Lerum, Ø., Aadland, E., Andersen, L. B., Anderssen, S. A., Resaland, G. K. (2017). Validity of noninvasive composite scores to assess cardiovascular risk in 10-year-old children. *Scandinavian Journal of Medicine & Science in Sports, 27*, 865-872.

Validity of noninvasive composite scores to assess cardiovascular risk in ten-year-old children

Øystein Lerum[a], Eivind Aadland[a], Lars B Andersen[a,b], Sigmund A Anderssen[a,b], Geir K Resaland[a]

[a] *Sogn og Fjordane University College, Faculty of Teacher Education and Sports, Sogndal, Norway.*
[b] *Norwegian School of Sport Sciences, Department of Sports Medicine, Oslo, Norway.*

Manuscript details:

Abstract: 234 words

Manuscript: 3684 words

Tables and figures: 3 tables and 1 figures

Corresponding author:

Øystein Lerum

Faculty of Teacher Education and Sports,

Sogn og Fjordane University College,

P.O. Box 133,

6851 Sogndal, Norway

E-mail: oystein.lerum@hisf.no

Office phone: +47 57676365

Cell phone: +47 91172244

Abstract

Agreement between and classification accuracy of six different noninvasive composite scores and a cardiovascular disease (CVD) risk factor score were investigated in 911 (466 boys and 445 girls) ten-year-old Norwegian children. A CVD risk factor score (triglyceride, total cholesterol/HDL ratio, homeostasis model assessment of insulin resistance, systolic blood pressure (SBP), waist-to-height ratio (WHtR) and cardiorespiratory fitness) and six noninvasive risk scores (fitness + three different measurements of fatness (body mass index (BMI), WHtR, skinfolds), with and without inclusion of SBP) were calculated (mean z-score by gender). Agreement was assessed using Bland–Altman plots. The ability of noninvasive scores to correctly classify children with clustered CVD risk was examined by Receiver Operating Characteristic (ROC) analysis and Cohen's kappa coefficient (k). For both sexes the noninvasive scores without SBP showed excellent AUC values (AUC = 0.93–0.94, 95% CI = 0.88–0.98), moderate kappa values (k = 0.49–0.64), and had limits of agreement of 0.0 ± 0.78–0.89 (arbitrary unit). Inclusion of SBP increased AUC values (AUC = 0.96–0.97, 95% CI = 0.94–0.99), kappa values (k = 0.58–0.69), and reduced limits of agreement (0.0 ± 0.68–0.76). Noninvasive scores that include fitness and fatness provide acceptable agreement and classification accuracy, allowing for widespread early identification of children that might be at risk for developing CVD later in life. SBP should be included in the noninvasive score to improve classification accuracy if possible.

Keywords: cardiorespiratory fitness, anthropometric measures, cardiovascular disease, systolic blood pressure, noninvasive score

Introduction

Cardiovascular disease (CVD) is the leading cause of premature death worldwide (Lozano et al., 2013). CVD risk factors have their roots in childhood and are caused by interplay between genetic and behavioral risk factors (Berenson et al., 1998; Camhi & Katzmarzyk 2010; Morrison et al., 2008; Raitakari et al., 2003). Independent associations between physical activity, fitness and fatness, and clustering of CVD risk factors in children have been established (Andersen et al., 2008; Eisenmann et al., 2007; Rizzo et al., 2007; Steene-Johannessen et al., 2009). Since 2000, great attention to identify those at risk has been devoted to find valid and practical measurement methods for use in various contexts such as research and clinical settings.

Clustering of CVD risk factors, typically components of the metabolic syndrome, for example, overweight, hypertension, insulin resistance, elevated triglycerides (TG) and lowered high-density lipoprotein cholesterol (HDL), has been suggested to reflect cardiovascular health in children better than single risk factors do (Andersen et al., 2008). A composite risk score may be a more solid indicator of the constellation of disturbances associated with CVD than single risk factors (Bailey et al., 2012). However, the assessment required to construct these composite risk scores is labor intensive and data are costly to obtain. Moreover, these measurements include invasive techniques, placing a significant burden upon some children. Thus, the feasibility of applying such scores for assessing risk in children on a large scale can be limited.

