Original research

BMJ Open Long-term effects of a primary schoolbased overweight preventive intervention on physical fitness and physical activity: a propensity scorematched retrospective cohort study within the Generation R Study

Michel Sebastiaan Smit ⁽¹⁾, ¹ Famke Mölenberg ⁽¹⁾, ¹ Rosalie Mensink-Bout, ² Daan Nieboer, ¹ Trudy Voortman, ² Liesbeth Duijts, ³ Hein Raat ⁽¹⁾, ¹ Wilma Jansen^{1,4}

ABSTRACT

Mensink-Bout R, et al. Longterm effects of a primary school-based overweight preventive intervention on physical activity (PA). physical fitness and physical activity: a propensity scorematched retrospective

cohort study within the Generation R Study. BMJ Open 2025;15:e088272. doi:10.1136/ bmjopen-2024-088272

Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/ bmjopen-2024-088272).

To cite: Smit MS, Mölenberg F,

Received 01 May 2024 Accepted 28 March 2025



C Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

For numbered affiliations see end of article.

Correspondence to Dr Wilma Jansen:

w.jansen.1@erasmusmc.nl

Objectives In this study, we evaluate the long-term effects (±1.5 years postintervention) of 6-year exposure to the Lekker Fit! intervention on physical fitness and

Design The retrospective intervention evaluation is embedded within the Generation R Study in Rotterdam, the Netherlands, a population-based prospective birth cohort study.

Setting Measurements took place in the research centre of the Generation R cohort study.

Participants 5489 adolescents from the Generation R Study were eligible for inclusion within this study. Successful linking to school career data was possible for 4129 adolescents who were then retrospectively subdivided into a Lekker Fit! group, mixed group and regular school group based on their primary school career.

Interventions The Lekker Fit! intervention is a multicomponent primary school-based intervention for the prevention of overweight. It focuses on a healthy diet and healthy lifestyle rather than focusing directly on the reduction of overweight. The intervention targets individual behaviour of children as well as their obesogenic environment and parental engagement in shaping their children's behaviour.

Primary and secondary outcome measures Aged 13/14 years old, physical fitness was measured with an incremental ergometer test. The actual highest achieved work rate was divided by the expected highest achieved work rate (age- and sex-related Dutch population-based reference data), and converted into z-scores. PA was determined by the number of days with at least 1 hour of PA, obtained by a self-reported questionnaire. Propensity score matching was performed to correct for non-random selection bias. Linear regression analyses were performed to estimate intervention effects.

Results Children from the Lekker Fit! group had significantly lower fitness z-scores (-0.18 (95% CI -0.29 to -0.06), n=1826) compared with children from the matched regular school group. No Lekker Fit! intervention

STRENGTHS AND LIMITATIONS OF THIS STUDY

- \Rightarrow This study examined the long-term effect of a widely implemented public health intervention targeting schools in deprived areas, reaching 18 000 children annually.
- \Rightarrow Data from an ongoing population-based prospective cohort study were used and linked to external school career data.
- \Rightarrow Physical fitness was measured by trained research personnel in a medical centre using incremental ergometer test and wearing heart rate sensors.

gometer test and wearing heart rate sensors.
 ⇒ A propensity score matching analysis was used to adjust for the issue of non-random selection bias for receiving the intervention, but some confounding may have remained.
 ⇒ The intervention specifically targeted children at risk of becoming overweight, and therefore we looked for differential effects, but some strata were relatively small.
 effect was found on PA (-0.12 (95% Cl -0.36 to 0.12), n=1258).
 Conclusions No evidence was found for long-term favourable effects of a school-based multicomponent intervention on physical fitness and PA. Recommendations for policy and future research are discussed.
 INTRODUCTION
 Physical fitness and obesity are both important determinants of child health and

important determinants of child health and future health.¹² Interventions starting already in childhood to improve physical fitness (PF) and to prevent obesity are warranted for a healthy future life. Schools have often been suggested as an important setting for the delivery of lifestyle interventions in children. This is due to the continuous contact and because children with a wide range

Protected by copyright, including for uses related to text and

of backgrounds are reached.^{3 4} Two systematic reviews reported evidence for a promising role of school-based lifestyle interventions in the reduction of overweight in children.⁵ ⁶ Furthermore, school-based interventions with physical activity (PA) components can also positively contribute to physical fitness. Various intervention characteristics can improve physical fitness. This includes additional exercises (such as rope skipping and running during breaks or recess) in addition to regular physical education (PE) lessons, PE lessons with a minimal duration of 60 min, PE lessons at a minimal frequency of three times per week or combining aerobic and resistance exercises.⁷ Also, a Cochrane review reported a positive effect of school-based PA interventions on levels of PA engagement and PF after intervention durations of more than 12 weeks.⁸ However, evidence regarding the long-term effects following a period of at least 12 months without intervention is scarce.

