residual variability.

Results

Cardiorespiratory and most muscular fitness outcomes improved, but gains were greater among children with NW than EW, resulting in persistent or widening disparities. Intervention settings differed more in their influence on inequalities than average performance. Combined HLS and HSN generally produced the most balanced improvements, whereas no intervention more often showed widening inequalities. Children with EW also exhibited greater heterogeneity in developmental trajectories.

Conclusions

School-based policies differed more in their influence on fitness inequalities than average fitness levels. Combined HLS and HSN generally produced the most equitable trajectories, but no intervention consistently reduced heterogeneity, highlighting the importance of evaluating both average effects and variability.

Funding

Horizon 2020 (Grant no. 774548), Slovenian National Research and Innovation Agency (Grant no. P5-0142), Croatian Science Foundation (Grant no: DOK-2020-01-3728).

Introduction

Obesity-related health disparities are a persistent global problem, particularly among disadvantaged groups. Although healthy lifestyle interventions aim to improve population health, their effects on health inequalities are not always predictable.¹ Having overweight or obesity during childhood affects cardiometabolic risk at those ages,^{2–4} higher all-cause mortality later on,² with substantial social and economic consequences related to an earlier chronic disease burden.⁵ Importantly, the combination of excess weight and low physical fitness may further increase cardiometabolic risk,⁶ while low fitness is also associated with reduced physical functioning, earlier frailty, and fewer healthy and productive years throughout adulthood.⁷ Physical fitness is therefore a well-established marker of health throughout childhood and adolescence⁸, where higher levels of cardiorespiratory fitness⁷ and muscular fitness⁹ are associated not only with immediate health but also with long-term physical, cognitive, and functional outcomes.

School-based physical activity interventions can improve physical fitness while reducing fatness and cardiometabolic risk factors among children and adolescents.¹⁰ Schools provide a potentially effective setting for such interventions because of their broad reach and regular contact with children.¹¹ However, the effectiveness and equity of different intervention approaches remain uncertain.¹² Although school-based programmes can reduce sedentary behaviour, their effects on moderate-to-vigorous physical activity vary, with greater effectiveness reported when interventions include multiple components or improve programme quality.¹³ Some interventions have also reduced physical activity inequalities, particularly among low-active and socioeconomically disadvantaged children.^{14,15} Population-scale school-based interventions may similarly have greater effects on BMI among children with excess weight.¹⁶ However, little is known about whether school-based healthy lifestyle interventions influence variability in physical fitness, and therefore whether improvements are distributed similarly across children.

Focusing only on mean changes may obscure important differences in individual responses. An intervention may increase average fitness while simultaneously reducing variability, suggesting broadly shared improvements, or increase variability, indicating that some children benefit substantially while others show little progress. Thus, examining mean trajectories together with variability may provide a more complete assessment of whether interventions reduce or widen disparities within groups.

We tried to address priorities identified in the Sustainable Development Goals, particularly SDGs 3 and 10, concerning health and well-being and the reduction of inequalities.^{17,18} We therefore examined whether nationwide school-based healthy lifestyle policies were associated with changes in physical fitness and its variability among children with excess weight (EW) and normal weight (NW). Specifically, we assessed whether mean fitness trajectories and within-group variability changed differently according to intervention exposure and weight status, and whether disparities between children with EW and NW widened or narrowed over time. To our knowledge, no previous study has explicitly examined temporal changes in fitness variability across both weight status and school-based intervention exposure. We hypothesised that improvements in physical fitness would be smaller among children with EW than NW and that intervention-related differences in variability would emerge between the two groups. By examining both mean performance and variability, we aimed to provide evidence relevant to the design and evaluation of school-based policies that promote more equitable physical fitness improvements.

Methods

Study design, participants, and measurements

We conducted a population-based study using data from the Slovenian SLOfit national monitoring system, which annually assesses physical fitness and motor development in schoolchildren.¹⁹ Nearly 300,000 children aged 6–14 years, approximately equally divided by sex, were assessed annually between 2010 and 2019, representing more than 94% of Slovenian schoolchildren in ISCED levels 1 and 2.²⁰ We classified weight status using age- and sex-specific Slovenian national BMI standards.²¹ We combined children with overweight or obesity (BMI > 85th or > 95th centile, respectively) into a group living with excess weight, with all remaining children classified as living with normal weight. Children were further categorised according to exposure to two nationwide school-based

healthy lifestyle policies: Healthy Lifestyle (HLS) and Healthy Schools Network (HSN), resulting in four groups: no intervention, HSN only, HLS only, and both interventions.

HLS, introduced in 2010/11, provided two or three additional weekly physical education lessons, depending on age/grade, taught by specialist physical education teachers. Participation was voluntary at the school level, with schools deciding whether to implement the programme; once enrolled, the intervention was available to all children in the school. HSN, established in 1993 as part of a WHO initiative, promotes school-based health through annual programmes addressing topics such as healthy diet, physical activity, interpersonal relationships, and mental health, but does not generally provide additional physical education lessons. A detailed description of both interventions is provided in Additional file 3.

The fitness battery included body composition (BMI and triceps skinfold thickness [TSF]), cardiorespiratory fitness (600-m run [R600]), and muscular fitness (60-s sit-ups [SU60], bent-arm hang [BAH], and standing broad jump [SBJ]). These tests, all part of the SLOfit system test battery,²² were developed years before Eurofit²³ and were modelled after the extensive physical fitness testing conducted in Slovenia in the 1960s and 1970s. Centile values converted from raw test results are based on the Slovenian national fitness standards.²⁴ All measurements are conducted in schools with identical protocols across all assessment waves. Physical education teachers, with extensive training in anthropometry and a level of precision beyond the requirements of the SLOfit system, perform the measurements using the necessary measuring tools each April. More details on protocols and outlier treatment can be found elsewhere.²⁴

Statistical analysis

Our primary aim was to determine whether interventions reduced or exacerbated disparities in fitness development over time. We therefore examined both predicted mean differences and residual variability, assessing whether developmental trajectories were more heterogeneous among children with EW and whether this varied by intervention exposure.

We used multilevel models (MLMs) to account for the hierarchical structure of the data. Models were developed sequentially from null to final specifications using likelihood ratio tests and the Akaike Information Criterion (AIC), with model assumptions assessed using diagnostic plots. Analyses were stratified by sex to account for established differences in physical fitness and developmental trajectories between boys and girls.²⁵

Due to complex developmental patterns, we opted to use age- and sex-specific centile values instead of raw scores for each physical fitness test. The centile values were derived from the Slovenian national reference standards based on SLOfit measurements collected between 1988 and 2019.¹⁹ Age- and sex-specific centile values were used as dependent variables for BMI, TSF, R600, SU60, BAH, and SBJ. For BMI, TSF, and R600, higher values indicate poorer performance, whereas the opposite applies to muscular fitness measures. Time was treated as a categorical factor, with primary comparisons between 2010 and 2019. Models included weight status, intervention group, age group (i.e., International Standard Classification of Education - ISCED-level²⁰), and height percentile as a covariate to account for maturation, given known differences in developmental timing across weight status groups.²⁶ Fixed effects included the three-way interaction between weight status, intervention, and time, while group-specific variance structures were used to assess differences in residual variability across subgroups and over time. Estimated coefficients are reported with 95% confidence intervals (CIs).

Model-based predicted values were extracted for presentation in Table S2.4 and to visualise time trends in Figure S1.1. Complete model results are provided in Additional Material 2.

All statistical analyses were performed with R version 4.3.0.²⁷

Role of the funding source

The study's design, data collection, management, analysis, interpretation, report writing, and publication submission were all unrelated to the funding bodies.