Epidemiological studies have evaluated simple anthropometric measurements (i.e., body mass index (BMI), skinfold, waist circumference) and field-based indirect cardiorespiratory fitness (CRF) tests as indexes for clustering of CVD risk factors in children and adolescents (Gonçalves et al., 2015; Sardinha et al., 2016). These findings are equivocal and single risk factors' ability to classify high-risk children and adolescents varies (Eisenmann 2007). Recently, Andersen et al., (2015) tested a noninvasive measurement of cardiovascular health in a study using pooling of a number of cohorts from Northern, Eastern, Central and

Southern Europe and North America (n = 9871) that included youths aged 6–18. The mean z-score of a composite noninvasive risk score based on fatness (waist-to-height ratio [WHtR]) and CRF was examined in a Receiver Operating Characteristic (ROC) analysis against the International Diabetes Federation (IDF) definition of the metabolic syndrome and a clustered CVD risk score including TG, waist circumference, systolic blood pressure (SBP), homeostasis model assessment (HOMA), HDL, and CRF. Andersen et al., (2015) observed areas under the curve (AUC) of 0.92 and 0.94, and sensitivity and specificity of 0.85 and 0.87, respectively, suggesting that the noninvasive score may be useful as a prescreening tool for identifying youths with clustered CVD risk. However, the study by Andersen et al., (2015) has some limitations; only one measure of fatness (WHtR) was examined in the composite noninvasive risk score, although SBP data were available, this simple noninvasive risk variable was not included, differences by sex was not investigated. Furthermore, to provide justification for the set of inferences intended to be drawn from the purposed noninvasive score, agreement for such a score across all levels of risk should be tested (e.g., by including Bland–Altman plots) (Bland & Altman, 1986; 1999).

Therefore, the present investigation's purpose was to evaluate the validity of six noninvasive composite scores (fitness + three different measures of fatness, with and without inclusion of SBP) in assessing CVD risk in ten-year-old children 1) by determining the agreement between different noninvasive risk scores and a comprehensive CVD risk score, and 2) by examining these risk scores' ability to correctly classify children with elevated risk.

Materials and methods

Setting and participants

Baseline data were obtained from the Active Smarter Kids (ASK) Study, a study that have previously been presented in detail (Resaland et al., 2015). A total of 1129 (588 boys and 541 girls) ten-year-old healthy pre-pubertal children were included. Before any testing was performed, written informed consent was obtained from each child's parents or legal guardians after they were given a detailed oral and written explanation of the study. Assent was obtained from the children. The study was approved by the Regional Committee for Medical Research Ethics (ID number 2013/1893). The ASK Study registration number at Clinicaltrails.gov ID number is NCT02132494.

Anthropometric assessment

Weight was measured (when children were in light clothing) to the nearest 0.1 kg with an electronic scale (Seca 899, SECA GmbH, Hamburg, Germany). Height was measured (when children were in their stockinged feet) to the nearest 1 mm with a transportable stadiometer (Seca 217, SECA GmbH, Hamburg, Germany). BMI (kg $\cdot$ m^{-2}) was calculated as weight (kg) divided by the height squared (m^2). Waist circumference 2 cm over the level of the umbilicus was measured to the nearest 0.5 cm with the child's abdomen relaxed at the end of a gentle expiration, using an ergonomic circumference measuring tape (Seca 201, SECA GmbH, Hamburg, Germany). Two measurements from each child were collected. If the difference between measurements was greater than 1 cm, a third measurement was obtained; the average of the two closest measurements was used for analysis. Skinfold thickness was measured at the left side of the body using a Harpenden skinfold caliper (Bull: British Indicators Ltd., West Sussex, England). Two measurements were taken at each position (biceps, triceps, subscapular, and suprailiac). If the difference between measurements was greater than 2 mm, a third measurement was obtained; the average of the two closest measurements was used for analysis.

Pubertal status

The Tanner pubertal stages self-assessment questionnaire was used to determine pubertal status (Tanner 1962). Boys were presented with five pictures of Tanner staging for pubic hair and external genitalia development, whereas girls were presented with five pictures representing breast development and pubic hair using color images proposed by Carel and Leger (Carel & Leger 2008). The children were asked to indicate which stage best referred to their own pubertal stage. The procedure took place in a private space with sufficient time to self-assess the pubertal stage. For analysis, children were classified as pre-pubertal (Tanner 1–2) or pubertal (3–5) by pubic hair and external genitalia development for boys and by breast development for girls.