Randomised controlled trials (RCTs) are regarded as the gold standard for intervention effect studies.⁹ However, the performance of RCTs studying school intervention effects with long-term follow-up is challenging. This is due to the transition of children to secondary schools, due to funding issues, and because it might be considered unethical to withhold (control) schools from implementing effective interventions.¹⁰ Therefore, information and evidence regarding the long-term effects and sustainability of school intervention effects after transitioning to secondary education is often lacking.⁵⁶¹⁰ Other types of designs, such as quasi-experimental and observational studies, may be used to study long-term effects of school interventions and can provide evidence with good external validity.^{11 12}

The Lekker Fit! (LF) intervention (translated as 'enjoy being fit intervention') is a primary school-based obesity prevention intervention in the city of Rotterdam, the Netherlands, containing PA and diet components.¹³ The LF intervention was widely implemented after the shortterm beneficial effects among children aged 6-9 years old on overweight prevalence, waist circumference and performance on the 20 m shuttle run test were established in a cluster RCT in 2006-2007 with a follow-up period of 9 months.¹⁴ However, it is currently unknown whether the beneficial effects of this intervention on overweightrelated outcomes and PF are sustained after transitioning to secondary school.⁸¹⁵

In this study, data from an ongoing birth cohort study (the Generation R Study) in Rotterdam, the Netherlands, were used to assess the long-term effects of the LF intervention. The aim of this research study was to determine the long-term effects of the LF intervention on PF and PA. There is an unequal distribution of health across society, and the equity impact of effective and sustainable interventions is unclear.¹⁶ Therefore, in explorative subgroup analysis, we assess whether associations differ by (a) gender, (b) household income, (c) ethnic background and (d) preintervention weight status.

MATERIALS AND METHODS Study design

A propensity score-matched design was used, embedded within the Generation R Study. The Medical Ethics Committee of Erasmus University Medical Centre in Rotterdam approved the study (MEC 217.595/2002/20). Written informed consent for participation was obtained from parents until the child was 12 years old, and from parents and children hereafter in accordance with Dutch legislation.

The Generation R Study is an ongoing populationbased prospective cohort study from fetal life until young adulthood.^{17 18} Pregnant women with an expected delivery date between April 2002 and January 2006 and Z living in Rotterdam, the Netherlands, were invited for 8 their participation and their newborn. The overall aim of the Generation R Study is to identify early environmental and genetic causes of normal and abnormal growth, development and health from fetal life until young adulthood.^{17 18} The large database from the Generation R Study includes children who were exposed to the LF intervention and comparable children who were not. The Generation R Study includes a very broad range of data Generation R Study includes a very broad range of data collected during the preschool period (ie, preinterven-tion data) that can be used for 'matching' of individuals who were not exposed to LF to those who were. Matching was important as the LF intervention was mostly implemented in more deprived neighbourhoods. Furthermore, the Generation R Study enables the inclusion of ē outcome variables that are measured after the transition and from primary schools to secondary schools.

Study population We included adolescents from the Generation R Study for whom informed consent was given for linking their cohort data with school data. Figure 1 presents a flow chart for the current study. For 5489 children in total, training, parental consents were given for linking to school career data. 1360 children were excluded because of unsuccessful linkage (n=1316) or for incomplete school career data (n=44). A total of 4129 children were included for group assignment in the current analysis.

Exposure LF intervention

In this study, children were considered part of the LF group or regular school group. An overview of the timeline of intervention, data collection and group assignment is presented in figure 2. By linking cohort data to school & career data from the municipal registration databases, we obtained information on all primary schools attended by each child. Combined with information about the implementation status of the LF intervention, three comparison groups were identified, the number of years of LF intervention exposure per child was calculated and the last age of primary school education was determined.

The three comparison groups that were identified were the LF group, a mixed group and the regular school group. Following Prochaska and DiClemente, we assumed that

≥

and similar

, tech



Figure 1 Flow chart of the participants in this study. PA, physical activity.

a duration of at least 6 months is needed for behaviour change to stabilise when it involves PA.¹⁹ In line with this theory, we included children with <6 months of LF intervention exposure in the regular school group. The children who attended an LF intervention school during all their primary school years were considered the LF group. The mixed group contained children with some years of LF intervention exposure during their primary school career and some years of regular school exposure. Due to very limited variability in years of LF intervention exposure, the mixed group was excluded from further analysis in this study (n=364). 3765 children remained available for data analysis. Using the last age of primary school education, we calculated the time since primary school.

e

Ing,

S

Time since primary school was used to create a sensitivity analysis study sample with participants with ≥ 1 year since primary school to avoid the risk that the estimated treatment effect was attributable to the children who were relatively recently exposed to the intervention.

The LF intervention

The LF intervention has been developed by the Sports Department of the City of Rotterdam, the Netherlands in close collaboration with primary schools.²⁰ The number of primary schools that implemented LF has increased from 20 in 2006 to 94 in 2020.¹³ The intervention targets children aged 6/7 to 12/13 years old, primarily in socioeconomically disadvantaged neighbourhoods because of a higher prevalence of overweight and obesity.²¹ In the 8 Dutch school system, children attend primary school for 8 years (usually between ages 4 and 12/13 years) and afterward start secondary school. Until 2012, the LF intervention was only implemented in grades 3-8 (approximately ages 6/7 to 12/13 years). From 2012 onwards, the LF intervention was additionally implemented in grades Бu 1 and 2 (approximately ages 4 to 5/6 years).

for uses r LF focuses on a healthy diet and healthy lifestyle rather than focusing directly on the reduction of overweight. The intervention targets the individual behaviour of children as well as their obesogenic environment and parental engagement in shaping their children's behaviour,^{14 20} as described in detail elsewhere.²² Briefly, LF entails multiple components including but not limited to an additional third PE lesson in comparison with regular school programmes, professional PE teachers instead of regular classroom teachers providing the PE lessons, additional voluntary PA sessions outside school hours in cooperation with sports clubs, special themed education on healthy Ξ lifestyle topics and since 2013 the promotion of drinking water.^{13 22} A complete overview of components of the LF intervention and the regular curriculum are published elsewhere.22

