**SUPPLEMENTAL INFORMATION**

**Sample size calculation**

The original aim of the PANIC Study was to investigate the effects of combined dietary and physical activity intervention on metabolic syndrome and insulin resistance. Therefore, our sample size calculations were based the effects of a dietary intervention on fasting serum insulin and HOMA-IR among children in the Special Turku Coronary Risk Factor Intervention Project (STRIP). Because of a larger number of children in our study than in the STRIP study, we approximated a slightly smaller difference for the change in fasting serum insulin and HOMA-IR of 0.3 SD between the intervention group (60% of children) and the control group (40% of children) with a power of 80% and a two-tailed p-value for the difference between the groups of 0.05, allowing for a 20% loss to follow-up or missing data. However, these calculations did not allow for non-independence within schools, and therefore the power could be lower than 80%. According to our calculations, we resulted in a sample size of at least 275 children in the intervention group and at least 183 children in the control group at baseline.

**Allocation of children to the intervention and the control groups**

We allocated the children from nine schools to a combined physical activity and dietary intervention group (306 children, 60%) and the children from seven schools to a control group (198 children, 40%) to avoid contamination in the control group by any local or national health promotion programmes that could have been initiated in the study region during the follow-up period1. We also proportionally matched the intervention and control group according to the size (small vs large) and the location of the schools (urban vs rural) to minimise sociodemographic differences between the groups. We included more children in the intervention group than in the control group because of a larger number of dropouts expected in the intervention group and to retain a sufficient statistical power for comparison between the groups. The children, their parents or caregivers, or people carrying out the examination visits or doing the measurements were not blinded to the group assignment.

**Assessment of physical activity and sedentary behavior**

Sedentary time (ST), light physical activity (LPA), moderate PA (MPA), and vigorous PA (VPA) were assessed at baseline and 2-year follow-up (Supplemental figure) using individually calibrated combined heart rate and body movement monitor (Actiheart, CamNtech Ltd, Papworth, UK)2,3 which was attached to the chest with standard electrocardiogram (ECG) electrodes (Bio Protech Inc, Wonju, South Korea). The monitor was set to record heart rate and body movement in 60‐second epochs. The participants were instructed to carry on with their usual behavior and to wear the monitor during all daily activities, including sleep, shower, sauna, and swimming. The activity patterns of school-aged children are known to vary markedly between weekdays and weekend days4. The participants were therefore requested to wear the monitor continuously for a minimum of four consecutive days, including two weekdays and two weekend days, to obtain more representative information on PA and ST. We accepted PA and ST data for the statistical analyses if there was a minimum of 48 hours of activity recording in weekday and weekend day hours that included at least 12 hours from morning (3 am‐9 am), noon (9 am‐3 pm), afternoon (3 pm‐9 pm), and night (9 pm‐3 am) to avoid potential bias from over‐representing specific times and activities of the days.

Heart rate data were cleaned. and individually calibrated with sleeping heart rate and parameters obtained from maximal exercise tests performed by the Ergoselect 200 K® electromagnetic bicycle ergometer (Ergoline, Bitz, Germany) and the Cardiosoft® V6.5 Diagnostic System ECG device (GE Healthcare Medical Systems, Freiburg, Germany)2. The heart rate data were finally combined with trunk acceleration data in a branched equation model to estimate activity intensity time‐series5. PA energy expenditure was calculated by integrating the intensity time‐series, where time distribution of activity intensity was generated by using standard metabolic equivalents (METs) in 0.5 increments. Sleep duration was analyzed from the Actiheart recordings by a trained exercise specialist and confirmed by a physician, if necessary. The time of falling asleep was defined as accelerometer counts decreasing to zero and heart rate to a plateau level. The time of waking up was defined as accelerometer counts increasing and remaining above zero and heart rate increasing and remaining above the plateau level. We defined total ST as time spent in activity ≤1.5 METs excluding sleep and LPA, MPA, and VPA as time spent in activity >1.5 and ≤4.0 METs, >4.0 and ≤7.0 METs, and >7.0 METs, respectively, by defining 1 MET as 71 J/kg−1/min−1 or oxygen uptake of 3.5 ml/kg−1/min−1. Moderate‐to‐vigorous PA (MVPA) included MPA and VPA.

**Assessment of body size and composition**

Body weight was measured twice with the children having fasted for 12 hours, emptied the bladder, and standing in light underwear using a weight scale integrated into a calibrated InBody® 720 bioelectrical impedance device (Biospace, Seoul, South Korea) to an accuracy of 0.1 kg. The mean of these two values was used in the analyses. Body height was measured three times with the children standing in the Frankfurt plane without shoes using a wall-mounted stadiometer to an accuracy of 0.1 cm. The mean of the nearest two values was used in the analyses. BMI was calculated by dividing weight (kg) by height (m) squared. The prevalence of overweight and obesity was defined using the cut-off values provided by Cole and co-workers6. Body fat percentage (BF%) was measured by the Lunar® DXA device (GE Medical Systems, Madison, WI, USA) using a standardized protocol7.

**Other assessments**

The parents were asked to report in a questionnaire their annual household income (categorized as ≤ 30 000 €, 30 001–60 000 €, and ≥ 60 001 €) and completed or ongoing educational degrees (categorized as vocational school or less, polytechnic, and university). The degree of the more educated parent was used in the analyses. A research physician assessed pubertal status according to breast development for girls (scored M 1–5) and according to testicular volume measured by an orchidometer for boys (scored G 1–5) using the staging method described by Tanner8,9.