Blood pressure

SBP and diastolic blood pressure (DBP) were measured using the Omron HEM-907 automatic blood pressure monitor (Omron HEM-907, Omron Healthcare, Inc., Veron Hills IL, USA). The children sat in a relaxed position in a quiet environment without distractions for ten minutes. Four measurements were taken at one-minute intervals on the upper right arm using an appropriately sized cuff, with the average of the final three measurements within 5 mmHg used in all analyses. If a difference > 5 mmHg between measurements was found, we obtained one extra measurement, in which case the average of the last four measurements was used.

Blood sample

Blood samples were obtained between 08.00 and 11.00 a.m. after an overnight fast and were immediately stored in a -20ºC freezer and within 48 hours in a -80ºC freezer until analysis. All blood samples was analyzed at the Endocrine Laboratory of the VU University Medical Center (VUmc) (Amsterdam, Netherlands). This laboratory is subject to external quality assessment and has the CCKL accreditation. Low-density lipoprotein cholesterol (LDL) was estimated from total cholesterol (TC), HDL, and TG by the Friedewald formula (Friedewald, Levy & Fredrickson, 1972). Insulin was measured using an automated

immunoassay (Centaur, Siemens Diagnostics). Insulin resistance was estimated according to HOMA as the product of fasting glucose (mmol/L) and insulin (pmol/L) divided by the constant 22.5 (Matthews et al., 1985). To calculate cholesterol ratio, HDL was divided by total cholesterol.

Cardiorespiratory fitness

CRF was assessed using the Andersen intermittent-running field test (Andersen et al., 2008), which has been validated in the target age group (Aadland et al., 2014). The Andersen test was administered according to standard procedures. The children were tested indoors on a wooden or rubber floor in groups of 10–20 children. Children ran from one end line to another (20 m apart) in an intermittent to-and-fro movement, with 15-second work periods and 15-second breaks (standing still) for a total duration of 10 minutes. The total distance covered in meters was used in all analyses. The equation suggested by Aadland et al., (2014) (VO_{2peak} = 23.262 + 0.050*Andersen distance -3.858*gender -0.376*body weight) was used to estimate VO_{2peak}.

Clustering of cardiovascular risk factors

Z-scores by gender were computed for all risk factors. BMI, WHtR, CRF, sum of four skinfold measurements, TC, insulin, and HOMA scores were positively skewed and were thus transformed (natural log) before z-scores were computed. Seven scores were computed. First, we computed a comprehensive score (the criterion to which the noninvasive scores were compared) as the mean of the following six CVD risk factors: 1) TG, 2) TC/HDL ratio, 3) HOMA score, 4) SBP, 5) WHtR, and 6) inverse CRF ("comprehensive score"). In addition, we computed three noninvasive scores: 1a) inverse CRF and WHtR, 2a) inverse CRF and BMI, and 3a) inverse CRF and sum of four skinfold measurements, and the three scores with SBP incorporated (1b, 2b, and 3b) ("noninvasive scores").

Statistical analysis

All statistical analyses were performed in IBM SPSS Statistics version 23.0 (SPSS Inc., Chicago IL, USA). Variables are described as group means and standard deviations (SDs) if normally distributed. Median and interquartile ranges were calculated for skewed data. Differences between genders and between included and excluded data were analyzed using the student's *t*-test for independent samples for normally distributed data, whereas Mann–Whitney U tests were performed for skewed data. Differences between genders in puberty status and clustered CVD risk were analyzed using the Pearson's chi-square test.

To evaluate the extent of clustering of risk factors in our sample, we compared the observed number of children who had zero to six risk factors (arbitrarily defined as the least favorable quartile) with the expected number calculated from an assumed independent distribution of the risk factors according to the binominal formula (Altman 1991),

$$(n!*p^{r}*(1 - p)^{n - r})/(r!*(n - r)!),$$

where *n* is the possible number of risk factors (6), *p* is the probability of having a risk factor (0.25), and *r* is the number of the risk factors for which the probability is calculated (0 through 6). The expected proportions having 0 to 6 risk factors were 0.178, 0.356, 0.296, 0.132, 0.033, 0.004, and 0.002, respectively. Corresponding observed proportions were 0.332, 0.285, 0.175, 0.091, 0.060, 0.037, and 0.020.