Results

Study sample

We included 265,206 children (girls: 127,676; 48.1%), aged 10.1 ± 2.4 years (range: 6–14), who had their fitness measured annually from 2010 to 2019. In the Additional Material, we show the exact number of participants divided by sex, age, weight status and intervention group for all years observed in this study. About 90,000 children were included per year on average, with yearly totals ranging from 71,000 to 107,000.

Regarding age groups, henceforth ISCED level, over the time course of the study, $n = 190,362$ ($n = 91,802$ girls) children belonged to the younger group, i.e., 6–10 years old, leaving approximately the same number of children in the older group ($n = 179,468$ [$n = 85,842$ girls]), i.e., 11–14 years old. Over the whole study period, we included $n = 202,437$ ($n = 97,794$ girls) children with NW, and $n = 62,769$ ($n = 29,882$ girls) children with EW. In total, we included children who were solely involved in one of the interventions, only HLS $n = 62,107$ ($n = 29,907$ girls), only HSN $n = 84,732$ ($n = 40,762$ girls), both $n = 78,164$ ($n = 37,866$ girls) or none $n = 40,203$ ($n = 19,141$ girls) of the interventions. The number of children and descriptive statistics are shown in Additional Tables S1.1, S1.2 and S1.3 for each level of interest. We reported the results from the final, fully adjusted models for each outcome in Additional Material 2.

Within-group Differences Analyses Interpretation

Residual variance estimates for all fitness outcomes were examined at baseline (2010) and follow-up (2019) among boys, stratified by weight status and intervention exposure. Attention was given to changes in variance over time as an indicator of divergence or convergence in developmental trajectories, and to whether intervention exposure modified these temporal patterns. When we present the results, the reference group was one chosen by default in our models, i.e., the first level of each categorical variable of interest in our data. Based on this point, the reference group for all models are older children living with NW in 2010, going to HLS-only schools. Baseline (2010) and final (2019) results regarding within-group differences for each physical fitness test across weight status and intervention groups stratified by sex can be seen in Additional Material 2, Table S2.1. For ease of interpretation, in the Results section we have also presented changes in within-group differences as the percentage change in the estimated variance relative to baseline (2010). We want to note that any percentage we report is the relative change in the model-estimated residual variance between 2010 and 2019, not a direct percentage change in observed variability. Results of multilevel regression analysis for each physical fitness test with estimated differences between intervention groups by weight status and sex can also be seen in Additional Material 2, Table S2.2. For between-group differences, we have provided the detailed results of multilevel

regression analysis for each physical fitness test with estimates and corresponding 95% CIs by sex in Additional Material 2, Table S2.3.

Cardiorespiratory fitness

600-meter run (R600)

Within-group differences

From 2010 to 2019, within-group variability among boys with NW remained broadly stable across intervention groups: no intervention – 3.2% (316.8 to 306.7), HSN only – 1.6% (330.5 to 325.3), HLS only – 2.9% (326.8 to 317.3), and both interventions + 0.7% (328.8 to 331.1). In contrast, variability increased among boys with EW in all groups: +5.1% (260.4 to 273.6), + 13.0% (250.3 to 282.9), + 9.8% (238.0 to 261.4), and + 6.1% (255.6 to 271.3), respectively. Thus, the difference between boys with NW and EW was more evident under all intervention scenarios, particularly HSN only. A similar pattern was observed among girls with NW, although variability increased under no intervention (+ 3.9%; 306.6 to 318.5) and HSN only (+ 0.3%; 332.9 to 333.9), remained unchanged under HLS only (0.0%; 336.7 to 336.7), and decreased under both interventions (– 10.5%; 364.4 to 326.3). Among girls with EW, variability increased under no intervention (+ 6.9%; 248.7 to 265.9), HSN only (+ 8.9%; 255.1 to 277.7), and HLS only (+ 0.4%; 257.0 to 258.1), but decreased slightly under both interventions (– 2.1%; 284.7 to 278.6). Overall, increases in variability were more pronounced among EW children, whereas the combined intervention was associated with the most favourable pattern, particularly among girls.

Between-group differences

From 2010 to 2019, average cardiorespiratory fitness increased across all intervention groups, but improvements were generally greater among boys with NW. Among boys, improvements in NW versus EW groups were + 5.0 versus + 2.8 centiles under no intervention (51.4 to 56.4; 35.1 to 37.9), + 4.4 versus + 2.6 under HSN only (50.7 to 55.1; 34.1 to 36.7), + 3.6 versus + 2.7 under HLS only (50.2 to 53.8; 32.0 to 34.7), and + 3.9 versus + 3.7 under both interventions (51.4 to 55.3; 32.4 to 36.1). Accordingly, the NW–EW difference increased from approximately 16.3 to 18.5 under no intervention, 16.6 to 18.4 under HSN only, and 18.2 to 19.1 under HLS only, but remained essentially stable under both interventions (~ 19.0 to 19.2). The more balanced improvement under both interventions was particularly evident between 2013 and 2016, when boys with EW improved relatively more rapidly, after which trajectories became largely parallel (Additional Material 1, Figure S1.1).

Among girls, NW versus EW improvements were + 5.3 versus + 2.3 centiles under no intervention (51.5 to 56.8; 38.7 to 41.0), + 4.9 versus + 2.2 under HSN only (51.0 to 55.9; 38.1 to 40.3), + 4.4 versus + 1.0 under HLS only (49.1 to 53.5; 35.7 to 36.7), and + 4.7 versus + 2.9 under both interventions (49.6 to 54.4; 36.2 to 39.1). The corresponding NW–EW differences changed from 12.9 to 15.8, 12.9 to 15.6, 13.4 to 16.8, and 13.4 to 15.3, respectively. Thus, the combined intervention again resulted in the smallest increase in the NW–EW difference, whereas the largest widening occurred under HLS only. Detailed model estimates and 95% CIs are provided in Additional Material 2, Table S2.3.

Muscular fitness

60s sit-ups (60SU)

Within-group differences

From 2010 to 2019, within-group variability showed a consistent contrast by weight status across intervention groups. Among children with NW, variability generally decreased, with changes of – 7.1% (410.3 to 381.1), – 4.2% (390.1 to 373.6), – 8.6% (383.6 to 350.7), and – 7.7% (409.6 to 377.9) under no intervention, HSN only, HLS only, and both interventions, respectively, among boys; corresponding changes among girls were – 7.7% (409.6 to 377.9), – 11.8% (396.1 to 349.3), – 5.8% (364.2 to 343.0), and – 10.4% (399.5 to 357.9). The only exception was a slight increase under HLS only among boys with NW (+ 2.6%; 367.8 to 377.5). In contrast, variability among children with EW generally increased under no intervention (+ 3.4%; 419.6 to 433.9 in boys; +4.3%; 400.0 to 417.1 in girls), HSN only (+ 1.3%; 391.3 to 396.3; –0.7%; 409.2 to 406.2), and HLS only (+ 22.8%; 340.0 to 417.6; +14.1%; 364.8 to 416.3), whereas both interventions reduced variability (– 8.8%; 393.6 to 358.9 in girls) or produced only a small increase (+ 3.4%; 380.2 to 393.2 in boys). Thus, the contrast between intervention groups was most pronounced among children with EW, with HLS only showing the largest increase in variability and the combined intervention the most favourable pattern (Fig. 2).

