Obesity THE OBESITY SOCIETY **WILEY**

Effectiveness of a population-scaled, school-based physical activity intervention for the prevention of childhood obesity

Petra Jurić[1] | Gregor Jurak[2] | Shawnda A. Morrison[2] | Gregor Starc[2] |
Maroje Sorić[1,2]

[1]Faculty of Kinesiology, University of Zagreb, Zagreb, Croatia

[2]Faculty of Sport, University of Ljubljana, Ljubljana, Slovenia

Correspondence
Petra Jurić, Faculty of Kinesiology, University of Zagreb, Horvaćanski zavoj 15, 10000 Zagreb, Croatia.
Email: petra.juric@kif.unizg.hr

Abstract

Objective: The aim of this study was to examine the effectiveness of a real-world, population-scaled, school-based physical activity (PA) intervention that provided two to three additional physical education lessons per week to children aged 6 to 14 years in Slovenia.

Methods: More than 34,000 participants from over 200 schools were compared with a similar number of nonparticipants from the same schools. Generalized estimating equations were used to estimate the effects of differing levels of exposure to the intervention (i.e., from 1–5 years) on BMI in children with normal weight, overweight, or obesity at baseline.

Results: BMI was lower in the intervention group, irrespective of participation duration or baseline weight status. The difference in BMI increased with the program duration, with maximal effects being seen after 3 to 4 years of participation, and was consistently larger for children with obesity (peaking at 1.4 kg/m^2 [95% CI: 1.0–1.9] for girls with obesity and peaking at 0.9 kg/m^2 [95% CI: 0.6–1.3] for boys with obesity). The program started to become effective at reversing obesity after 3 years, whereas the lowest numbers needed to treat (NNTs) were observed after 5 years (NNTs = 17 for girls and 12 for boys).

Conclusions: The population-scaled, school-based PA intervention was effective in preventing and treating obesity. The effects were the greatest in children initially presenting with obesity, such that the program was able to benefit children needing support the most.

INTRODUCTION

The global prevalence of childhood obesity has increased at a startling rate, moving from 0.7% to 5.6% in boys and 0.9% to 7.8% in girls from 1975 to 2016 [1]. Projected obesity-related morbidity incidence rates and all-cause mortality are very high, as are the anticipated costs for

health care and economic losses [2]. It is, therefore, critical to implement convenient and controlled approaches on a global scale in order to slow down and ultimately reverse this costly pandemic.

Most children and adolescents go to school daily, which means that school-based interventions can impact many children simultaneously, including hard-to-reach groups. Numerous systematic reviews have shown that school interventions involving physical activity (PA) produce larger effects on body weight than interventions

Gregor Starc and Maroje Sorić had equal authorship.

without PA [3–5]. Therefore, a school-based program that includes PA can serve as an important contributing feature in childhood obesity management, especially when led by experienced physical education (PE) professionals [6]. Moreover, fitness-oriented interventions that focus on improving the components of physical fitness and typically involve more vigorous PA (e.g., structured exercise) have demonstrated greater potential for improving body composition than other types of PA interventions, such as interventions focusing on merely increasing the volume of PA (e.g., providing more time for unstructured play) or interventions with exclusively educational content (e.g., providing information on the benefits of PA) [7].

Although the beneficial health effects of initiating PA intervention programs during childhood are well documented, most of this evidence comes from short-term efficacy trials conducted in well-controlled settings, usually without implementing large-scale, or scalable, population-based approaches. The current lack of successfully implemented school-based PA interventions in real-world settings impedes the fight against the childhood obesity pandemic [8, 9]. The present study leverages a natural experiment in Slovenia that provided the opportunity to examine the effectiveness of the Healthy Lifestyle intervention, a real-world, population-based, longitudinal PA intervention, on body mass index (BMI) in children aged 6 to 14 years; this intervention was derived from a previous successful, small-scale PA intervention in individual Slovenian schools that provided an above-standard program of one PE lesson per day, which was also delivered by PE specialist teachers in the lower classes of primary school [10].

METHODS

Intervention

Healthy Lifestyle was a nationwide intervention introduced in Slovenia from 2011 to 2018. The intervention provided two additional PE lessons in grades 1 to 6 and three in grades 7 to 9, thus providing one PE lesson per day to children aged 6 to 14 years (Table 1). The additional lessons were not part of the obligatory curricula but they were organized to take place immediately after the end of regular school hours and were thus within the time frame of the ordinary school day. The intervention was financed by the European Social Fund with the aim of increasing initial employment opportunities for recently graduated PE teachers. In order to get the funding, schools needed to employ recently graduated PE teachers who were the only teachers delivering the intervention lessons. The intervention was offered to all children in an individual school and was organized in the form of an elective course. After the children joined the program upon their parental written consent, their participation became compulsory, but they were not graded as in regular PE classes. The maximum number of children per class was between 16 and 30, and the school was allowed to shape multiple joint classes composed of children from grades 1 to 3, 4 to 6, or 7 to 9 if the number of enrolled children per grade was lower than 16. The program required teachers to provide at

Study Importance

What is already known?

- The global prevalence of obesity has increased at a significant rate, growing from 0.7% to 5.6% in boys and 0.9% to 7.8% in girls from 1975 to 2016.
- The current lack of successfully implemented interventions in real-world settings prevents quality management of the childhood obesity pandemic.

What does this study add?

- A school-based physical activity (PA) intervention that provided additional physical education lessons remained effective in the prevention of obesity after scaling up.
- The greatest effect was present in children initially presenting with obesity, such that the program was able to benefit children needing support the most, whereas the number needed to treat for obesity reversal decreased with intervention duration, emphasizing the need for long-term PA programs.

How might these results change the direction of research or the focus of clinical practice?