Validity of the noninvasive scores was assessed using the Pearson's correlation coefficient (*r*), linear regression, Bland Altman plots, ROC analyses and Cohen's kappa coefficient (*k*) through the following six steps: 1) We initially randomly split our sample in two to construct one training dataset (n = 456) and one test dataset (n = 455). 2) Then, we used linear regression to make equations to predict the comprehensive risk score based on the non-invasive scores (6 in boys and 6 in girls) in the training dataset. 3) These equations were then used to predict the comprehensive risk score in the test datasets for boys (n = 231) and girls (n = 224) separately, using linear regression. 4) Agreement for the continuous scores

8

(predicted vs. measured comprehensive scores) were investigated using Bland Altman plots (Bland & Altman 1986) with 95% limits of agreement (LoA) (± 1.96 SD of the differences). Means were compared using t-test. 5) The predicted noninvasive scores were further examined in ROC analyses to examine their ability to correctly classify children with clustered cardiovascular risk. The applied cutoff point in the comprehensive score was computed from the normal distribution curve of the respective proportion of children having ≥4 risk factors in the training dataset (10.3%) (z = x − mean / standard deviation). The estimated cutoff point was 0.77. The AUC with 95% confidence intervals (CI) are reported. The AUC represents the ability of the test to correctly classify children having elevated CVD risk. The values of AUC range from 0.5 (no classification ability beyond chance) to 1 (perfect agreement). 6) Finally, we calculated the Cohen's kappa coefficient (k), ranging from 0 (no classification ability beyond chance) to 1 (perfect agreement). Agreement was interpreted as moderate for $k = 0.41 − 0.60$, and good for $k = 0.61 − 0.80$ (Altman, 1991).

Results

A complete data set for all six CVD risk factors was available for 911 children (466 boys and 445 girls). Thus, 218 children were excluded in the present analysis. Children with incomplete data had significantly lower Andersen test results (866 vs. 897 m, P < 0.001) than children with complete data, although no other significant differences were found. Characteristics of the included children are shown in Table 1; girls had higher mean levels of the sum of four skinfold measurements, fasting insulin, TG, and HOMA score (P < 0.001), and lower Andersen test distance (P < 0.001), $VO_{2\,peak}$ (P < 0.001), glucose (P < 0.001), and HDL (P < 0.001) than boys. An excess number of children were identified with clustering of risk factors. In all, 11.7% of the population (n = 107) had ≥ four risk factors (odds ratio 3.1, 95% CI = 2.5–3.8 compared to the expected number). The corresponding proportion in the training and test datasets was 10.3% (odds ratio 2.7, 95% CI = 2.0–3.7 compared to the expected number) and 13.2% (odds ratio 3.5, 95% CI = 2.6–4.6 compared to the expected number), respectively.

All results were virtually equal for boys and girls. Table 2 shows the bivariate relationship between the six predicted comprehensive scores and the measured comprehensive score. All predicted scores showed strong associations with the comprehensive score for both sexes ($r = > 0.75$, all $P < 0.001$), and were almost identical with the values obtained in the training dataset (supplementary table 1). Stronger associations ($r = 0.82–0.86$, all $P < 0.001$) were observed for noninvasive scores that included SBP.

The Bland–Altman plots show the agreement between the noninvasive scores and the comprehensive score in terms of systematic error (bias) and random error (95% LoA), displayed in Figure 1 (noninvasive scores 1a and 1b shown). LoA (arbitrary unit) were smaller for all noninvasive scores that included SBP (boys: 1b, 2b, 3b: LoA ± 0.68–0.71; girls: 1b, 2b, 3b: LoA ± 0.68–0.76) than they were for all noninvasive scores that did not include SBP (boys: 1a, 2a, 3a: LoA ± 0.78–0.85; girls: 1a, 2a, 3a: LoA ± 0.80–0.89).

Table 3 shows the AUC as well as sensitivity and specificity for the predictions made by the six noninvasive scores (ROC curves are given as supplementary figure 1). The classification accuracy of noninvasive scores to identify clustered CVD risk in boys and girls was excellent for all scores (AUC > 0.93). Sensitivity was moderate and specificity was high for the threshold analyzed. In general, noninvasive scores without SBP had moderate agreement according to the Kappa coefficient ($k = 0.49–0.64$), whereas scores including SBP had good agreement ($k = 0.58–0.69$).