Between-group differences

Average sit-ups performance increased across all intervention groups, but improvements were consistently greater among children with NW than EW, resulting in a widening of the NW–EW difference under every intervention. Among boys, NW versus EW improvements were + 9.4 versus + 2.3 centiles under no intervention (53.8 to 63.1; 48.6 to 50.9), + 6.1 versus – 0.8 under HSN only (55.3 to 61.4; 50.9 to 50.0), + 6.2 versus + 2.8 under HLS only (54.5 to 60.8; 46.9 to 49.7), and + 6.5 versus + 2.5 under both interventions (55.3 to 61.7; 48.6 to 51.1). The corresponding NW–EW differences changed from approximately 5.2 to 12.3, 4.4 to 11.3, 7.6 to 11.1, and 6.7 to 10.6, respectively. Among girls, improvements were + 6.6 versus + 2.3 under no intervention (55.0 to 61.6; 49.7 to 52.0), + 5.2 versus – 1.2 under HSN only (56.0 to 61.2; 52.3 to 51.0), + 4.5 versus + 0.9 under HLS only (55.5 to 60.0; 49.9 to 50.7), and + 5.3 versus + 2.2 under both interventions (56.0 to 61.2; 51.0 to 53.2). The corresponding differences changed from approximately 5.3 to 9.7, 3.7 to 10.2, 5.6 to 9.3, and 4.9 to 8.0. Thus, although the NW–EW difference increased under all intervention scenarios, the combined intervention consistently showed the smallest widening in both boys and girls, indicating relatively greater improvements among EW children. This pattern was particularly evident between 2013 and 2016, after which trajectories became largely parallel (Additional Material 1, Figure S1.2).

Bent-arm hang (BAH)

Within-group differences

From 2010 to 2019, within-group variability was generally more stable or decreased among children with NW, whereas changes among those with EW were more variable and tended to increase under intervention exposure. Among boys with NW, variability changed by – 10.3% (332.7 to 298.4), 0.0% (318.6 to 318.6), – 0.2% (312.1 to 311.5), and – 1.6% (312.2 to 307.1) under no intervention, HSN only, HLS only, and both interventions, respectively. Among girls with NW, corresponding changes were – 9.3% (308.1 to 279.5), + 1.7% (297.2 to 302.3), – 0.7% (290.3 to 288.3), and + 0.4% (291.4 to 292.5). In contrast, among boys with EW, variability decreased under no intervention (– 10.7%; 213.3 to 190.5) but increased under HSN only (+ 4.1%; 204.3 to 212.6), HLS only (+ 13.4%; 192.8 to 218.7), and both interventions (+ 4.4%; 192.8 to 201.3). Among girls with EW, variability decreased under no intervention (– 4.7%; 197.3 to 188.1) and HLS only (– 9.9%; 217.6 to 196.0), but increased under HSN only (+

2.6%; 213.3 to 218.8) and both interventions (+ 4.0%; 198.0 to 205.9). Thus, intervention-related increases in variability were more evident among children with EW than with NW, particularly under HLS only (Fig. 3).

Between-group differences

Average upper body strength improved among children with NW but showed little improvement or declined among those with EW, resulting in a widening of the NW–EW difference across all intervention groups. Among boys, NW versus EW changes were + 4.0 versus – 0.1 centiles under no intervention (49.4 to 53.4; 29.7 to 29.6), + 2.6 versus – 1.0 under HSN only (50.0 to 52.6; 30.3 to 29.3), + 1.9 versus – 0.3 under HLS only (50.5 to 52.4; 29.0 to 28.7), and + 2.4 versus 0.0 under both interventions (50.3 to 52.7; 29.1 to 29.1). The corresponding NW–EW differences changed from approximately 19.7 to 23.8, 19.7 to 23.3, 21.4 to 23.7, and 21.1 to 23.6, respectively. Among girls, NW versus EW changes were + 5.1 versus – 1.2 under no intervention (50.6 to 55.7; 33.1 to 31.9), + 4.4 versus – 2.2 under HSN only (51.4 to 55.8; 34.1 to 31.9), + 3.9 versus – 3.4 under HLS only (51.6 to 55.5; 34.5 to 31.1), and + 4.0 versus – 2.4 under both interventions (50.8 to 54.8; 34.0 to 31.6). The corresponding differences changed from approximately 17.5 to 23.8, 17.3 to 23.9, 17.1 to 24.4, and 16.8 to 23.1. Thus, the NW–EW difference increased under every intervention, although the combined intervention showed the smallest widening among girls and a modest slowing of divergence among boys (Fig. 3; Additional Material 1, Figure S1.3).

Standing broad jump (SBJ)

Within-group differences

From 2010 to 2019, within-group variability generally increased among children with NW, whereas changes among those with EW were smaller and less consistent. Among boys with NW, variability changed by + 12.4% (260.5 to 292.7), – 2.2% (284.2 to 278.0), + 5.5% (266.4 to 281.1), and + 6.2% (275.6 to 292.8) under no intervention, HSN only, HLS only, and both interventions, respectively; among girls with NW, corresponding changes were + 6.1% (286.8 to 304.4), + 4.5% (282.1 to 294.9), + 13.0% (262.4 to 296.6), and + 13.2% (265.9 to 300.9). Among boys with EW, changes were + 1.9% (255.3 to 260.1), + 3.7% (232.7 to 241.4), + 3.0% (232.1 to 239.0), and – 2.2% (225.9 to 220.9), respectively, while among girls with EW they were – 8.0% (266.6 to 245.3), + 1.1% (240.4 to 243.0), + 5.9% (250.4 to 265.1), and + 22.4% (223.8 to 273.9). Thus, variability increased more consistently among children with NW, whereas changes among children with EW were more heterogeneous, including a marked increase among girls under both interventions (Fig. 4).

Between-group differences

Average lower body power remained stable or declined among children with NW but declined more consistently among those with EW, resulting in wider NW–EW differences across all intervention groups. Among boys, NW versus EW changes were – 0.7 versus – 3.6 centiles under no intervention (51.3 to 50.6; 37.4 to 33.8), – 0.6 versus – 3.6 under HSN only (50.4 to 49.8; 37.4 to 33.8), – 1.7 versus – 3.4 under HLS only (50.3 to 48.6; 35.8 to 32.4), and – 2.7 versus – 3.7 under both interventions (50.6 to 48.0; 35.9 to 32.1). The corresponding NW–EW differences changed from approximately 13.9 to 16.8, 13.0 to 16.0, 14.5 to 16.2, and 14.8 to 15.9, respectively. Among girls, NW versus EW changes were + 0.9 versus – 5.1 under no intervention (49.9 to 50.8; 41.4 to 36.2), – 1.0 versus – 6.1 under HSN only (50.1 to 49.1; 41.3 to 35.2), 0.0 versus – 5.1 under HLS only (49.2 to 49.1; 40.1 to 35.0), and – 2.1 versus – 5.2 under both interventions (49.6 to 47.6; 40.0 to 34.8). The corresponding differences changed from approximately 8.5 to 14.6, 8.8 to 13.9, 9.1 to 14.1, and 9.7 to 12.8, respectively. Thus, the NW–EW difference increased under all intervention scenarios, with no clear evidence that any intervention substantially

reduced this inequality; the smallest widening occurred under the combined intervention in both boys and girls (Fig. 4; Additional Material 1, Figure S1.4).

Body Composition

Body mass index (BMI) and triceps skinfold (TSF)

Changes in body composition from 2010 to 2019 were generally small and showed limited differences between intervention groups (Additional Material 2, Figure S2.1 and Figure S2.2). For BMI, between-group differences remained relatively stable or increased slightly over time, reflecting declining or stable BMI centiles among children with NW and stable (boys) or increasing (girls) BMI centiles among those with EW. In contrast, TSF centiles changed little across follow-up, with between-group differences showing slight convergence among boys and remaining largely unchanged among girls. Overall, none of the intervention approaches was associated with substantial improvements in body composition, and intervention-related differences were considerably smaller than those observed for physical fitness outcomes. We have provided a more detailed explanation of the results in Additional Material 2.

Discussion

In this comprehensive quasi-experimental study, we evaluated physical fitness trajectories among Slovenian schoolchildren between 2010 and 2019 using more than 4.2 million observations from over three decades of national surveillance. Beyond mean performance, we examined interindividual variability to determine whether fitness became more homogeneous or disparate according to weight status and exposure to HLS, HSN, both interventions, or no intervention. This allowed us to assess whether interventions were associated not only with changes in average fitness but also with changes in disparities within and between weight-status groups.