- To be effective at reducing obesity for children of both sexes aged 6 to 14 years, the intervention must last a minimum of three consecutive years without funding interruption because even temporary disruptions in long-lasting interventions attenuate their long-term effectiveness.
- Policy makers and funding bodies should be aware that obesity is a chronic condition that needs to be dealt with over a longer time frame and that easy solutions and immediate effects are neither realistic nor sustainable.

least 12 different sports per triennium, and they had to prioritize the three most established sports in the local environment. It also included the presentation of urban sports (such as in-line skating, parkour, and other sports suitable in urban settings) that were currently not specifically covered in the PE curricula at the time, and PE specialist teachers also had to provide limited information on healthy dietary and lifestyle habits regarding energy balance, limiting the consumption of snacks and sugar-sweetened beverages, and promoting a diverse diet. Teachers were free to choose how to provide this information, but they typically delivered it in the form of short group conversations at the beginning of the lessons or during short breaks between activities. Childhood obesity was not a specific target of the intervention per se. However, increasing overall PA with the addition of two or three PE lessons was considered a helpful "by-product" strategy for

TABLE 1 Participation in Healthy Lifestyle intervention and related costs in the period of 2010–2018

School year	2010/2011	2011/2012	2012/2013	2013/2014	2014/2015	2015/2016	2016/2017	2017/2018
Participants (N)	18,993	24,202	26,000	27,600	30,261	29,549	35,640	32,245
Lessons (N)	33,190	60,505	68,306	70,866	72,054	53,527	69,613	51,893
Annual costs (€)	1,156,322	1,754,087	2,007,291	2,026,940	2,070,681	1,752,964	2,618,384	2,341,557
Annual costs per child (€)	60.88	72.48	77.20	73.44	68.43	59.32	73.47	72.62

maintaining optimal weight and improving the energy imbalance in participating children. All in all, the intervention was designed through a bottom-up approach, meaning that schools were totally independent in selecting the contents and form of work. This approach was chosen to facilitate adaptation to local contexts and settings. Parents only provided consent for their children to be involved in the intervention; they did not receive any educational materials and were not involved in the intervention program in any other way.

New schools were joining the program yearly, so in the final year of implementation, the total number of involved schools was 216 (48% of the total number of primary schools). Only two schools decided to discontinue the intervention in the entire 2011 to 2018 period. On the other hand, the intervention faced a serious challenge in the school year 2015/2016 when financing was suspended for several months for administrative reasons. This resulted in a considerable reduction of delivered lessons compared with previous years (Table 1).

Study design and sample

The protocol, measurement procedures, and data management of the SLOfit surveillance system were approved by the National Medical Ethics Committee of the Republic of Slovenia (number 52/03/14), and they are in accordance with the Declaration of Helsinki. The Healthy Lifestyle intervention did not require ethical approval because it was not an experimental study and was independently evaluated by using the SLOfit system. This intervention was implemented at a national level on a voluntary basis, and we treat this as a natural experiment; here we compare weight outcomes in children involved in the intervention with the outcomes of their peers who did not participate. In Slovenia, there are 451 primary schools, 216 of which volunteered to implement the intervention for at least 1 school year. Each year, the Slovenian Sports Office published a public call, and schools were free to submit their applications. All the schools that applied in any of the years were granted the funding, and no school was ever refused the funding. Participating schools did not differ from nonparticipating schools in terms of regional distribution, size, or urbanization level, but participating schools did show higher levels of baseline obesity (data not shown). Between 18,000 and 35,000 children were included in the intervention in each single year, and around 96% of the children from participating schools were measured within the SLOfit system every year. Because of the natural experiment design, intervention assignments on both the school and individual levels were outside of our control (with schools joining in different years and children joining, leaving, and rejoining the intervention at different ages), which resulted in a large and unfeasible number of durations to consider when evaluating the intervention effects. Hence, we opted to include only children who participated in the intervention continuously over a certain period, but we restricted the analyses to children who were enrolled in a participating school at least a year before the specific school joined the intervention to have true baseline values. We compared these children with a control group of children who attended participating schools but who were not involved in the intervention at any time. All analyses were stratified by the number of consecutive years of participation or nonparticipation in the intervention. The longer the participation duration, the fewer generations of children there were who had the opportunity to join the intervention. This resulted in very low statistical power in models that were restricted to groups with 6, 7, and 8 years of participation. Hence, we decided to restrict the analyses to a consecutive 5 years of participation.

Anthropometric measurements

Height, weight, and triceps skinfold measurements were obtained through the SLOfit system—the Slovenian national fitness surveillance system—in accordance with the standardized and uniform protocol [11], and data were anonymized regarding the intervention involvement. The SLOfit measurements are organized in all Slovenian schools every April, assuring an identical time interval between measurements in all schools with standard equipment [12]. The measurements in schools are performed by the regular PE teachers with the support of classroom teachers, and all of the schools are equipped with standard measuring equipment. All PE teachers in Slovenia are educated in a single 5-year study program at the Faculty of Sport, University of Ljubljana, and are all thoroughly trained in measurement procedures in three different study courses.

The SLOfit systematic measurement protocol requires children to be tested barefoot and wearing only light clothes during the anthropometric measurements. Height is measured to the nearest millimeter in the standing position with a stadiometer, and weight is measured to the nearest 0.1 kg with a medical scale. The measurements are sent to the Laboratory for Diagnostics of Somatic and Motor Development at the Faculty of Sport, where the data are checked for logistical errors; any errors are communicated to teachers for immediate
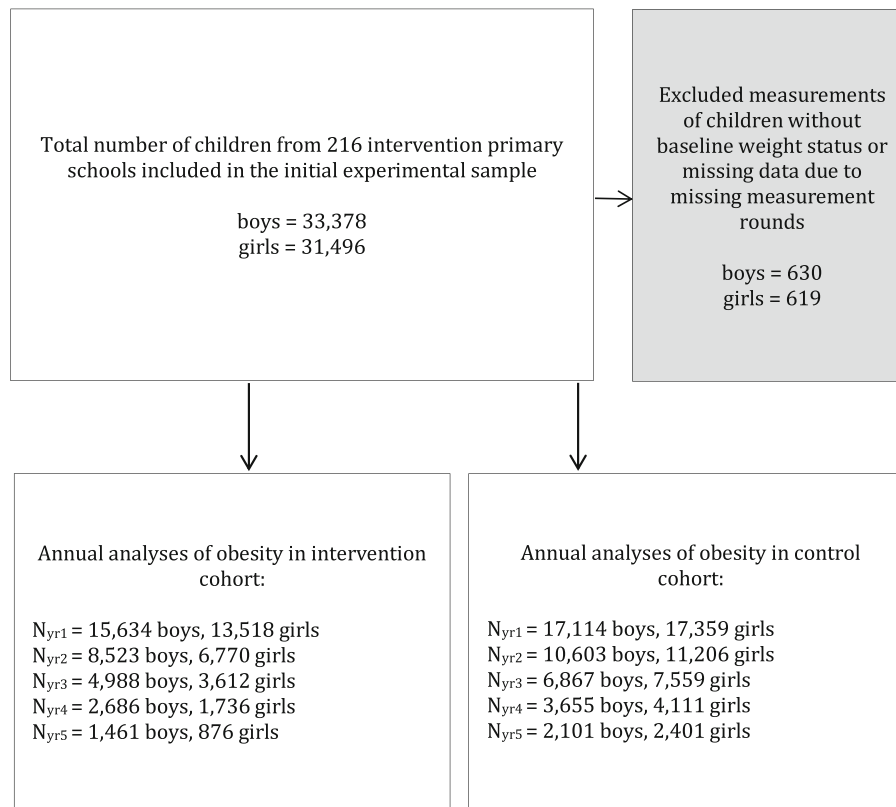
Total number of children from 216 intervention primary schools included in the initial experimental sample