Physical fitness by the steep ramp test

Al training, and The outcome variable for PF was derived from the steep ramp test (SRT). The SRT is a cycle ergometer test with electronic resistance and was used according to the paediatric modified SRT protocol.²³ In the SRT, children were instructed to keep a pedalling frequency of 60-80 circulations per minute. After warming at 25 Watt, the resis-Ino tance of the ergometer slowly increased every 2s with increments depending on the body height of the child. The SRT was ended when the child was not able to maintain the minimum of 60 circulations per minute. Heart rate was monitored during the whole SRT using an elastic belt with a Polar heart rate sensor. The complete used protocol of the paediatric modified SRT is published elsewhere.²⁴

First, the highest achieved work rate measured in Watt during the test (Abs-WRpeak) was derived from the SRT. Furthermore, we calculated the predicted highest achieved work rate measured in Watt during



Figure 2 Groups, data collection and timeline within the Generation R Study to evaluate the LF intervention. LF, Lekker Fit!.

the test (Pred-WRpeak), based on sex- and age-specific population-based Dutch reference data.²⁵ By dividing Abs-WRpeak by Pred-WRpeak, we calculated the actual achieved proportion (Prop-WRpeak) of Pred-WRpeak and multiplied it by 100%. We standardised Prop-WRpeak into z-scores, using formula 1, to create our PF outcome variable WRpeak z-score. A higher WRpeak z-score indicates a higher anaerobic power and leg muscle strength, reflecting better PF of the child as compared with their peers.

z-score=(observed value-mean)/SD

Additionally, we derived the absolute heart rate peak (Abs-HRpeak) of participants during the SRT. Based on sex- and age-specific population-based data, we calculated the predicted absolute heart rate peak (Pred-HRpeak).²⁶ Hereafter, we calculated the actual achieved proportion (Prop-HRpeak) of Pred-HRpeak and multiplied it by 100%. To account for the role of motivation and secure maximum effort on the SRT, we used Prop-HRpeak for sensitivity analysis in which we repeat the analyses including only the participants with a Prop-HRpeak $\geq 85\%$.

Daily PA

Children completed a questionnaire at age 13 years. For the outcome PA, children completed questions adapted from the Activity Questionnaire for Adults & Adolescents and Health Behaviour in School-Aged Children.^{27 28} The outcome variable of PA in this study was the number of days with at least 1 hour of PA (PA-Days).

Sociodemographic characteristics and preintervention measures

Data on the sociodemographic characteristics included child's age, gender, ethnic background, parental educational level and net household income, which were obtained using parental questionnaires of the Generation group Dies rty Physical fitness (Steep Ramp Test) & Physical activity (Questionnaires) s old 13/14 years old children ion R Study to evaluate the LF intervention. LF, Lekker Fit!. R Study at baseline and during a follow-up visit when the children were 6 years old. Additionally, we collected data on the child's preintervention measures for sport participation, playing outside, body mass index (BMI) z-scores and weight status.

ated to A child's ethnic background was based on the country of birth of the parents and categorised into the largest groups in Rotterdam: native Dutch, Turkish, Moroccan, Surinamese, Antillean, Cape Verdean and Other. Maternal country of birth was chosen when parents were born in different countries. Maternal and paternal education levels were determined by the highest education level obtained. According to Statistics Netherlands, it was categorised into high level (university degree), midhigh level (higher vocational training), mid-low level (>3 \triangleright years general secondary school, intermediate vocational training) and low level (≤ 3 years of general secondary education, intermediate general school, lower vocational training, primary school or no education).²⁹ Net Dd household income was categorised into low (<€2000 per month), intermediate (€2000-€3200 per month) and high (> \in 3200 per month). The cut-off point of \in 3200 per month distinguishes between lower and higher than average Dutch net household income.

Data on sport participation and outdoor play were obtained using parental questionnaires. Parents were asked if their child participates in sport (yes/no). For outdoor play, parents were asked 'On average, how many weekdays per week does your child play outside?' and 'On average, how many week-end days per week does your child play outside?'. The answers were combined into an average amount of days per week in which the child plays outside. Finally, the variable was categorised into 'Up to 5 days per week' and 'More than 5 days per week'.

Children's body height and body weight measurements (without wearing shoes and heavy clothing) were centre in the Erasmus Medical Center. Preintervention BMI was calculated in kg/m^2 , and sds-scores (BMIsds) were calculated based on sex- and age-specific population-based Dutch reference models.³⁰ Based on BMI, children's weight status was categorised into child with normal weight or child with overweight, using standardised international sex- and age-specific populationbased cut-offs.³¹ **Statistical analysis** We created two separate study samples for analysis based on data availability on the outcome variables. The first consisted of all children with data on the outcome vari-

conducted by trained staff during a visit at the research

able WRpeak z-score and the second of all children with data on the outcome variable PA-Days. We described the unmatched and unimputed sociodemographic characteristics and baseline characteristics for both study samples. Hereafter, we conducted multiple imputation for missing sociodemographic characteristics and preintervention data in both study samples by chained equations (M=5) using the RStudio package mice.

For both study samples, we conducted propensity score matching³² to match children from the LF group and regular school group with similar input characteristics to adjust for the issue of non-random selection bias for receiving the LF intervention. To do so, we used logistic regression models in which treatment status was regressed on the sociodemographic and baseline characteristics of age, gender, ethnic background, maternal educational level, paternal educational level, net household income, sport participation, playing outside and BMI-sds. Propensity scores were then used to match the children in the LF group and regular school group with 1:2 nearest neighbour matching without replacement and calliper width 0.1, by using the RStudio package MatchThem. The matched samples were checked on balance in sociodemographic and baseline characteristics by visual inspections and by standardised mean differences. We described the matched sociodemographic characteristics and preintervention measures in the matched study samples. Additionally, we reported the mean values of Prop-WRpeak per group for comparison purposes with Dutch reference models from healthy populations.