Discussion

This study examined the validity of six noninvasive composite risk scores to assess CVD risk in a large group of apparently healthy ten-year-old children from one county in western Norway. The good agreement and high discriminant ability with the criterion measure, especially for scores that include SBP, demonstrate that the noninvasive composite risk scores have good utility for identifying high-risk children. Accuracy for the noninvasive scores

of boys was similar to that of girls, suggesting that relatively simple and feasible measurements can be used to identify children at elevated risk, irrespective of sex.

Results from the present study are in good agreement with results from the study by Andersen et al., (2015), who found an AUC of 0.94 for their composite noninvasive risk score analyzed against the mean z-score of a more comprehensive (TG, waist circumference, SBP, HOMA, HDL and CRF) CVD risk score. Similarly, our results showed that noninvasive score 1a (CRF and WHtR) provided an AUC of 0.93 for boys, and an AUC of 0.94 for girls. We extend these results by demonstrating good agreement and classification accuracy in both girls and boys, and for different measurements of fatness. The lower sensitivity and the very high specificity compared to Andersen et al (2015), results from a slightly lower proportion of children at risk in the training dataset than in the test dataset (10.3 vs. 13.2% having clustering of ≥ 4 risk factors, respectively), which is a direct result of performing a cross-validation within a sample being limited in size. This led to a somewhat high threshold to classify children at risk in the test dataset, resulting in a somewhat increased false negative rate, but a very low false positive rate. Nevertheless, the cut point applied in the present study is sample-specific and calculated for the purpose of the present analyses. We recommend the cut point suggested by Andersen et al (2015) to be applied in future studies. Kappa coefficients seem to contradict the excellent AUC values obtained from the ROC analyses, as the kappa values indicated only moderate to good agreement. Worth noting, however, is that the kappa coefficient is vulnerable to class skew, that is different proportions classified as positive and negative cases (in our study ~10 to 90%), which will clearly attenuate the score (Altman, 1991), despite an overall high classification accuracy (90-94% of cases correctly classified from the noninvasive scores).

Previous investigations have identified CRF as a strong predictor of CVD risk factor clustering in children and adolescents with AUC values from 0.67–0.84 (Ruiz et al., 2007; Sasayama et al., 2015; Welk et al., 2011). Similarly, anthropometric measures also seem to be valuable predictors of CVD risk factor clustering with AUC values from 0.62–0.80 (Goncalves et al., 2015; Maffeis et al., 2008; Sardinha et al., 2016). Compared to these previous studies investigating the validity of single variables as predictors for CVD clustering in children and adolescents (Gonçalves et al., 2015; Maffeis et al., 2008; Ruiz et al., 2007; Sardinha et al., 2016; Sasayama et al., 2015; Welk et al., 2011), the present study indicates that using a "noninvasive mini-cluster" of fitness and fatness is superior to using each of these variables alone. Thus, our results extend these previous observations with even higher AUC values and by suggesting that the strength of associations between WHtR, BMI, and skinfold with clustered CVD risk is similar for all anthropometric variables; hence, they are interchangeable. Thus, we argue that it is a matter of practical convenience to choose one measure over another. However, we still recommend the use of waist circumference or BMI, because skinfold measurement might be difficult to standardize (Castro–Piñero et al., 2009). BMI is the most common measure of fatness because it is easy to measure, which makes it a practical measure in many contexts, while also comparable across studies. Still, WHtR has been advocated as an effective and convenient measure of central adiposity that could potentially be superior to BMI alone in determining CVD risk (Khoury et al., 2013), and simultaneously as a variable independent of age (Savva et al., 2000), which may favor WHtR over BMI.

Data from the present study showed that a simple risk score consisting of fitness and fatness could discriminate high-risk children with high accuracy (AUC = 0.93–0.94) and is practically superior to an advanced CVD risk score. Adding SBP to the noninvasive scores further improved the scores, naturally, since three of the variables are identical in the noninvasive variables and the criterion.

Strengths and weaknesses

The major strength of this study is that it measures blood lipids and other cardiovascular risk factors from a relatively large sample with a high participation rate (97%), of which 89% provided valid blood samples.