The main finding was that patterns of variability were outcome-specific and often differed from those based on mean performance alone. Across several fitness domains, variability was more likely to remain stable or decrease among children with NW, whereas it more often increased among children with EW, particularly in HSN-only and no-intervention schools. Combined interventions generally produced the most balanced trajectories among children with EW and the smallest increases in between-group disparities, although they did not consistently reduce variability across all outcomes. Across both boys and girls, variance structures differed markedly by weight status and fitness domain, revealing heterogeneity patterns that were not always aligned with mean-level trends (shown in Figs. 1–4, Figure S2.1 and S2.2 for within-group differences and Figure S1.1 for between-group differences).

For cardiorespiratory fitness, variance remained relatively stable among children with NW across intervention groups, whereas it increased more clearly among children with EW, especially in HSN-only and no-intervention schools. This indicates widening dispersion in endurance performance within the excess-weight group over time. Among muscular fitness components, patterns differed by test. For body core strength endurance, variance decreased among children with NW but increased markedly among children with EW across most intervention settings, pointing to growing heterogeneity in abdominal strength endurance among the latter. For upper body strength endurance, dispersion generally declined among children with NW but showed mixed changes among children with EW, with increases in some intervention groups. For lower body power, variance increased modestly in most groups, though increases were more pronounced among children with EW in several settings. Changes in

body composition were comparatively small. BMI and TSF showed limited differences between intervention settings, and none of the interventions was consistently associated with meaningful reductions in inequalities. Overall, intervention-related differences were more apparent for physical fitness than for body composition.

In summary, from baseline to 2019, variance patterns point to a recurring theme: dispersion frequently narrowed or remained stable among children with NW, whereas it more often increased among children with EW, particularly within certain intervention contexts. These findings underscore that changes over time were not limited to mean shifts but involved evolving heterogeneity within weight-status and intervention groups across multiple domains of physical fitness.

Across the 10-year follow-up, cardiorespiratory fitness improved in children with both EW and NW regardless of intervention exposure, although intervention groups differed more in the extent to which they influenced inequalities than average fitness levels. Among children with EW, schools implementing both HLS and HSN showed the most favourable trajectories in both sexes, with the largest improvements in average cardiorespiratory fitness and a temporary acceleration of gains during the middle years of follow-up. Improvements were smaller in HLS-only, HSN-only and particularly no-intervention schools. Despite these gains, variability generally increased over time, indicating increasingly heterogeneous responses within the EW group. This increase in variability was most evident in HSN-only schools, whereas schools implementing both interventions showed the smallest increases and, among girls, even slight reductions in variability. Among children with NW, cardiorespiratory fitness also improved across all intervention settings, with relatively small differences between intervention groups. Variability remained largely stable throughout follow-up, with only modest changes across intervention settings. The greatest reduction in variability was observed among girls attending schools implementing both interventions, whereas boys generally showed little change regardless of intervention exposure. Overall, children with NW consistently achieved larger improvements in average cardiorespiratory fitness than children with EW, resulting in persistent or slightly widening disparities under most intervention settings. Schools implementing both interventions demonstrated the most equitable trajectories by producing relatively greater improvements among children with EW and limiting increases in within-group variability, although differences between weight groups were not eliminated.

Across all intervention settings, body core strength improved over time, although gains differed substantially according to weight status and intervention exposure. Among children with EW, improvements were generally modest. Schools implementing both HLS and HSN showed the most balanced trajectories in both sexes, whereas HSN-only schools showed little or no improvement and HLS-only schools produced only limited gains. Variability increased under most intervention settings, particularly among boys attending HLS-only schools. Among girls, schools implementing both interventions represented the only setting in which variability decreased, indicating more homogeneous responses over time. Among children with NW, average performance improved across all intervention groups, with the largest gains generally observed under no intervention and combined interventions. Variability decreased consistently in most intervention settings, particularly among girls, indicating increasingly homogeneous performance over time. Only boys attending HLS-only schools showed a slight increase in variability. Overall, disparities between weight groups widened under every intervention scenario because improvements were consistently greater among children with NW. Nevertheless, schools implementing both interventions showed the smallest increases in inequalities by combining relatively larger gains among children with EW with reductions in within-group variability.

Upper-body strength demonstrated less favourable trajectories than cardiorespiratory or body core strength, particularly among children with EW. Among children with EW, average performance remained largely unchanged in boys and declined among girls regardless of intervention exposure. Combined interventions showed the least pronounced decline, whereas HLS-only schools generally demonstrated the poorest trajectories. Variability showed mixed patterns. Boys experienced increasing variability under all intervention settings except no intervention, while girls demonstrated reductions under no intervention and HLS-only schools but modest increases in HSN-only and combined intervention schools. Among children with NW, average performance improved modestly across all intervention settings. Variability generally decreased or remained stable, particularly in boys attending no-intervention schools and girls attending no-intervention or HLS-only schools. Overall, inequalities between weight groups increased consistently because children with NW improved while children with EW showed little improvement or declining performance. Although combined interventions appeared to slow this divergence slightly, none of the intervention approaches substantially reduced these inequalities.

Lower-body power showed the least favourable trends across the study period. Among children with EW, average performance declined across all intervention settings in both sexes. Combined interventions produced only marginally smaller declines than the remaining intervention groups. Variability changed inconsistently. Boys showed relatively small changes regardless of intervention, whereas girls experienced increasing variability under every intervention except no intervention, where variability declined. Among children with NW, average lower-body power remained relatively stable or declined only slightly across intervention settings. Variability generally increased, particularly among girls, indicating increasingly heterogeneous performance despite relatively stable average scores. Overall, disparities between weight groups widened under every intervention scenario because declines were consistently greater among children with EW. None of the intervention approaches substantially altered these trajectories, suggesting limited effectiveness for preserving lower-body power among children with EW.

Changes in body composition were considerably smaller than those observed for physical fitness outcomes and showed limited differences between intervention settings. Among children with EW, BMI trajectories remained relatively stable over time, while triceps skinfold thickness generally declined, particularly under combined interventions and HSN-only schools. Variability in BMI showed only modest changes across intervention settings, whereas variability in triceps skinfold thickness generally decreased, indicating increasingly homogeneous adiposity within the EW group. Among children with NW, BMI changed little across follow-up, with only small intervention-related differences. Triceps skinfold thickness also demonstrated relatively modest changes, with variability generally decreasing under intervention conditions. Overall, intervention-related differences in body composition were considerably smaller than those observed for cardiorespiratory and muscular fitness. None of the intervention approaches was consistently associated with meaningful reductions in inequalities in BMI or triceps skinfold thickness, suggesting that long-term school-based healthy lifestyle policies had a greater influence on physical fitness than on body composition.