With both the two propensity score-matched models, we examined the causal effect of the LF intervention in comparison with regular school on the outcome variable WRpeak z-score and the outcome variable PA-Days in the five multiple imputed datasets. The average pooled treatment effects using t-tests were reported to compare the outcomes between the LF group and regular school group. Data analyses were performed with RStudio software V.3.4.1. Level of statistical significance was set at p<0.05.

Exploratory subgroup analyses

In explorative subgroup analyses, we examined whether the observed effects differed between (a) gender (girls

vs boys), (b) net household income ($\leq \in 3200/\text{month}$ vs $\geq 3200/\text{month}$), (c) ethnic background (native Dutch vs non-native Dutch) and (d) weight status (children with normal weight vs children with overweight). Differences were assessed by conducting stratified analysis and by testing for differences in effect parameters using z-tests.

Sensitivity analysis

In a first sensitivity analysis, we repeated the main analysis for the outcome variable WRpeak z-score from the SRT, including only the participants with a Prop-HRpeak \geq 85%, to account for the potential role of motivation and maximum effort on the SRT.

In a second sensitivity analysis, we repeated the main copyr analysis for both the outcome variables WRpeak z-score and PA-Days, including only the participants with a time ʻight, since primary school of ≥ 1 year to avoid the risk that the estimated treatment effect was attributable to the children including for uses related to text who were relatively recently exposed to the intervention.

Patient and public involvement

None.

RESULTS

Description of the analytical samples and propensity score matching

Table 1 shows both the unmatched and matched sociodemographic characteristics and preintervention measures of the study sample for PF. After matching, the characteristics were balanced between the groups. 51.5% of the childata dren were girls, 43.6% had a native-Dutch background, and the mean BMI-sds score was 0.27. Approximately half \exists of the children had parents with low or mid-low educational levels (59.6% of the fathers and 55.9% of the mothers), and a third of the households (34.3%) had a ≥ training net income of <€2000/month. The characteristics of the study sample for PA are included in online supplemental additional file 1.

We were able to match the vast majority of the participants in the LF group with one or two peers in the regular group (online supplemental additional files 2 and 3). Sample sizes in the matched study samples for PF and PA were, respectively, n=1826 and n=1258 (figure 1). With the propensity score matching, we have been able to reduce all the absolute standardised mean differences to below the 0.1 threshold in both the matched study gamples (online supplemental additional files 4 and 5).

Long-term LF intervention effect on physical fitness and PA

The mean values for Prop-WRpeak were lower than 100% for both children in the LF group (92.2±15.3) and children in the regular school group (97.7±14.9). The same findings were found for all different subgroups in exploratory subgroup analyses. These proportions are calculated based on sex- and age-specific population-based Dutch reference data. Therefore, these scores indicate

and

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies Enseignement Superieur (ABES

Sociodemographics and preintervention measures of unmatched and PSM samples in study sample for physical Table 1 fitness*

	Unmatched†			PSM‡	
	Lekker Fit! group (n=693)	Regular school group (n=2302)	Missing data (%)	Lekker Fit! group (n=679)	Regular school group (n=1147)
Age (years), mean (SD)	6.17 (0.50)	6.06 (0.40)	5.8	6.15	6.16
Sex, %			0.0		
Girls	52.2	50.8		52.1	51.2
Ethnicity, %			1.4		
Native Dutch	42.7	69.3		43.7	43.5
Turkish	8.8	2.5		8.0	7.6
Moroccan	10.4	2.9		9.7	8.8
Surinamese	11.9	6.1		12.1	12.7
Antillean	4.2	2.4		4.3	4.3
Cape Verdean	4.9	1.9		5.0	4.9
Other	17.1	14.9		17.2	18.1
Maternal education level, %			12.9		
High	21.4	33.9		20.3	20.8
Mid-high	23.1	31.1		24.0	24.2
Mid-low	37.1	28.0		37.6	37.6
Low	18.4	7.0		18.1	17.4
Paternal education level, %			20.0		
High	25.3	39.4		21.5	21.4
Mid-high	18.7	25.2		18.7	19.0
Mid-low	33.3	24.7		33.1	33.4
Low	22.7	10.8		26.7	26.1
Net household income, %			17.5		
<€2000/month	31.7	14.2		34.2	34.3
€2000–€3200/month	34.3	22.7		32.8	32.3
>€3200/month	34.0	63.1		33.0	33.4
BMI-sds, mean (SD)	0.28 (0.96)	0.20 (0.86)	5.9	0.28	0.26
Sports participation, %			13.7		
Yes	35.5	47.9		34.9	35.3
No	64.5	52.1		65.1	64.7
Outdoor play, %			15.0		
<5 days/week	37.7	31.3		38.6	38.0
≥5 days/week	62.3	68.7		61.4	62.0
Weight status, %			5.9		
No overweight	79.0	86.1		78.2 ^{n.m.}	81.8 ^{n.m.}
Overweight/obese	21.0	13.9		21.8 ^{n.m.}	18.2 ^{n.m.}

*Study sample for physical fitness contains participants with data on the outcome measure WRpeak z-scores.

†Non-imputed sample.

‡Imputed sample.

BMI, body mass index; n.m., sample is not matched on weight status, as it was matched on BMI-sds; PSM, propensity score matched.

that children in both groups scored lower than their peers on the Dutch national level.