boys = 33,378
girls = 31,496

Excluded measurements of children without baseline weight status or missing data due to missing measurement rounds

boys = 630
girls = 619

Annual analyses of obesity in intervention cohort:

N_{yr1} = 15,634 boys, 13,518 girls
N_{yr2} = 8,523 boys, 6,770 girls
N_{yr3} = 4,988 boys, 3,612 girls
N_{yr4} = 2,686 boys, 1,736 girls
N_{yr5} = 1,461 boys, 876 girls

Annual analyses of obesity in control cohort:

N_{yr1} = 17,114 boys, 17,359 girls
N_{yr2} = 10,603 boys, 11,206 girls
N_{yr3} = 6,867 boys, 7,559 girls
N_{yr4} = 3,655 boys, 4,111 girls
N_{yr5} = 2,101 boys, 2,401 girls

FIGURE 1 Study design and flowchart of the study sample. yr, year

correction. The participation rate of children in SLOfit for the period studied here (i.e., 2010–2018) surpassed 94% in all years.

Statistical methods

Because the aim of this study was to examine the effects of an intervention scaled up to a population level, we opted to use generalized estimating equations (GEE) [13]—one of the population-average models (or marginal models) that tests the average effects on a population level, instead of examining individual-level effects. GEE models [14] also provide robust parameter estimates regardless of the assumed variance–covariance correlation matrix and deal well with missing data. Unlike traditional basic regression models, a GEE model can handle multilevel, clustered, and autocorrelated data and it does not require distributional assumptions (e.g., normally distributed data). We accounted for the multilevel structure of the data by considering clustering at the school level, but we did not cluster at the exercise-group level, as this information was not available. The change in BMI was analyzed by using a linear scale response. Because of the balanced data, we specified a first-order autoregressive correlation structure for all GEE models [14]. Time—as a within-subject variable—was categorized into five categories, contrasting baseline with the first, second, third, fourth, and fifth year of children's participation or nonparticipation in the intervention. Covariates were selected a priori based on expert knowledge, and each model was adjusted for age, baseline school obesity prevalence, economic

affluence of the local environment (Municipality Development Index), individual risk of obesity (baseline percentile of triceps skinfold thickness of an individual), proxy of individual maturation rate (body height percentile rank of an individual in a certain year), and intervention disruption (designation of whether an individual was exposed to disturbance of the intervention in 2016). Because we tried to establish the possible differences in the effects of the intervention on BMI in children who had normal weight, overweight, or obesity at baseline, stratified by sex, we produced 30 different models with all potentially moderating covariates included. The criterion for normal weight was having BMI below the 85th percentile of the national age- and sex-specific BMI values, which were calculated by using the data of more than 7.5 million measurements in the period of 1989 to 2020. The criterion for overweight was having BMI above or equal to the 85th percentile but below the 95th percentile, and the criterion for obesity was having BMI equal to or above the 95th percentile. As the GEE model does not contain an intuitive statistical effect size metric, we report and discuss the effect size in terms of clinical significance of observed effects (i.e., β coefficients) and related uncertainty estimates (i.e., 95% confidence intervals [CIs]). These effects denote the differences in BMI between the intervention group and control group at a given time point.

The number of obesity cases reversed was calculated as the difference between the number of obesity cases at the baseline year and the number of obesity cases at the final year for all five participation durations. The χ^2 test was used to assess the difference in the number of obesity cases between the baseline and final year in each duration

scenario. The Cramer V value was calculated as a measure of effect size. The number needed to treat (NNT; number of children need to be involved in the intervention to reverse one additional obesity case) was calculated by using a modification to the standard equation. First, the difference in the favorable clinical outcome rate between the experimental group and the control group was calculated (i.e., the difference between obesity reversal rates). In the next step, the NNT was calculated as the inverse of this number and rounded up to the higher whole number [15].

An independent-sample t test was used to check baseline differences in age, triceps skinfold measurements, height, BMI, and school baseline obesity prevalence between the intervention group and the control group. All statistical analyses were stratified by sex and were performed by using SPSS Statistics 26.0 (IBM Corp.), and statistical significance was set at $\alpha = 0.05$.

RESULTS

There are 451 primary schools in Slovenia, out of which 216 (48%) were part of Healthy Lifestyle during the final year of the intervention. At baseline, the intervention cohort included 29,152 children, and the control cohort included 34,473 children. The number of participants in both the intervention cohort and control cohort dropped continuously as the duration of the intervention increased and amounted to 2337 (8%) and 4502 (13%) participants, respectively, in a 5-year participation scenario (Figure 1). This phenomenon was largely due to children finishing primary education and was in part due to children leaving the intervention. The differences between children who adhered to the intervention and their peers who dropped out at any point are shown in Supporting Information Tables S1 and S2. Children who dropped out were about 1 year older and they came from schools with slightly lower levels of baseline obesity.