An illustrative study focusing on the temporal trends in the means and distributional characteristics of cardiorespiratory fitness for French children and adolescents between 1999 and 2022 suggested that the trends between children with low, moderate, and high cardiorespiratory fitness were not always uniform over time.²⁸ Although the results suggested that there has been a substantial decline since baseline, which slowed considerably after 2010, a larger decline was found in children with low cardiorespiratory fitness compared to those with high cardiorespiratory fitness, whereas post-2010, trends were similar across all children.²⁸ To gain

insights into the potential causes of decline in Brazilian children's physical fitness over the past 17 years, a recent study observed distributional changes in the data using the standard deviations as well as the means.²⁹ As presented in the study, the standard deviations of all six physical fitness tests were systematically increasing over time, i.e., despite a steady mean decline in five of the physical fitness tests, the distribution of the physical fitness scores was increasing/expanding.²⁹ The positively skewed distributions observed for five physical fitness measures, together with increasing variability over time, may indicate diverging fitness trajectories within the population, whereby a relatively small proportion of fitter children continue to improve while less-fit children fall further behind.²⁹ A previous study that examined data based on the implementation of a school-wide intervention, Alberta Project Promoting active Living and healthy Eating in Schools (APPLE Schools), in a socioeconomically disadvantaged neighbourhood, found reduced differences between within-group changes in equity, calculated as the sum of the relative difference in 2009 and 2011.¹⁴ These reductions in inequalities were shown between children living with normal and EW (– 2.7%), but also between active and least active, highly active and least active, highly active and active boys and girls (– 2%, – 2.9% and – 0.9%, respectively). When divided by sex, these reductions were not equally represented, and there were more notable negative outcomes for girls. The authors of the mentioned study concluded based on their knowledge that this was the first experimental study at the time that demonstrated improvements in physical activity and reduced health inequalities among students by implementing school-based health promotion, which more favourably improved the physical activity levels amongst the most disadvantaged subgroups.¹⁴ Most of these previously reported studies examined trends in distributional characteristics either visually, e.g., using LOWESS curves where trends were plotted across a range of percentiles or as trends in the variability, e.g., the ratio of standard deviations or coefficients of variation. Some of the analyses were either cross-sectional or did not track participants over several decades, some did not consider examining the comparative effectiveness of interventions, others did not include weight status as a factor and, to our knowledge, even if all of these reasons were considered, physical fitness was not the main outcome. Therefore, we found it difficult to draw relevant comparisons with our study, and we can discuss the worthwhile contribution of our study to the literature.

Explanation of findings

Across the study period, the most notable finding was that intervention effects differed more in their influence on variability than on mean performance. Combined interventions showed some evidence of reduced heterogeneity, although this pattern was not consistent across fitness domains. Children with EW generally showed increasingly heterogeneous developmental trajectories, particularly in HSN-only and non-intervention schools, whereas children with NW showed more stable or decreasing variability. Importantly, changes in variability were often not aligned with changes in mean performance, indicating that similar average trajectories may mask substantial differences in individual responses.

One possible explanation is that school-based interventions are generally designed for the whole student population rather than specifically addressing more vulnerable subgroups. As children with EW represent a minority within classrooms, standardised programmes may insufficiently accommodate their physical, motivational, or psychosocial needs. Consequently, some children with EW may benefit while others respond less successfully, resulting in greater variability. The lower heterogeneity observed in schools implementing both HLS and HSN may suggest that combining health literacy with additional opportunities for physical activity provides a more supportive environment that better accommodates individual differences. Differentiated instruction and

activities adapted to different levels of physical competence may promote more inclusive participation and reduce unequal intervention responses, particularly among children with EW.

Strengths and limitations

We should note several strengths of our study. Physical fitness was assessed annually using the SLOfit system, with measurements conducted by trained physical education teachers and data quality checked by competent experts. The study included more than 135,000 children and adolescents annually from all regions of Slovenia and allowed analyses by age, sex, and weight status. Importantly, we evaluated nationwide healthy lifestyle policies under real-world school conditions rather than a tightly controlled experimental setting. Although this introduces greater heterogeneity and potential confounding, it increases external validity and applicability to routine school practice and public health policy. The long-term implementation of both programmes also reduces the likelihood that findings primarily reflect novelty or initial enthusiasm effects. Finally, by examining variability alongside mean trends, we addressed an aspect of intervention effects that has received limited attention in physical fitness research.

Moreover, we want to consider several limitations. We did not account for changes in physical activity, diet, sleep, sedentary behaviour, pubertal status, socioeconomic status, or differences between fit and unfit children. We also lacked information on the duration of individual exposure to HLS, as schools joined the programme at different times and children could enter the programme later, which may partly explain heterogeneous intervention effects. The lower residual variance observed among boys with EW, particularly for adiposity-related outcomes, should also be interpreted cautiously, as it may partly reflect the smaller size of the EW subgroup, regularisation of variance estimates, and range restriction resulting from weight-status classification. The study period ended in 2019, before COVID-19 restrictions were introduced in Slovenia in March 2020; therefore, the observed trends were not affected by pandemic-related disruptions. HLS also ended in 2018/19 and was replaced by a less effective extracurricular programme from a physical fitness perspective, preventing comparable analyses beyond this period. Finally, comparisons with other studies should account for differences in the fitness tests used across countries and between sexes.³⁰

Intervention settings differed more in their influence on variability than on mean performance, with heterogeneity often increasing among children with EW, particularly in HSN-only and non-intervention schools. Although combined interventions generally produced more balanced fitness trajectories, they did not consistently reduce variability. Children with EW entered the study with lower fitness and therefore had greater potential for improvement, yet generally showed smaller gains than children with NW, indicating that ceiling effects alone cannot explain the observed inequalities. School-based interventions may therefore require greater adaptation to the needs of children with EW to avoid reinforcing existing disparities. Future interventions and evaluations should explicitly consider variability in individual responses alongside mean effects to better assess and promote equitable improvements in physical fitness.

Conclusion

We examined whether improvements in physical fitness were accompanied by more equitable outcomes among children in real-world school settings. By considering both mean performance and variability, we assessed whether children moved forward together or whether disparities within groups widened over time. Improvements in fitness did not necessarily translate into more equal outcomes. In several domains, particularly

cardiorespiratory fitness and body core strength, variability increased among children with EW despite improvements in mean performance, whereas children with NW generally showed more stable or decreasing variability. Thus, interventions did not consistently reduce disparities within groups and, in some cases, differences between children became more pronounced.

These findings highlight the need to move beyond one-size-fits-all approaches and consider whether interventions benefit children with different weight statuses equitably. Policies should prioritise inclusive strategies that support children with EW and reduce unequal responses to interventions. Although our comparisons did not identify a consistently superior intervention for promoting equity, they highlight the importance of considering both mean effects and variability when evaluating school-based physical fitness interventions.

Abbreviations

BMI body mass index

TSF triceps skinfold thickness

R600 600-m run

SU60 60 sec sit-ups

BAH bent arm hang

SBJ standing broad jump

WHO World Health Organization

EW excess weight

NW normal weight

HLS Healthy lifestyle intervention

HSN Healthy schools network intervention

Declarations

Ethics approval and consent to participate

SLOfit maintains strict ethical standards focused on informed consent, data privacy, and voluntary participation. While schools mandate physical development monitoring, central database inclusion requires explicit parental or student consent, and participants retain full rights to access, manage, or request anonymisation of their personal fitness data at any time. All the data in the SLOfit database is automatically anonymised one year after an individual finishes their education. All the users are notified before anonymisation and are given an option to permanently store and access the data in the application. Anonymised SLOfit data are afterwards used for statistical and scientific research in accordance with the agreement of the Commission on Ethical Issues in Sport No. 6-2019-539.

Consent for publication

Not applicable.

Availability of data and materials

Data used in this study are available from the first author upon reasonable request. We will upload and make the code we used in this paper openly available on GitHub (<https://github.com/antonio-martinko>) upon publication.

Competing interests

No conflicts of interest are declared.

Funding

This work was funded by the European Commission's Horizon 2020 research and innovation programme (Grant no. 774548) within the Science and Technology in childhood Obesity Policy (STOP) project and by the Slovenian National Research and Innovation Agency within Bio-psycho-social research programme, no. P5-0142. The work of A.M. was funded by the Croatian Science Foundation, under grant no: DOK-2020-01-3728.

Author Contribution

With the intention to recognise the individual contributions of authors, the CRediT (Contributor Roles Taxonomy) framework is followed. Antonio Martinko: Conceptualization, Methodology, Software, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization; Gregor Jurak: Conceptualization, Writing - Review & Editing; Gregor Starc: Conceptualization, Writing - Review & Editing; Maroje Sorić: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration.