Table 2 presents the estimated long-term treatment effect of ±6 years of LF intervention on the outcomes

assessed ±1.5 years after transitioning to secondary education. Regarding physical fitness, WRpeak z-scores in the LF group were -0.26 (5% CI -0.33 to -0.17), whereas scores in the regular school group were -0.08

Table 2 Estimated long-terr	2 Estimated long-term treatment effect of 6 years of the LF intervention on Z-WRpeak and PA-Days								
	N	Regular school group (mean)	Lekker fit! group (mean)	Effect of intervention (95% CI)					
WRpeak z-score	Х	-0.08 (-0.17 to 0.01)	-0.26 (-0.33 to -0.17)	-0.18 (-0.29 to -0.06)					
PA-Days (days/week)	Х	4.11 (3.96 to 4.27)	3.99 (3.82 to 4.17)	-0.12 (-0.36 to 0.12)					

Results are the average treatment effects on the standardised achieved proportion of the predicted highest achieved work rate, and number of days with at least 1 hour of PA, obtained from a linear regression model on propensity score matched sample. The propensity score matched sample was matched on age, gender, ethnic background, maternal educational level, paternal educational level, net household income, sport participation, playing outside and BMI-sds.

Significance at the 0.05 level is indicated in bold.

85%-HR-sample: sample only including participants with $\ge 85\%$ achieved proportion of the predicted absolute heart rate to account for the role of motivation.

LF, Lekker Fit!; PA-Days, days with at least 1 hour of physical activity; WRpeak z-score, standardised achieved proportion of the predicted highest achieved work rate.

(95% CI –0.17 to 0.01). The children from the LF group had significantly lower WRpeak z-scores (–0.18 (95% CI –0.29 to –0.06)) compared with children from the regular school group. Regarding PA, PA-Days in the LF group was 3.99 (95% CI 3.82 to 4.17), whereas it was 4.11 (95% CI 3.96 to 4.27) in the regular school group. No LF intervention effect was found on PA-Days (–0.12 (95% CI –0.36 to 0.12)).

Exploratory subgroup analysis

Table 3 presents results from the subgroup analyses by gender, ethnic background, net household income and pre-intervention weight status on the outcomes. There were no differences in the estimated treatment effect of the LF intervention on all the outcomes between any of the explored subgroups.

Sensitivity analysis accounting for motivation

Online supplemental additional file 6 presents the estimated long-term treatment effect of the LF intervention on zWRpeak z-score in a population with \geq 85% achieved proportion of the predicted absolute heart rate, to account for the potential role of motivation. The result for WRpeak z-score (-0.18 (95% CI -0.28 to -0.08)) was similar to that of the main analysis.

Sensitivity analysis accounting for motivation

Online supplemental additional file 7 presents the estimated long-term treatment effect of the LF intervention on zWRpeak z-score in a population with ≥ 1 year since primary school. In contrary to the main analysis, the result for WRpeak z-score (-0.10 (95% CI -0.24 to 0.01)) was not significant in the sensitivity analysis. The result for

 Table 3
 Subgroup analysis for the estimated long-term treatment effect of the LF intervention on WRpeak z-score and PA-Days

		WRpeak z-score		PA-Days	
Subgroups		Effect of intervention (95% CI)	Subgroup effect Z-test (P value)	Effect of intervention (95% CI)	Subgroup effect Z-test (P value)
Gender	Boys	-0.17 (-0.34 to -0.00)	0.22 (0.83)	0.10 (-0.27 to 0.46)	1.66 (0.10)
	Girls	-0.19 (-0.33 to -0.06)		-0.33 (-0.69 to 0.03)	
Ethnic background	Native Dutch	-0.17 (-0.29 to -0.04)	1.15 (0.25)	-0.05 (-0.37 to 0.26)	0.36 (0.72)
	Non-native Dutch	-0.28 (-0.42 to -0.13)		-0.14 (-0.49 to 0.21)	
Net household income	< €3200 per month	-0.13 (-0.27 to 0.01)	0.08 (0.94)	-0.23 (-0.58 to 0.12)	0.35 (0.72)
	> €3200 per month	-0.14 (-0.34 to 0.06)		-0.13 (-0.58 to 0.32)	
Pre-intervention weight status	No overweight	-0.18 (-0.29 to -0.06)	0.07 (0.94)	-0.16 (-0.44 to 0.12)	0.51 (0.61)
	Overweight	-0.19 (-0.46 to 0.09)		0.02 (-0.63 to 0.67)	

Results are the average treatment effects on the standardised achieved proportion of the predicted highest achieved work rate, and number of days with at least 1 hour of PA, obtained from a linear regression model on propensity score matched sample. The propensity score matched sample was matched on age, gender, ethnic background, maternal educational level, paternal educational level, net household income, sport participation, playing outside and BMI-sds. Differences were assessed by conducting stratified analysis and by testing for differences in effect parameters using z-tests.

Significance for the subgroup effect at the 0.05 level is indicated in bold.

LF, Lekker Fitl; PA-Days, days with at least 1 hour of physical activity; WRpeak z-score, standardised achieved proportion of the predicted highest achieved work rate.

PA-Days (-0.16 (95% CI -0.45 to 0.14)) was similar to that of the main analysis.

DISCUSSION

In this study, we assessed the long-term effects of a 6-year primary school-based intervention (LF), containing PA and diet components, on PF and PA. Children in the LF group had lower WRpeak z-scores compared with children in the regular school group, whereas we found no treatment effect on PA-Days. Findings from the exploratory subgroup analyses indicated that no differences in treatment effect were apparent between subgroups of gender, ethnic background, net household income and preintervention weight status. Our sensitivity analysis including only the participants with an achieved proportion of $\geq 85\%$ of the predicted absolute heart rate peak during the SRT to account for motivation differences among participants taking the test showed similar findings. In contrary to the main analysis, in our sensitivity analysis including only participants with ≥ 1 year since primary school, no treatment effect was found on WRpeak z-scores.