The characteristics of participants at baseline according to consecutive years of participation in the intervention are shown in Table 2. Age differences were apparent between the intervention and control groups in all five analyzed participation durations, with the age difference declining between groups with longer participation duration. The intervention group had a higher height percentile, BMI percentile, and prevalence of obesity in all five participation durations (Table 2).

The unadjusted prevalence of obesity that shows the obesity trends in intervention schools compared with the trends in all other Slovenian schools that were never included in the intervention is presented in Figure 2. The schools that decided to join the intervention typically had a higher-than-average prevalence of obesity (except for schools joining in 2014 and 2016). The prevalence of obesity declined in the years after joining the intervention in schools that joined the intervention before or after 2016 but not in schools that joined in 2016 when the intervention was disrupted because of delayed and reduced financing. Next, in schools that joined the intervention from 2011 to 2015, a temporary increase of obesity prevalence in 2016 was seen, ranging from 0.2 to 0.3 percentage points. Last, despite

TABLE 2 Characteristics of participants at baseline according to consecutive years of participation in the intervention

Duration	Age (y)		Height percentile (%)		BMI percentile (%)		Triceps skinfold percentile (%)		School-level obesity (%)	
	Intervention	Control	Intervention	Control	Intervention	Control	Intervention	Control	Intervention	Control
Baseline	9.06 (2.25)*	10.37 (2.26)*	52.32 (28.74)	52.43 (28.79)	52.94 (29.30)*	52.41 (29.72)*	55.17 (28.63)*	54.66 (28.85)*	7.65 (3.11)**	7.12 (2.80)**
1 Year	10.06 (2.25)**	11.37 (2.26)**	53.93 (28.64)*	53.30 (28.78)*	53.05 (29.46)*	52.32 (29.78)*	55.47 (28.39)	55.19 (28.66)	7.65 (3.11)**	7.12 (2.80)**
2 Years	10.55 (1.96)**	11.98 (1.96)**	54.90 (28.41)**	53.81 (28.69)**	53.72 (29.43)**	52.24 (29.73)**	55.74 (28.81)**	54.45 (28.91)**	7.99 (3.28)**	7.04 (2.75)**
3 Years	11.19 (1.70)**	12.42 (1.70)**	55.56 (28.38)**	53.72 (28.86)**	53.46 (29.44)**	51.68 (29.77)**	55.64 (28.61)*	54.61 (28.95)*	8.22 (3.35)**	7.03 (2.76)**
4 Years	11.86 (1.46)**	12.82 (1.44)**	55.86 (28.26)**	53.38 (28.79)**	52.49 (29.74)*	50.66 (29.78)*	54.94 (28.94)	54.31 (28.87)	8.42 (3.47)**	7.01 (2.78)**
5 Years	12.60 (1.20)**	13.19 (1.20)**	56.06 (28.10)**	53.07 (28.73)**	53.23 (29.50)**	50.16 (29.85)**	55.78 (29.34)*	53.88 (28.66)*	8.49 (3.56)**	7.03 (2.75)**

Note: Data presented as mean (SD). Differences between groups were tested by using an independent-sample t test.

*Significant difference between intervention group and control group, $p < 0.05$.

**Significant difference between intervention group and control group, $p < 0.005$.
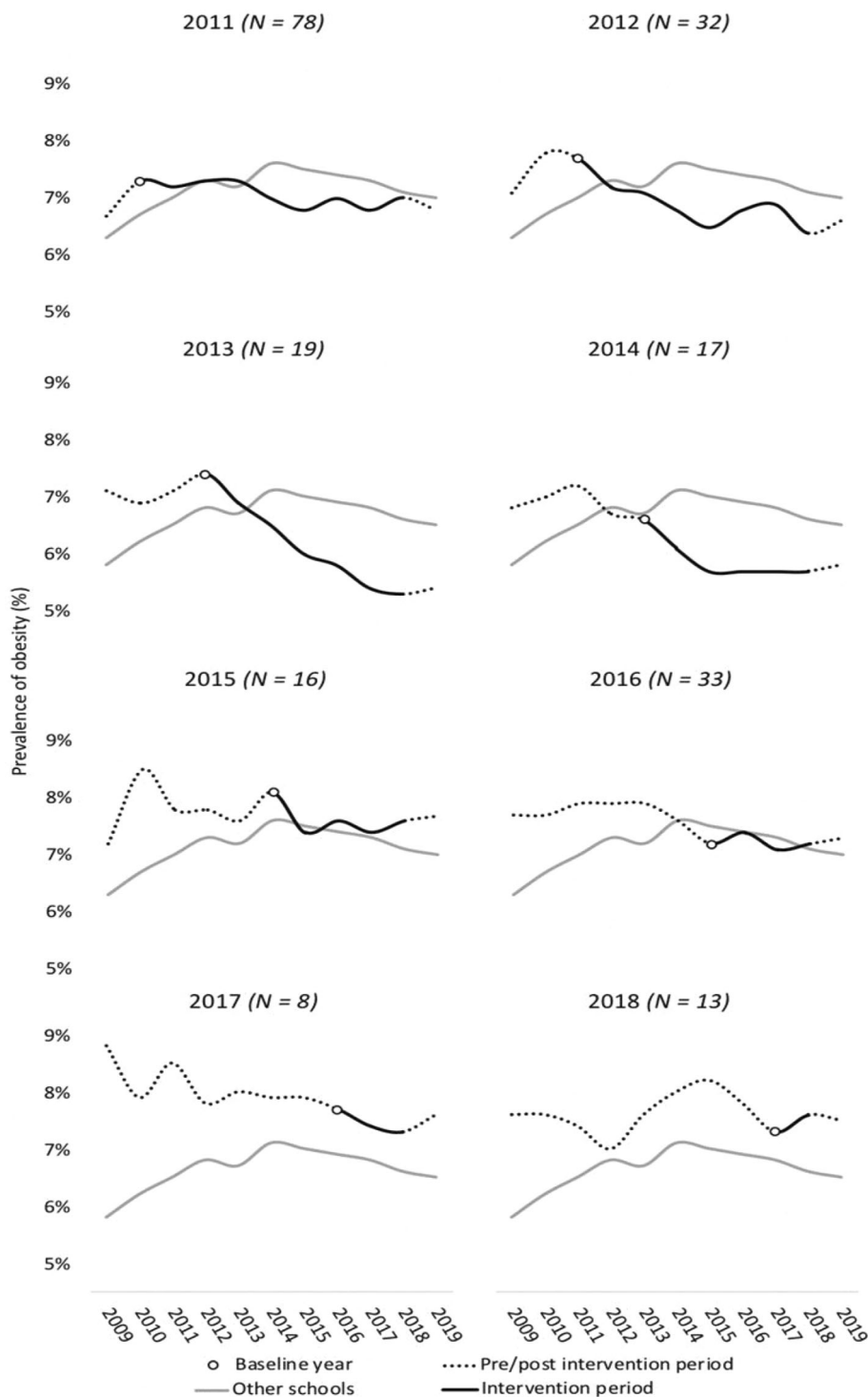
FIGURE 2 Time trends in obesity prevalence in intervention schools vs. other schools, stratified by intervention start year