Acknowledgement

The authors wish to thank the schools, teachers, children, and their parents who generously agreed to be involved in the SLOfit system and the Ministry of Education, Science, and Youth of the Republic of Slovenia for funding the annual data analysis of the SLOfit measurements.

Data Availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. Lobstein T, Gabrijelčič M. A policy assessment tool to identify causes of inequities that influence obesity prevalence. *Bull World Health Organ.* 2025;103:57–65.

2. Johansson L, Putri RR, Danielsson P, Hagströmer M, Marcus C. Associations between cardiorespiratory fitness and cardiometabolic risk factors in children and adolescents with obesity. *Sci Rep.* 2023;13:7289.
3. Lund MAV, Thstrup AH, Frithioff-Bøjsøe C, et al. Low-grade inflammation independently associates with cardiometabolic risk in children with overweight/obesity. *Nutr Metab Cardiovasc Dis.* 2020;30:1544–53.
4. Skinner AC, Perrin EM, Moss LA, Skelton JA. Cardiometabolic Risks and Severity of Obesity in Children and Young Adults. *N Engl J Med.* 2015;373:1307–17.
5. Chung ST, Onuzuruike AU, Magge SN. Cardiometabolic risk in obese children. *Ann N Y Acad Sci.* 2018;1411:166–83.
6. Bagatini NC, Feil Pinho CD, Leites GT, da Cunha Voser R, Gaya AR, Santos Cunha GD. Effects of cardiorespiratory fitness and body mass index on cardiometabolic risk factors in schoolchildren. *BMC Pediatr.* 2023;23:454.
7. Mintjens S, Menting MD, Daams JG, van Poppel MNM, Roseboom TJ, Gemke RBBJ. Cardiorespiratory Fitness in Childhood and Adolescence Affects Future Cardiovascular Risk Factors: A Systematic Review of Longitudinal Studies. *Sports Med.* 2018;48:2577–605.
8. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes.* 2008;32:1–11.
9. García-Hermoso A, Ramírez-Campillo R, Izquierdo M. Is Muscular Fitness Associated with Future Health Benefits in Children and Adolescents? A Systematic Review and Meta-Analysis of Longitudinal Studies. *Sports Med.* 2019;49:1079–94.
10. Sun C, Pezic A, Tikellis G, et al. Effects of school-based interventions for direct delivery of physical activity on fitness and cardiometabolic markers in children and adolescents: a systematic review of randomized controlled trials. *Obes Rev.* 2013;14:818–38.
11. Van Sluijs EMF, McMinn AM, Griffin SJ. Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. *BMJ.* 2007;335:703.
12. Wu J, Yang Y, Yu H, Li L, Chen Y, Sun Y. Comparative effectiveness of school-based exercise interventions on physical fitness in children and adolescents: a systematic review and network meta-analysis. *Front Public Health.* 2023;11:1194779.
13. Rodrigo-Sanjoaquin J, Corral-Abós A, Aibar Solana A, Zaragoza Casterad J, Lhuisset L, Bois JE. Effectiveness of school-based interventions targeting physical activity and sedentary time among children: a systematic review and meta-analysis of accelerometer-assessed controlled trials. *Public Health.* 2022;213:147–56.
14. Vander Ploeg KA, Maximova K, McGavock J, Davis W, Veugelers P. Do school-based physical activity interventions increase or reduce inequalities in health? *Soc Sci Med.* 2014;112:80–7.
15. Gorely T, Nevill ME, Morris JG, Stensel DJ, Nevill A. 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.
16. Jurić P, Jurak G, Morrison SA, Starc G, Sorić M. Effectiveness of a population-scaled, school-based physical activity intervention for the prevention of childhood obesity. *Obesity.* 2023;31:811–22.
17. Lang JJ, Zhang K, Agostinis-Sobrinho C, et al. Top 10 International Priorities for Physical Fitness Research and Surveillance Among Children and Adolescents: A Twin-Panel Delphi Study. *Sports Med.* 2023;53:549–64.
18. The UN Sustainable Development Goals. 2015. <https://sdgs.un.org/2030agenda>
19. Jurak G, Leskošek B, Kovač M. SLOfit surveillance system of somatic and motor development of children and adolescents: Upgrading the Slovenian Sports Educational Chart. *Acta Univ Carol Kin.* 2020;56:28–40.

20. International standard classification of education. ISCED 2011. Quebec: UNESCO Institute for Statistics,; Montreal; 2012.
21. Blagus R, Jurak G, Starc G, Leskošek B. Centile Reference Curves of the SLOfit Physical Fitness Tests for School-Aged Children and Adolescents. *J Strength Conditioning Res.* 2023;37:328–36.
22. Jurak G, Leskošek B, Kovač M, et al. SLOfit surveillance system of somatic and motor development of children and adolescents: Upgrading the Slovenian Sports Educational Chart. *AUC KINANTHROPOLOGICA.* 2020;56:28–40.
23. CDDS. TESTING PHYSICAL FITNESS: EUROFIT. Experimental Battery - PROVISIONAL HANDBOOK. Strasbourg: Council of Europe. 1983 https://bitworks-engineering.co.uk/Products_files/eurofit%20provisional%20handbook%20leger%20beep%20test%201983.pdf
24. Radulović A, Jurak G, Leskošek B, Starc G, Blagus R. Secular trends in physical fitness of Slovenian boys and girls aged 7 to 15 years from 1989 to 2019: a population-based study. *Sci Rep.* 2022;12:10495.
25. Ortega FB, Leskošek B, Blagus R, et al. European fitness landscape for children and adolescents: updated reference values, fitness maps and country rankings based on nearly 8 million test results from 34 countries gathered by the FitBack network. *Br J Sports Med.* 2023;57:299–310.
26. Chung S. Growth and Puberty in Obese Children and Implications of Body Composition. *Journal Obesity Metabolic Syndrome.* 2017;26:243–50.
27. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2023. <https://www.R-project.org/>.
28. Vanhelst J, Lang JJ, Matelot D, et al. Cardiorespiratory fitness has declined among French children since 1999, although the decline appears to be getting smaller. *Scand J Med Sci Sports.* 2024;34:e14641.
29. Nevill AM, Duncan MJ, Gaya A, Mello JB. Secular trends in the physical fitness of Brazilian youth: Evidence that fitness is declining for the majority but not for a fit minority. *Scand J Med Sci Sports* 2023; published online July 4. 10.1111/sms.14440
30. Masanovic B, Gardasevic J, Marques A, et al. Trends in Physical Fitness Among School-Aged Children and Adolescents: A Systematic Review. *Front Ped.* 2020;8. 10.3389/fped.2020.627529.