The short-term effects of LF have previously been studied among primary school-aged children in 2006-2007 in a cluster RCT after 9 months of intervention exposure. Significant positive effects were reported for a lower percentage of children with overweight, lower waist circumference and higher PF among 6-9 yearold pupils, whereas no effects were found among 9-12 year-old pupils.¹⁴ The current study found no evidence for the long-term effects of LF on PF and PA. This may suggest that the positive short-term intervention effects on PF are not sustained after the transition from primary school to secondary school. This may be due to different factors, of which we discuss two major points here. First of all, after the transition to secondary schools, the children become exposed to the Rotterdam environment, which became increasingly obesogenic, especially in the lower socioeconomic areas.³³ The obesogenic environment that surrounds children and shapes their behaviour certainly plays an important role.^{34 35} It may oppose the beneficial effects that an ended intervention might yet have accomplished. Otherwise, the intervention may not have adapted enough to the changing environment and consequently have become less effective over the years. Another factor possibly contributing to the low measures of PA and fitness may be the change in the level of PA after the transition from primary school to secondary school. According to Statistics Netherlands, the percentage of children that meet the Dutch PA guidelines is minimally 20% lower among children in secondary school (12-16 year-olds) compared with children in primary school (4-12 year-olds) for the past 4 years.³⁶ The intervention may not have been effective to counteract this. Differences in results might also be partly explained by the study samples that were different between the 2006-2007 RCT (10.4% children with Dutch background and 24.4% children with preintervention overweight in grades 3-5

and 31.2% in grades 6–8) and the current study (63.6% children with native-Dutch background and 15.5% children with overweight preintervention and 56.9% of the parents with mid-high to high levels of education). The current study population received the intervention after a scale-up of the LF intervention to half of all primary schools in the city, whereas it was originally targeted at the disadvantaged neighbourhoods.

disadvantaged neighbourn An earlier review among 3–18 year-old children rock on the >6 months postintervention sustained impact of PA and PF interventions on PA and PF.³⁷ there may be various explanations. At first, there is uncertainty about the implementation status of intervention **Z** components. Unfortunately, no detailed information 8 is available about the actual implementation status and quality of implementation of the different LF intervention components. Further, we cannot exclude the possibility that regular schools in Rotterdam, the Netherlands, had implemented diet- or PA-related programmes or interventions into their curricula. Due to these uncer-٥ tainties, the differences between the curricula of LF schools and regular schools may have been smaller than anticipated. Furthermore, the effectiveness of the LF intervention may be subject to the environment, which may have become more obesogenic over the years. For instance, as measured in the Generation R cohort, for children, the exposure to fast food outlets increased over $\overline{\mathbf{a}}$ the past years, an increase that was even stronger for ŧ children with lower-educated parents.³³ This potentially opposes the accomplished effects immediately postintervention. Finally, the low scores on Prop-WRpeak, which were lower than 100% in both children in the LF group \vec{a} and regular school group, mean that children performed worse than expected based on Dutch population-based reference models. Sensitivity analysis demonstrated that ≥ this was irrespective of motivation. It indicates a relatively unfavourable state of PF in the Rotterdam children population. In line with a recent systematic review on secular trends in PF, it warrants further initiatives in PA and fitness promotion for children.³⁸

A main strength of this study was the use of a valid incremental ergometer test for the objective determination of outcome variable PF to gain valuable new insights into the effects of the intervention on so far less studied outcomes. Furthermore, the collection of heart rate data during the performance in the SRT allowed us to account for the role that motivation may play in $\ensuremath{\underline{\mathsf{G}}}$ physical incremental tests and forms a major strength of **3** this study. Finally, a major strength of this study was the propensity score matching procedure to account for the issue of non-random selection bias in the assignment for intervention. Using the large database with a very broad range of data collected during the preschool period, we were able to create two similar comparison groups regarding preintervention measures and characteristics. We cannot rule out the possibility that important unmeasured bias-sensitive information was missed. For instance,

preintervention data on PF was not available to include in the matching procedure. Presented means should therefore be interpreted with caution. Also, the representativeness of the cohort for the eligible LF population of Rotterdam, the Netherlands, may be limited. Our study sample includes considerable numbers of adolescents with higher socioeconomic backgrounds. Still, the presented effects were adjusted for selection bias, and subgroup analyses were performed. Our subgroup analysis has given no indication of the influence of net household income on LF intervention effectiveness. Also, the absence of data on intervention implementation, adherence and quality of teacher training in this study may be considered a limitation of this intervention effectiveness study. Due to the retrospective design of this study, this data was unavailable for the researchers. Furthermore, the use of self-reported data regarding PA behaviour might be considered a limitation of this study. This was due to the unavailability of objectively measured PA behaviour data in the Generation R cohort at the age of 13 years old. The use of self-reported data might have introduced underestimation, overestimation, recall bias and social desirable answering.³⁹ Recall bias can be a factor if the respondents have to respond based on previous experiences, memories or habitual patterns.⁴⁰ Because the length of the recall period in this questionnaire was very short, the influence of recall bias probably remains limited. As a final limitation, we could not match children on school-level indicators because of too low number of children in each school.