having a much higher obesity prevalence at baseline, schools that joined the intervention in the first 5 years managed to achieve and sustain lower obesity rates than nonparticipating schools. The same was not seen for schools joining in 2016, 2017, or 2018.

Next, the GEE analysis showed an increasingly lower average BMI in the intervention group than in the control group, with longer participation durations being shown in the intervention group

across all three weight groups and in both sexes (Tables 3 and 4). Generally, in girls, the magnitude of effects plateaued after 3 years (girls with normal weight and girls with obesity) or 4 years of participation (girls with overweight) at around 1 to 1.4 kg/m^2 (girls with normal weight: $\beta = 0.937$, 95% CI: 0.845–1.029; girls with overweight: $\beta = 1.151$, 95% CI: 0.785–1.517; girls with obesity: $\beta = 1.417$, 95% CI: 0.959–1.875). In boys, the effects plateaued

TABLE 3 Generalized estimating equation analysis of effects on BMI change in control versus intervention in girls

Duration by weight status	β	SE	95% CI		Wald χ^2	p value
			Lower	Upper		
Normal weight						
1 Year	0.235	0.0219	0.192	0.278	114.487	<0.001
2 Years	0.831	0.0341	0.764	0.898	595.722	<0.001
3 Years	0.937	0.0467	0.845	1.029	402.059	<0.001
4 Years	0.807	0.0665	0.677	0.938	147.313	<0.001
5 Years	0.554	0.0875	0.383	0.725	40.130	<0.001
Overweight						
1 Year	0.157	0.0437	0.071	0.242	12.876	<0.001
2 Years	0.789	0.0742	0.644	0.935	113.004	<0.001
3 Years	1.097	0.1264	0.849	1.345	75.330	<0.001
4 Years	1.151	0.1867	0.785	1.517	37.966	<0.001
5 Years	0.887	0.2293	0.437	1.336	14.959	<0.001
Obesity						
1 Year	0.544	0.1212	0.306	0.781	20.115	<0.001
2 Years	1.333	0.1858	0.969	1.698	51.519	<0.001
3 Years	1.417	0.2338	0.959	1.875	36.744	<0.001
4 Years	0.953	0.325	0.316	1.590	8.594	<0.001
5 Years	0.397	0.4192	−0.424	1.219	0.898	0.340

TABLE 4 Generalized estimating equation analysis of effects on BMI change in control versus intervention in boys

Duration by weight status	β	SE	95% CI		Wald χ^2	p value
			Lower	Upper		
Normal weight						
1 Year	0.219	0.0211	0.178	0.26	107.688	<0.001
2 Years	0.717	0.0321	0.654	0.779	499.277	<0.001
3 Years	0.851	0.0433	0.766	0.935	386.191	<0.001
4 Years	0.807	0.0595	0.691	0.924	184.226	<0.001
5 Years	0.574	0.0772	0.422	0.725	55.238	<0.001
Overweight						
1 Year	0.413	0.0688	0.278	0.548	36.092	<0.001
2 Years	0.413	0.0688	0.278	0.548	36.092	<0.001
3 Years	0.766	0.1139	0.542	0.989	45.195	<0.001
4 Years	0.591	0.1636	0.270	0.912	13.039	<0.001
5 Years	0.240	0.2288	−0.209	0.688	1.099	0.295
Obesity						
1 Year	0.272	0.1031	0.070	0.474	6.965	0.008
2 Years	0.715	0.1514	0.419	1.012	22.344	<0.001
3 Years	0.889	0.2182	0.461	1.316	16.584	<0.001
4 Years	0.630	0.2964	0.049	1.211	4.514	0.034
5 Years	0.834	0.3707	0.107	1.56	5.056	0.025

TABLE 5 Number of reversed cases of obesity in children across five different durations of participation/nonparticipation

Duration	Sex	Control			Intervention			
		Number of obesity cases	Number of obesity cases reversed	Obesity reversal rate (per 10,000)	Number of obesity cases	Number of obesity cases reversed	Obesity reversal rate (per 10,000)	NNT
1 Year	Male	1111	9	81	1004	7	69	−886
	Female	1119	1	8	829	2	24	659
2 Years	Male	666	9	135	573	10	174	254
	Female	692	3	43	420	8	190*	68
3 Years	Male	428	10	233	338	20	591*	28
	Female	466	5	107	203	10	492*	26
4 Years	Male	195	12	615	154	16	1038	24
	Female	240	6	250	103	8	776*	19
5 Years	Male	104	5	480	94	13	1382*	12
	Female	138	2	144	54	4	740*	17

*$p < 0.05$.

Abbreviation: NNT, number needed to treat.

after 3 years at around 0.8 to 0.9 kg/m^2 (boys with normal weight: $\beta = 0.851$, 95% CI: 0.766–0.935; boys with overweight: $\beta = 0.766$; 95% CI: 0.542–0.989; boys with obesity: $\beta = 0.889$, 95% CI: 0.461–1.316).