Figures

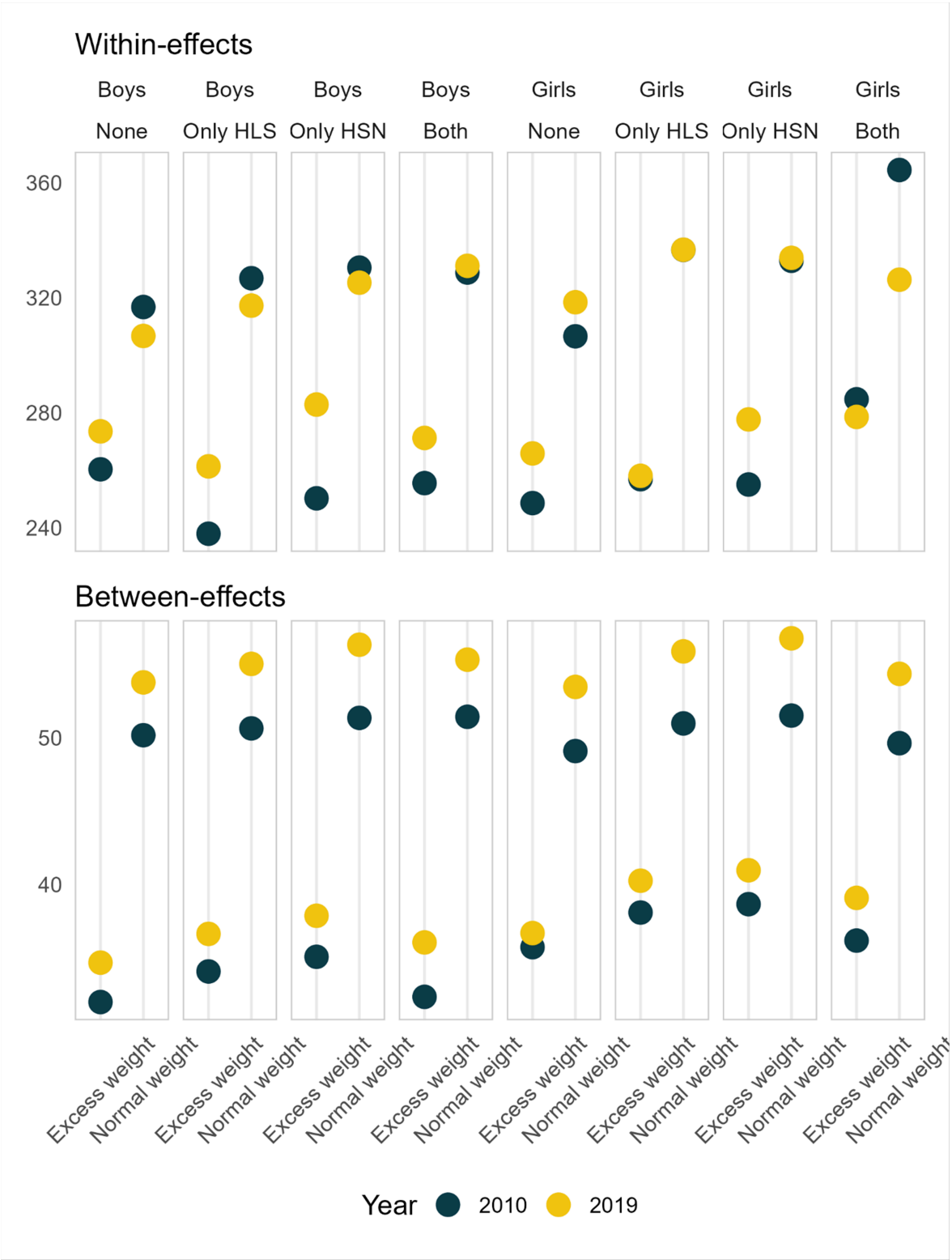


Figure 1

Within-group variability and between-group differences in cardiorespiratory fitness by intervention, weight status, and sex, 2010–2019

Upper panel: within-group variability (residual variance). Lower panel: between-group differences in model-predicted centile values. Results are shown by intervention, weight status (NW and EW), and sex at baseline and final follow-up.

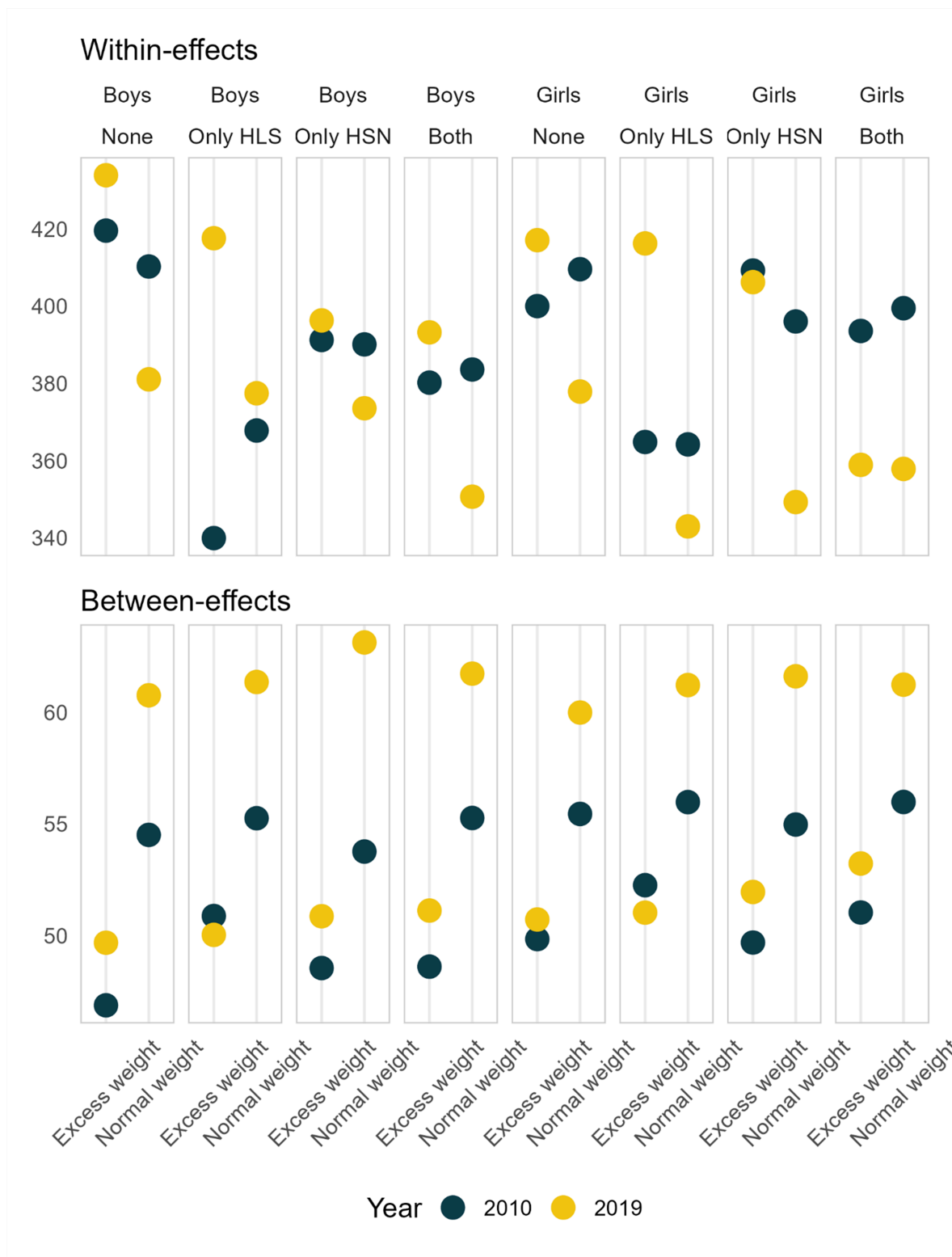


Figure 2

Within-group variability and between-group differences in body core strength by intervention, weight status, and sex, 2010–2019

Upper panel: within-group variability (residual variance). Lower panel: between-group differences in model-predicted centile values. Results are shown by intervention, weight status (NW and EW), and sex at baseline and final follow-up.