A black background with a blue square

Description automatically generated

Supplemental Figure. Schematic illustration of the main time points of the study. Diet quality, physical activity, and cognition were assessed at baseline (0 months) and the 2-year follow-up (24 months). The dietary and physical activity counseling sessions occurred at 0.5, 1.5, 3, 6, 12, and 18 months after baseline.

**STATISTICAL METHODS**

Statistical analyses were performed using the IBM SPSS statistics for Windows, version 25.0 (IBM Corporation, Armonk, NY, USA). In all analyses, differences and associations with P-values <0.05 are considered statistically significant. Basic characteristics between boys and girls were compared using the Student´s t-test for continuous variables and the Chi-Square test for categorical variables.

The effects of the intervention on the RCPM score and the longitudinal associations of total BSDS and BSDS components, including red meat and sausages, fruit and berries, vegetables, high-fiber (≥5 %) grain products, fish, PUFA-SFA ratio, total fat, and milk (<1% fat), TPA, LPA, MVPA, VPA, ST, media time (screen time), reading time, writing time, and computer use time with the RCPM score were analyzed using linear mixed-effects models. We used the Bayesian information criterion (BIC) as a measure of model adequacy, a lower BIC value indicating a better model with a better balance between complexity and good fit. We predetermined the model with the lowest BIC as our final model for a given outcome. That is, we did not force the more complex data structure to our model if it did not improve the model fit. First, the clustering effect of individuals and schools were ruled out using the mixed-effects model analyses four times: one model without random factors, one model with only subject as a random factor, one model with only school as a random factor, and one model with both subject and school as random factors on different pages. The analyses showed no clustering effect of individuals and schools. These variables were therefore excluded from the models. We then analyzed the effect of the 2-year dietary and PA intervention on cognition using a linear mixed-effects model adjusted for age, sex, parental education, household income, and BF% at baseline and pubertal status at 2-year follow-up.

To analyze the longitudinal associations of total BSDS and BSDS components, including red meat and sausages, fruit and berries, vegetables, high-fiber (≥5%) grain products, fish, PUFA-SFA ratio, total fat, milk (<1% fat), TPA, LPA, MVPA, VPA, ST, media time (screen time), reading time, writing time, and computer use time with the RCPM score, the intervention and control groups were pooled and treated as one group in the analyses. The linear mixed-effects model analyses adjusted for age, sex, parental education, household income, and BF% at baseline and pubertal status at 2-year follow-up with maximum likelihood estimation were conducted according to a two-level structure, i.e., repeated (baseline and 2 years) measures of diet quality, PA, ST, or SB with the RCPM score. Next, we explored the within-subject and between-subject relationships of the variables demonstrating statistically significant associations with the RCPM score, which included BSDS, the consumption of red meat and sausages, ST, reading, writing, and computer use. The standard linear mixed-effects model pool together within-subject relationships (i.e whether a change in one variable over 2 years is associated with a change in another one) and between-subject relationships (i.e. whether an overall level of one variable over 2 years is associated with an overall level of another one) in such a way that no separation can be made between the two aspects of longitudinal relationships. Because of this limitation, we carried out additional analyses to sort out the within-subject and between-subject aspects of the relationships of BSDS, the consumption of red meat and sausages, ST, reading, writing, and computer use with the RCPM score by using a simple method described by van de Pol and Wright.10 By this method, subtracting the subject’s mean value from each observation value (i.e., within-subject centering) effectively eliminates any between-subject variation thus creating a new variable that expresses only the within-subject variation component. On the other hand, a second new variable expressing only the between-subject variation component is simply the mean of baseline and two-year observations (i.e., baseline and two-year observations for the same subject are both given the same value). Whenever the parameter estimates of these two effects seem to differ, it is possible to compare them to see whether they are statistically different from each other.

**References**

1. Lakka TA, Lintu N, Väistö J, et al. A 2 year physical activity and dietary intervention attenuates the increase in insulin resistance in a general population of children: the PANIC study. *Diabetologia*. 2020;63(11):2270-2281.

2. Collings PJ, Westgate K, Väistö J, et al. Cross-Sectional Associations of Objectively-Measured Physical Activity and Sedentary Time with Body Composition and Cardiorespiratory Fitness in Mid-Childhood: The PANIC Study. *Sports Med*. 2017;47(4):769-780.

3. Väistö J, Haapala EA, Viitasalo A, et al. Longitudinal associations of physical activity and sedentary time with cardiometabolic risk factors in children. *Scand J Med Sci Sports*. 2019;29(1):113-123.

4. Brooke HL, Corder K, Atkin AJ, van Sluijs EMF. A Systematic Literature Review with Meta-Analyses of Within- and Between-Day Differences in Objectively Measured Physical Activity in School-Aged Children. *Sports Med*. 2014;44(10):1427-1438.

5. Brage S, Ekelund U, Brage N, et al. Hierarchy of individual calibration levels for heart rate and accelerometry to measure physical activity. *J Appl Physiol*. 2007;103(2):682-692.

6. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*. 2000;320(7244):1240.

7. Tompuri TT, Lakka TA, Hakulinen M, et al. Assessment of body composition by dual-energy X-ray absorptiometry, bioimpedance analysis and anthropometrics in children: the Physical Activity and Nutrition in Children study. *Clinical Physiology and Functional Imaging*. 2015;35(1):21-33.

8. Marshall WA, Tanner JM. Variations in the Pattern of Pubertal Changes in Boys. *Arch Dis Child*. 1970;45(239):13-23.

9. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. *Arch Dis Child*. 1969;44(235):291-303.

10. van de Pol M, Wright J. A simple method for distinguishing within- versus between-subject effects using mixed models. *Animal Behaviour*. 2009;77(3):753-758.