To assess the effect of the intervention on obesity treatment, we analyzed the transition of children with obesity at baseline ($N = 4063$) to overweight or normal weight for each of the five participation durations. The results shown in Table 5 reveal that for the reversal of obesity, the intervention was more effective in girls for whom statistically significant differences between the control group and intervention group were seen after 2 years ($p = 0.016$, Cramer $V = 0.072$), 3 years ($p = 0.002$, Cramer $V = 0.120$), 4 years ($p = 0.024$, Cramer $V = 0.122$), and 5 years of participation ($p = 0.033$, Cramer $V = 0.154$). On the other hand, in boys, the intervention started to show effects a bit later, and the difference between the intervention group and the control group reached the significance threshold only in participation durations of 3 years ($p = 0.011$, Cramer $V = 0.092$] and 5 years ($p = 0.027$, Cramer $V = 0.157$). In line with this, the NNT also decreased steadily as the duration of the program increased for both sexes, with generally lower numbers being shown for girls. The lowest NNT was seen for 5 years of participation in the program (NNT = 17 and 12 for girls and boys, respectively).

DISCUSSION

This study investigated the effectiveness of a scaled-up, population-based, long-lasting PA intervention on obesity-related outcomes in children aged 6 to 14 years, while using a complex analytical design to reflect diverse, real-world scenarios. The principal result of the study was that children included in the Healthy Lifestyle intervention had a significantly lower BMI rise than the control group. The difference in BMI grew with the number of years of participation in the intervention and the difference was the largest in children who initially had

obesity. Furthermore, reversal of obesity was more common in the intervention group in cases in which children were involved in the program for at least a consecutive 3 years, whereas maximal treatment effects were seen after 4 years of participation in girls and after 5 years of participation in boys.

Our findings are in line with the observed (prepandemic) trends among Slovenian children, which show that the increase in the prevalence of overweight and obesity has been decreasing over the past decade, with a larger reduction being shown in boys than in girls [16]. The fact that childhood obesity has been declining in Slovenia throughout the period of the Healthy Lifestyle intervention at the population level [16] suggests that there were also other drivers that contributed to the reversal of the obesity trends.

We observed that the effect of the intervention on BMI was generally larger in girls than in boys, although this was more evident in those with overweight and obesity at baseline. Moreover, reversed obesity cases were also more frequent in girls and more consistent across different participation scenarios. The disparity between sexes in terms of different intervention effects could be due to differences in PA levels outside the school environment. During leisure time, girls are usually less active than boys [17], whereas PE participation is not related to overall PA levels for boys [18]. This implies that in relative terms, PA accumulated during the Healthy Lifestyle intervention constituted a higher share of daily PA in girls, causing larger intervention effects in girls than in boys. Contrary to our study, a recent meta-analysis of school-based PA programs reported that fitness-oriented interventions such as the one analyzed here produced larger effects in boys than in girls [7]. Whether the Healthy Lifestyle intervention might have provided a larger stimulus for girls than for boys given girls' lower daily PA level remains to be confirmed in future studies. Apart from the relative volume of PA delivered, boys and girls may need different types of PA to achieve the same effects, and the contents of the intervention examined here might have been more

appropriate for girls than for boys. In addition, the earlier maturation of girls might have confounded the intervention's effects on BMI, although we adjusted the analyses for height percentile to reduce this effect. Notwithstanding the increase in the BMI of girls during puberty due to the increase in subcutaneous fat, the reduction of BMI in this study was still more pronounced in girls than in boys.

We found that the Healthy Lifestyle intervention produced very large and clinically meaningful differences in BMI compared with the control condition, particularly in children with overweight and obesity. At the same time, it must be noted that BMI is not the most accurate estimate of adiposity because of a well-known limitation to its ability to distinguish between fat and muscle mass [19]. Thus, PA interventions could have positive effects on children with overweight or obesity by changing their body composition and altering their risk of the possible future health impairments associated with excess body fat, even in the absence of changes in body weight. Consequently, the effects of the Healthy Lifestyle intervention on body composition could be even higher than reported in this study if fat mass could have been used instead of BMI as an outcome measure. A study including children aged 7 to 17 years with obesity showed that even a small reduction in the BMI z score (≥0.00 to <0.10) improved health markers such as insulin sensitivity, which reduces the risk of future noncommunicable diseases [20]. Standards for BMI z score reduction for each baseline-weight-status subgroup of children and adolescents do not exist, which makes it difficult to evaluate the effectiveness of an intervention that is based exclusively on BMI values rather than on BMI values and additional health markers, limiting the reliability of conclusions about the intervention's clinical effectiveness. Although we adjusted for maturation effects in our models, we cannot rule out the residual effect of growth that could have resulted in the underestimation of the true effect of the intervention.

Although we focused our analyses on weight-related outcomes, there are several other benefits incurred by a fitness-oriented PA program such as the one analyzed here. An unhealthy weight status is associated with an increased risk of several chronic diseases, including diabetes, cardiovascular diseases, osteoarthritis, and some types of cancer [21]. Obesity is one of the main modifiable risk factors for insulin resistance in children and adolescents [22]. In addition, in children with obesity, adequate PA and an appropriate fitness level represent a favorable cardiovascular predictor despite excess adiposity, and good cardiovascular health can serve as a component that is protective against heart-related disease, even in childhood [23].

We found that Heathy Lifestyle was effective in reversing obesity, reaching its highest efficiency after 5 years of participation. Unsurprisingly, the longer the intervention lasts for a given child/youth, the more profound the potential treatment effect is, keeping in mind the concept of obesity as a chronic disease. Even though our intervention lasted 8 years, we were forced to limit data analyses to 5 years of participation because the number of children who persisted more than 5 years was less than 800 among boys and less than 500 among girls. Hence, it remains to be seen whether programs that last more than 5 years are accompanied with a further reduction in

NNT statistics. Our findings are in line with a small-scale study from Denmark that provided NNT estimates of an intensive PE program to prevent overweight or obesity in a cohort ($N = 1009$) of 5- to 12-year-old children [24]. The authors calculated that 18 children, irrespective of their weight status, would need to participate in 270 min/wk of PE for 2 years to prevent one additional case of overweight or obesity compared with the usual PE lasting 90 min/wk [24]. Although the study was not powered to detect effects exclusively in children with overweight, the fact that the effect was considerably higher when children with overweight were included in the analyses implies that there was also a significant treatment effect present. Taken together, these results, along with universal coverage and high participation rates, provide evidence for intensive compulsory PE as a viable and effective solution for childhood obesity prevention at the population level.