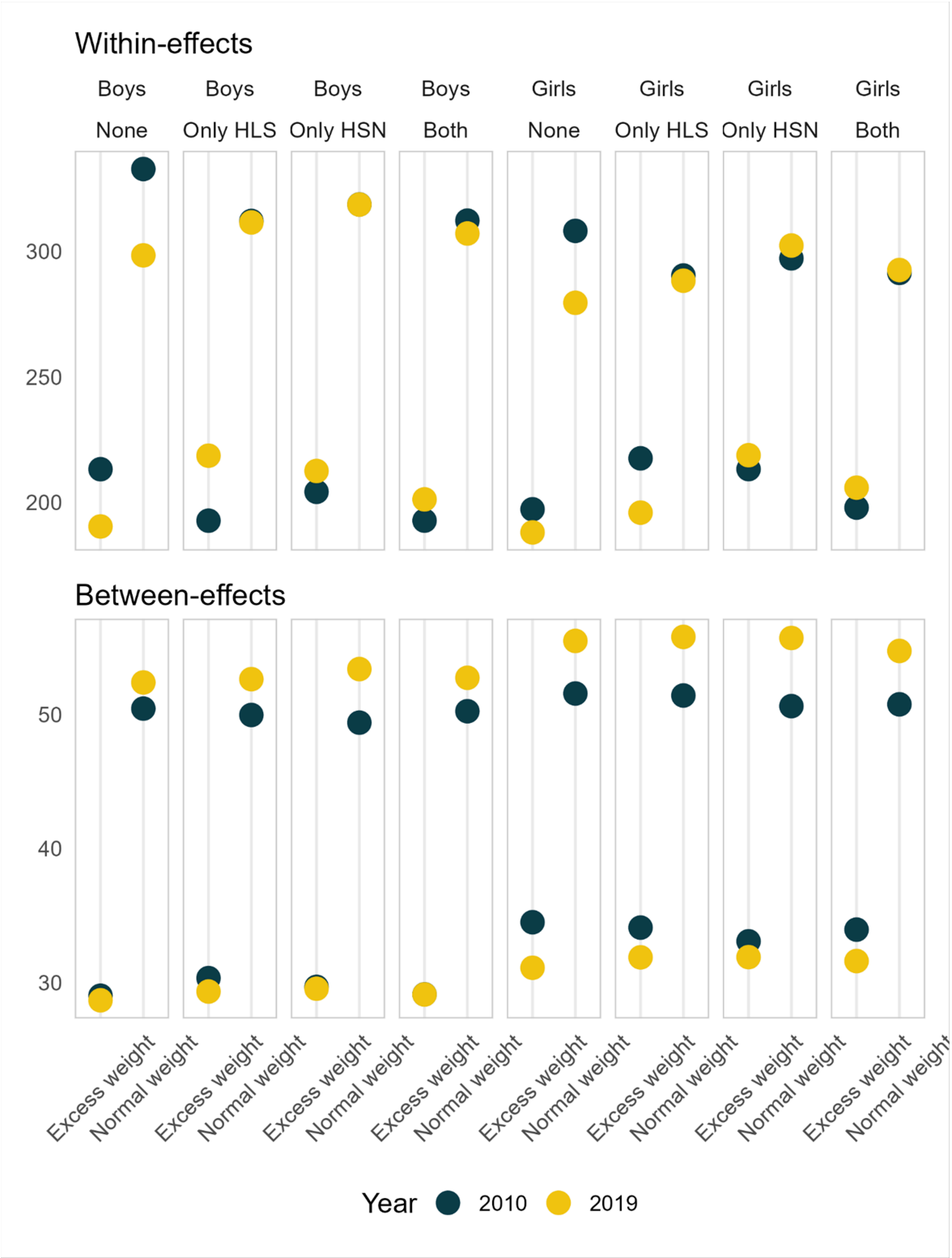


Figure 3

Within-group variability and between-group differences in upper body strength by intervention, weight status, and sex, 2010–2019

Upper panel: within-group variability (residual variance). Lower panel: between-group differences in model-predicted centile values. Results are shown by intervention, weight status (NW and EW), and sex at baseline and final follow-up.

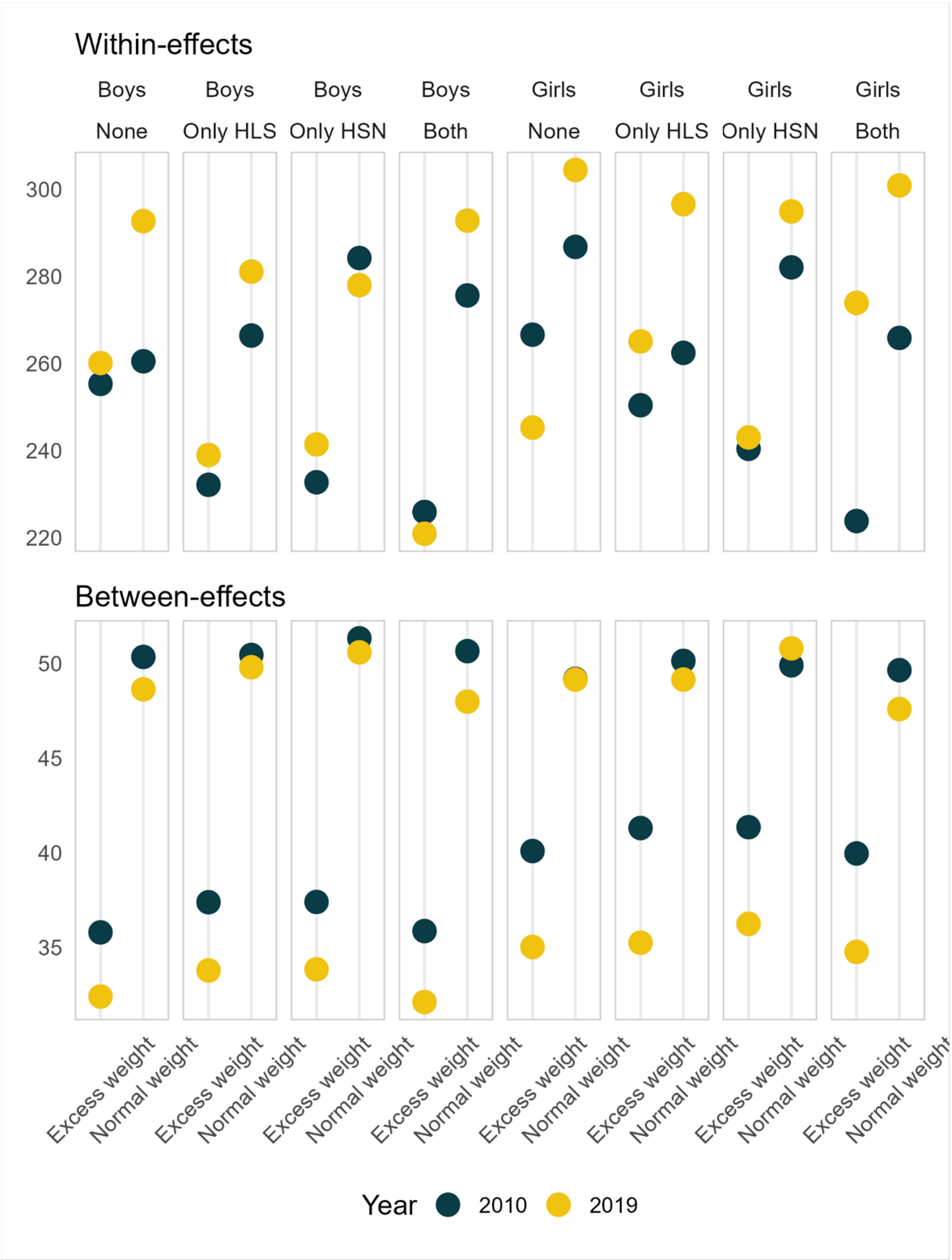


Figure 4

Within-group variability and between-group differences in lower body power by intervention, weight status, and sex, 2010–2019

Upper panel: within-group variability (residual variance). Lower panel: between-group differences in model-predicted centile values. Results are shown by intervention, weight status (NW and EW), and sex at baseline and final follow-up.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1MartinkoFitnessTrendsOfChildren.docx](#)
- [Additionalfile2MartinkoFitnessTrendsOfChildren.docx](#)
- [Additionalfile3MartinkoFitnessTrendsOfChildren.docx](#)