Based on this study, we have several recommendations for research and policy. Our study was not designed to give insight into the course of potential intervention effects over the years in primary and secondary school. However, our study results call for closer monitoring on desired quality of implementation and on desired intervention effects after wider implementation of initially promising interventions. Continued effort might be needed to retain an effective mixture of different intervention components as the contexts of children and schools might change. Therefore, in addition to evaluating overall intervention effects as in this study, more knowledge on the effectiveness of separate intervention components on behaviour, stratified for age groups, is needed. Knowledge on the efficacy of separate components may aid in the creation of new interventions or in the adjustment of current interventions (addition, alteration or removal of components) with greater effectiveness. To perform such efficacious research, researchers and municipal authorities should collaborate to gain detailed insight into the actual quality of intervention implementation and intervention adherence per school. Additionally, insight is needed regarding the potentially implemented programmes in regular schools to map in detail the differences between intervention schools and regular schools over time. Finally, more knowledge is required about intervention effects on the determinants of PA- and diet-related behaviour that may aid in the understanding of the pathways that drive

health behaviour change and sustainability of intervention effects.

In this study, we found no long-term effects of the LF intervention after the transition to secondary schools on PF and PA. The low PF scores for children, especially in the LF group, in comparison to population-based reference models warrant further initiatives in PA and fitness promotion for children. Prolonged interventions from primary schools to secondary schools or the implementation of new additional interventions may be needed to improve and maintain healthy lifestyles and PF among children and adolescents. Further research is needed into the sustained effects of up-scaled interventions and into the efficacy of different combinations of intervention components.

Author affiliations

¹Public Health, Erasmus MC, Rotterdam, The Netherlands ²Erasmus MC, Rotterdam, The Netherlands

³Pediatrics—Division of Respiratory Medicine, Erasmus MC, Rotterdam, The Netherlands

⁴Social Development, Municipality of Rotterdam, Rotterdam, The Netherlands

X Famke Mölenberg @FamkeMolenberg and Hein Raat @heinraat

Acknowledgements The Generation R Study is being conducted by the Erasmus Medical Centre in close collaboration with the School of Law and the Faculty of Social Sciences of the Erasmus University, Rotterdam; the Municipal Public Health Service, Rotterdam area; the Rotterdam Homecare Foundation and the Stichting Trombosedienst and Artsenlaboratorium Rijnmond, Rotterdam, The Netherlands. We gratefully acknowledge the contribution of the participating children, their mothers, general practitioners, hospitals, midwives and pharmacies in Rotterdam. Further, we thank the City of Rotterdam for sharing school data.

Contributors MSS, FM and WJ conceptualised and designed the study. MSS, FM and DN performed data analysis. MSS, FM, RM-B, DN, TV, LD, HR and WJ interpreted the analysis. MSS wrote the manuscript. FM, RM-B, DN, TV, LD, HR and WJ contributed to the manuscript by making critical revisions and giving comprehensive feedback on multiple drafts. MSS, FM, RM-B, DN, TV, LD, HR and WJ read and approved the final manuscript. WJ is the guarantor.

Funding The Generation R Study was made possible by financial support from the Erasmus Medical Centre, Erasmus University Rotterdam, the Netherlands Organisation for Health Research and Development, the Netherlands Organisation for Scientific Research, the Ministry of Health, Welfare and Sport and the Ministry of Youth and Families. The study presented in the current paper was supported by a grant from the City of Rotterdam, Department of Sport. FM received funding from the Erasmus Initiative Smarter Choices For Better Health. LD received funding for projects from the European Union's Horizon 2020 research and innovation programme (LIFECYCLE, grant agreement No 733206, 2016; EUCAN-Connect grant agreement No 824989; ATHLETE, grant agreement No 874583). The researchers are independent from the funders. The study sponsors had no role in the study design, data collection, data analysis, interpretation of data or writing of this report.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval This study involves human participants and was approved by the Medical Ethics Committee of Erasmus University Medical Centre in Rotterdam (MEC 217.595/2002/20). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those

Open access

of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs

Michel Sebastiaan Smit http://orcid.org/0000-0003-4231-3972 Famke Mölenberg http://orcid.org/0000-0002-5305-9730 Hein Raat http://orcid.org/0000-0002-6000-7445

REFERENCES

- 1 Kaminsky LA, Arena R, Ellingsen Ø, *et al.* Cardiorespiratory fitness and cardiovascular disease - The past, present, and future. *Prog Cardiovasc Dis* 2019;62:86–93.
- 2 Adhikary D, Barman S, Ranjan R, et al. A Systematic Review of Major Cardiovascular Risk Factors: A Growing Global Health Concern. Cureus 2022;14:e30119.
- 3 Verjans-Janssen SRB, Gerards SMPL, Kremers SPJ, *et al.* Effects of the KEIGAAF intervention on the BMI z-score and energy balancerelated behaviors of primary school-aged children. *Int J Behav Nutr Phys Act* 2020;17:105.
- 4 Corder K, Schiff A, Kesten JM, et al. Development of a universal approach to increase physical activity among adolescents: the GoActive intervention. BMJ Open 2015;5:e008610.
- 5 Bleich SN, Vercammen KA, Zatz LY, *et al.* Interventions to prevent global childhood overweight and obesity: a systematic review. *Lancet Diabetes Endocrinol* 2018;6:332–46.
- 6 Liu Z, Xu H-M, Wen L-M, et al. A systematic review and metaanalysis of the overall effects of school-based obesity prevention interventions and effect differences by intervention components. Int J Behav Nutr Phys Act 2019;16:95.
- 7 Minatto G, Barbosa Filho VC, Berria J, et al. School-Based Interventions to Improve Cardiorespiratory Fitness in Adolescents: Systematic Review with Meta-analysis. Sports Med 2016;46:1273–92.
- 8 Dobbins M, Husson H, DeCorby K, et al. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. Cochrane Database Syst Rev 2013;2013:CD007651.
- 9 Hariton E, Locascio JJ. Randomised controlled trials the gold standard for effectiveness research: Study design: randomised controlled trials. *BJOG* 2018;125:1716.
- 10 Smit MS, Boelens M, Mölenberg FJM, et al. The long-term effects of primary school-based obesity prevention interventions in children: A systematic review and meta-analysis. *Pediatr Obes* 2023;18:e12997.
- 11 Gilmartin-Thomas JF, Liew D, Hopper I. Observational studies and their utility for practice. *Aust Prescr* 2018;41:82–5.
- 12 Bärnighausen T, Tugwell P, Røttingen J-A, et al. Quasi-experimental study designs series-paper 4: uses and value. J Clin Epidemiol 2017;89:21–9.
- 13 Samenwerken aan een fitte generatie! over lekker fit!: gemeente rotterdam. n.d. Available: http://www.lekkerfit010.nl
- 14 Jansen W, Borsboom G, Meima A, et al. Effectiveness of a primary school-based intervention to reduce overweight. Int J Pediatr Obes 2011;6:e70–7.
- 15 Brown T, Moore TH, Hooper L, *et al.* Interventions for preventing obesity in children. *Cochrane Database Syst Rev* 2019;7:CD001871.
- 16 Love RE, Adams J, van Sluijs EMF. Equity effects of children's physical activity interventions: a systematic scoping review. Int J Behav Nutr Phys Act 2017;14:134.