Scaling up public health interventions within real-world settings addresses public health when these interventions are vast in size and when they have an adequate extension to reach a larger proportion of the targeted population [25]. The drawbacks and risks associated with scaling up and real-world setting immersion are commonly related to a lack of funding and the consequential poor intervention implementation and loss of effectiveness [26]. This is exactly what happened to the population-based, scaled-up intervention analyzed in this study; when funding was limited for 1 year (i.e., 2015/2016) because of a legislative impediment, a reduction of the beneficial effects of the intervention on obesity-related outcomes was observed. On the other hand, only 2 of the 216 schools voluntarily decided to discontinue the intervention. Moreover, the large dropout of children after 5 years of the program noted in this study is mostly a natural phenomenon. Namely, many children who were in grade 4 when they started the program finished primary school after 5 years and continued their high school education. The sustainability of interventions across longer periods, when implemented in an entire population of school-aged children and adolescents, must be ensured for favorable outcomes to be present continuously. This is supported by the World Health Organization's recommendation, which emphasizes that interventions lasting more than a year should provide larger effects than shorter ones [27], but at the same time it shows that the recommended 1-year minimum might not be enough to achieve substantial results in reversing the current childhood obesity epidemic. Additionally, because the largest differences in BMI reduction between groups were present after 3 years of involvement, a continuous, all-in approach on every level, including funding, of the intervention is crucial.

The successful implementation of the program was supported by PE teachers, who possess more specific knowledge on PA than classroom teachers. A meta-analysis of interventions focusing on PE lessons showed that an effective intervention approach included appropriate instructions alongside class organization and management [28]. Moreover, PE teachers are more successful than classroom teachers in terms of enhancing the physical fitness of children because of their higher competence in planning and delivering PE lessons [29].

This study has many strengths: (1) the population-level, scaled-up intervention was delivered exclusively by PE specialist teachers in real-world settings; (2) the large number of children of diverse ages contributed to a higher generalizability of our findings; (3) the longitudinal design implemented over a very long, 5-year period allowed us to infer a causal relationship between the intervention and obesity-related outcomes, while also allowing us to examine the sustainability of the effects over a longer intervention period; (4) the analyses were stratified by sex, providing even more specific insight into the effects of the intervention; (5) the models adjusted for several important covariates—the baseline school-level obesity prevalence, economic affluence of the local environment, individual baseline risk of obesity, and individual maturation rate; (6) we controlled for maturation effects, which can blur the actual decline in body mass because of the increased accumulation of subcutaneous fat before the growth spurt and because of the increased gains in muscle mass in boys entering puberty and in fat mass in girls entering puberty [30, 31]; and (7) the program was delivered by specialist PE teachers rather than by less competent (in terms of PA instruction) classroom teachers or other professionals.

There are also several limitations that deserve highlighting. First, BMI is not the most accurate estimate of adiposity because of a well-known limitation to its ability to distinguish between fat mass and muscle mass [19]. Still, BMI is considered a feasible outcome variable because measuring BMI is noninvasive and is an easily applicable method in the context of ethical requirements and legislative regulations. Second, we used the height percentile as a proxy for the maturation rate, which is not an ideal measure of maturation because it only shows the percentile deviation from the expected height. Although higher percentiles at the time of the pubertal growth spurt are indicators of accelerated growth, this indicator is less reliable in preadolescent individuals. Third, we were unable to collect information about dietary habits, PA outside the intervention, or screen time, which represent important factors in childhood obesity management. In addition, there are also several other important determinants of obesity and weight change (other than diet and PA) that were not recorded in this study and were thus not included in analyses. Examples are genetic variation, epigenetics, endocrine disease, central nervous system pathology, sleep, infection, and socioeconomic and cultural factors [32]. Fourth, as this is a natural experiment and not a randomized control trial, we acknowledge the possibility of sampling bias due to the nonrandom, voluntary enrollment approach used in the study. Thus, it is possible that some children and adolescents who may have been prone to behavior change would have wanted to be a part of this intervention, as opposed to children with the opposite characteristics, who might have been more likely to decline participation. Nevertheless, baseline differences between the control group and the experimental group in terms of BMI and triceps skinfold thickness were small to trivial. In addition, we included the baseline risk of obesity and school-level prevalence of obesity as covariates in all analyses to reduce the possible effects of sampling

bias. Fifth, we do not have information on adherence, but the fact that the intervention was mandatory after enrollment underlined the importance of attending the predetermined program and guaranteed high attendance rates. Sixth, although the additional PE lessons provided within this intervention were organized to take place immediately after the end of regular school hours, they were not integrated into the obligatory curriculum. Although this fact has not affected participation rates in our context, we acknowledge that this might raise additional implementation barriers in other educational systems (e.g., transportation issues). Last, while horizontally scaling up the intervention, we were not able to collect other implementation outcomes across the various schools and participation years.

CONCLUSION

A school-based PA intervention that provided additional PE lessons remained effective in the prevention of obesity after scaling up. The greatest effect was present in children initially presenting with obesity, such that the program was able to benefit children needing support the most. The NNT for obesity reversal decreased as the intervention duration increased, emphasizing the need for long-term PA programs. Policy makers and funding bodies should be aware that obesity is a chronic condition that needs to be dealt with over a longer time frame and that easy solutions and immediate effects are neither realistic nor sustainable. Last, the population-scaled PA intervention analyzed here was shown to be more effective among girls than among boys, particularly for children living with overweight or obesity at the start of the program. The reasons for this sex inequality need to be elucidated to guide the design of more equitable future PA programs in schools.○

CONFLICT OF INTEREST

The authors declared no conflict of interest.

ORCID

Petra Jurić https://orcid.org/0000-0002-6514-107X

REFERENCES

1. Abarca-Gómez L, Abdeen ZA, Hamid ZA, et al. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet*. 2017;390(10113):2627-2642. doi:10.1016/S0140-6736(17)32129-3
2. World Obesity and World Health Organization. *Taking Action on Childhood Obesity Report*. WHO; 2018.
3. Bleich SN, Vercammen KA, Zatz LY, Frelier JM, Ebbeling CB, Peeters A. Interventions to prevent global childhood overweight and obesity: a systematic review. *Lancet Diabetes Endocrinol*. 2018;6(4): 332-346. doi:10.1016/S2213-8587(17)30358-3
4. Wang Y, Cai L, Wu Y, et al. What childhood obesity prevention programmes work? A systematic review and meta-analysis. *Obes Rev*. 2015;16(7):547-565. doi:10.1111/obr.12277
5. Summerbell CD, Waters E, Edmunds LD, Kelly S, Brown T, Campbell KJ. Interventions for preventing obesity in children. *Cochrane Database Syst Rev*. 2005;(3):CD001871. doi:10.1002/14651858.CD001871.pub2
6. Sember V, Jurak G, Kovač M, Morrison SA, Starc G. Children's physical activity, academic performance, and cognitive functioning: a systematic review and meta-analysis. *Front Public Health*. 2020;14(8): 307. doi:10.3389/fpubh.2020.00307
7. Podnar H, Jurić P, Karuc J, et al. Comparative effectiveness of school-based interventions targeting physical activity, physical fitness or sedentary behaviour on obesity prevention in 6- to 12-year-old children: a systematic review and meta-analysis. *Obes Rev*. 2021; 22(2):e13160. doi:10.1111/obr.13160
8. Kohl HW III, Craig CL, Lambert EV, et al. The pandemic of physical inactivity: global action for public health. *Lancet*. 2012;380(9838): 294-305. doi:10.1016/S0140-6736(12)60898-8
9. 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
10. Jurak G, Cooper A, Leskosek B, Kovac M. Long-term effects of 4-year longitudinal school-based physical activity intervention on the physical fitness of children and youth during 7-year followup assessment. *Cent Eur J Public Health*. 2013;21(4):190-195. doi:10.21101/cejph.a3823
11. Strel J, Ambrožič F, Kondrič M, et al. *Sports Educational Chart*. Ministry of Education and Sport; 1997.
12. 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(1):28-40. doi:10.14712/23366052.2020.4
13. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73(1):13-22.
14. Grady JJ, Helms RW. Model selection techniques for the covariance matrix for incomplete longitudinal data. *Stat Med*. 1995;14(13):1397-1416. doi:10.1002/sim.4780141302
15. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med*. 1988; 318(26):1728-1733. doi:10.1056/NEJM198806303182605
16. Sorić M, Jurak G, Đurić S, Kovač M, Strel J, Starc G. Increasing trends in childhood overweight have mostly reversed: 30 years of continuous surveillance of Slovenian youth. *Sci Rep*. 2020;10(1):11022. doi:10.1038/s41598-020-68102-2
17. Mota J, Silva P, Santos MP, Ribeiro JC, Oliveira J, Duarte JA. Physical activity and school recess time: differences between the sexes and the relationship between children's playground physical activity and habitual physical activity. *J Sports Sci*. 2005;23(3):269-275. doi:10.1080/02640410410001730124
18. Lenhart CM, Hanlon A, Kang Y, Daly BP, Brown MD, Patterson F. Gender disparity in structured physical activity and overall activity level in adolescence: evaluation of youth risk behavior surveillance data. *ISRN Public Health*. 2012;2012(29):674936. doi:10.5402/2012/674936
19. Shephard RJ. The obesity epidemic: a challenge to pediatric work physiologists? *Pediatr Exerc Sci*. 2005;17(1):3-17. doi:10.12691/ajssm-2-1-5
20. Caprio S. Insulin resistance in childhood obesity. *J Pediatr Endocrinol Metab*. 2002;15(suppl 1):487-492.
21. Hruby A, Hu FB. The epidemiology of obesity: a big picture. *Pharmacoeconomics*. 2015;33(7):673-689. doi:10.1007/s40273-014-0243-x
22. Chiarelli F, Marcovecchio ML. Insulin resistance and obesity in childhood. *Eur J Endocrinol*. 2008;159(suppl 1):S67-S74. doi:10.1530/EJE-08-0245
23. Legantis CD, Nassis GP, Dipla K, Vrabas IS, Sidossis LS, Geladas ND. Role of cardiorespiratory fitness and obesity on hemodynamic responses in children. *J Sports Med Phys Fitness*. 2012;52(3):311-318.
24. Learmonth YC, Hebert JJ, Fairchild TJ, Møller NC, Klakk H, Wedderkopp N. Physical education and leisure-time sport reduce overweight and obesity: a number needed to treat analysis. *Int J Obes*. 2019;43(10):2076-2084. doi:10.1038/s41366-018-0300-1
25. Milat AJ, King L, Bauman AE, Redman S. The concept of scalability: increasing the scale and potential adoption of health promotion interventions into policy and practice. *Health Promot Int*. 2013;28(3): 285-298. doi:10.1093/heapro/dar097
26. Reis RS, Salvo D, Ogilvie D, et al. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. *Lancet*. 2016;388(10051):1337-1348. doi:10.1016/S0140-6736(16)30728-0
27. World Health Organizatin. *Interventions on Dietand Physical Activity: What Works: Summary Report*. WHO; 2009.
28. Lonsdale C, Rosenkranz RR, Peralta LR, Bennie A, Fahey P, Lubans DR. A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical education lessons. *Prev Med*. 2013;56(2):152-161. doi:10.1016/j.ypmed.2012.12.004
29. Starc G, Strel J. Influence of the quality implementation of a physical education curriculum on the physical development and physical fitness of children. *BMC Public Health*. 2012;12(1):61. doi:10.1186/1471-2458-12-61
30. Tanner JM. Sequence, tempo, and individual variation in the growth and development of boys and girls aged twelve to sixteen. *Daedalus*. 1971;100(4):907-930.

31. Tanner JM. Growth and maturation during adolescence. *Nutr Rev.* 1981;39(2):43-55.
32. Güngör NK. Overweight and obesity in children and adolescents. *J Clin Res Pediatr Endocrinol.* 2014;6(3):129-143. doi:10.4274/Jcrpe.1471

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.