- 17 Generation R, Erasmus Medical Center. The generation R study. 2022. Available: https://generationr.nl/researchers/
- 18 Kooijman MN, Kruithof ČJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. Eur J Epidemiol 2016;31:1243–64.
- 19 Prochaska JO, DiClemente CC. Stages of change in the modification of problem behaviors. *Prog Behav Modif* 1992;28:183–218.
- 20 Jansen W, Raat H, Zwanenburg EJ, *et al.* A school-based intervention to reduce overweight and inactivity in children aged 6-12 years: study design of a randomized controlled trial. *BMC Public Health* 2008;8:257.
- 21 Jansen W, Hazebroek-Kampschreur AA. Differences in height and weight between children living in neighbourhoods of different socioeconomic status. *Acta Paediatr* 1997;86:224–5.
- 22 Smit MS, Raat H, Mölenberg F, et al. Study protocol for the evaluation of long-term effects of the school-based obesity prevention program Lekker Fit! ('enjoy being fit'): a retrospective, controlled design. *BMJ Open* 2021;11:e046940.
- 23 Bongers BC, Takken T. The paediatric version of the steep ramp test. *J Physiother* 2014;60:113.
- 24 Mensink-Bout SM, Jahangir MR, de Jongste JC, *et al.* Associations of physical condition with lung function and asthma in adolescents from the general population. *Pediatr Allergy Immunol* 2022;33.
- 25 Bongers BC, de Vries SI, Obeid J, et al. The Steep Ramp Test in Dutch white children and adolescents: age- and sex-related normative values. *Phys Ther* 2013;93:1530–9.
- 26 Bongers B, van brussel M, Hulzebos HJ, et al. Pediatric norms for cardiopulmonary exercise testing in relation to sex and age. 2014.
- 27 Currie C, Inchley J, Molcho M, *et al.* Health behaviour in schoolaged children (HBSC) study protocol: background, methodology and mandatory items for the 2013/14 survey. Child and Adolescent Health Research Unit (CAHRU), 2014.
- 28 Chinapaw MJM, Slootmaker SM, Schuit AJ, *et al.* Reliability and validity of the Activity Questionnaire for Adults and Adolescents (AQuAA). *BMC Med Res Methodol* 2009;9:58.
- 29 Pleijers A, de Vries R. Indeling van opleidingen op basis van niveau en oriëntatie. Statistics Netherlands; 2021. Available: https://www. cbs.nl/nl-nl/longread/discussion-papers/2021/invulling-praktisch-entheoretisch-opgeleiden/3-indeling-van-opleidingen-op-basis-vanniveau-en-orientatie
- 30 Fredriks AM, van Buuren S, Wit JM, *et al*. Body index measurements in 1996-7 compared with 1980. *Arch Dis Child* 2000;82:107–12.
- 31 Cole TJ, Bellizzi MC, Flegal KM, et al. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1240–3.
- 32 Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res* 2011;46:399–424.
- 33 Mölenberg FJM, Mackenbach JD, Poelman MP, et al. Socioeconomic inequalities in the food environment and body composition among school-aged children: a fixed-effects analysis. Int J Obes (Lond) 2021;45:2554–61.
- 34 Hill JO, Peters JC. Environmental contributions to the obesity epidemic. Science 1998;280:1371–4.
- 35 Swinburn B, Egger G. Preventive strategies against weight gain and obesity. Obes Rev 2002;3:289–301.
- 36 Centraal Bureau voor de Statistiek. Beweeggedrag jeugd. 2022. Available: https://opendata.cbs.nl/
- 37 Lai SK, Costigan SA, Morgan PJ, et al. Do school-based interventions focusing on physical activity, fitness, or fundamental movement skill competency produce a sustained impact in these outcomes in children and adolescents? A systematic review of follow-up studies. Sports Med 2014;44:67–79.
- 38 Fühner T, Kliegl R, Arntz F, et al. An Update on Secular Trends in Physical Fitness of Children and Adolescents from 1972 to 2015: A Systematic Review. Sports Med 2021;51:303–20.
- 39 Paulhus D. Socially desirable responding on self-reports. 2017:1–5.
- 40 Khare SR, Vedel I. Recall bias and reduction measures: an example in primary health care service utilization. *Fam Pract* 2019;36:672–6.

Protected by copyright, including for uses related to text and data mining,

Al training, and similar technologies