The present study has limitations. Despite the large sample size, the age range included is narrow (10.2 ± 0.3), and therefore the results may have limited generalizability to other age groups. Thus, further investigation should include different populations and wider age range. Still, results in the present study were similar to those of Andersen et al., (2015), which were obtained from an international sample of children and youth. Unfortunately, since previous findings imply that the metabolic syndrome score differs by ethnicity and gender (Gurka et al., 2012), differences across age groups, gender, and ethnicity were not examined by Andersen et al., (2015). However, the AUC observed indicate a high discriminatory ability for identifying children likely to have clustering of CVD risk factors.

Further, composite risk scores have some limitations. It is sample-specific and highly dependent on the sample of children studied. Furthermore, each selected variable is equally weighted within the score, despite different variables might pose different risk of disease. Finally, although a child is being defined at elevated risk does not necessarily mean that the child will develop CVD later in life.

Only two children in this study were defined as having the metabolic syndrome according to the IDF definition. Thus, given our results, we agree with previous calls to use a continuous rather than a dichotomous approach to avoid underestimation when evaluating children's cardiovascular health (Eisenmann 2008).

Perspective

Together with previous findings, our findings suggest that although anthropometric variables and CRF seem to be strong predictors of CVD risk factors. Together with the findings shown

by Andersen et al., (2015), our findings add novel and important information to the field of pediatric epidemiology.

Conclusion

Given the current findings, different noninvasive scores consisting of measurements of fitness and fatness provide acceptable agreement and classification accuracy of children at metabolic risk, allowing for widespread early identification of children who might develop CVD later in life. SBP should be included in the score to improve classification accuracy if possible. Importantly, our findings indicate that it is possible to identify children at CVD risk reasonably well by relying only on noninvasive measures of risk, an approach that allows for reduced distress and possibly for an increased study participation rate among children.

Acknowledgement

References

Aadland E, Terum T, Mamen A, Andersen LB, Resaland GK. The Andersen aerobic fitness test: reliability and validity in 10-year-old children. PLoS One. 2014: 9: e110492.

Altman D. Theoretical distributions. Practical statistics for medical research. London: Chapman & Hall. 1991: 68–70.

Andersen LB, Andersen T-E, Andersen E, Anderssen SA. An intermittent running test to estimate maximal oxygen uptake: the Andersen test. The Journal of Sports Medicine and Physical Fitness. 2008: 48: 434–437.

Andersen LB, Lauersen JB, Brond JC, Anderssen SA, Sardinha LB, Steene-Johannessen J, McMurray RG, Barros MV, Kriemler S, Moller NC, Bugge A, Kristensen PL, Ried-Larsen M, Grontved A, Ekelund U. A new approach to define and diagnose cardiometabolic disorder in children. J Diabetes Res. 2015: 2015: 539835.

Andersen LB, Sardinha LB, Froberg K, Riddoch CJ, Page AS, Anderssen SA. Fitness, fatness and clustering of cardiovascular risk factors in children from Denmark, Estonia and Portugal: the European Youth Heart Study. International journal of pediatric obesity: IJPO: an official journal of the International Association for the Study of Obesity. 2008: 3 Suppl 1: 58–66.

Bailey DP, Boddy LM, Savory LA, Denton SJ, Kerr CJ. Associations between cardiorespiratory fitness, physical activity and clustered cardiometabolic risk in children and adolescents: the HAPPY study. European Journal of Pediatrics. 2012: 171: 1317–1323.

Berenson GS, Srinivasan SR, Bao W, Newman WP, Tracy RE, Wattigney WA. Association between Multiple Cardiovascular Risk Factors and Atherosclerosis in Children and Young Adults. New England Journal of Medicine. 1998: 338: 1650–1656.

Bland JM & Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986: 327: 307–310.

Bland JM & Altman D. Measuring agreement in method comparison studies. Statistical methods in medical research. 1999: 8(2): 135-160.

Camhi SM, Katzmarzyk PT. Tracking of cardiometabolic risk factor clustering from childhood to adulthood. International Journal of Pediatric Obesity. 2010: 5: 122–129.

Carel JC, Leger J. Precocious puberty. New England Journal of Medicine. 2008;358(22): 2366–77.

Castro-Piñero, J., Artero, E. G., España-Romero, V., Ortega, F. B., Sjöström, M., Suni, J., & Ruiz, J. R. Criterion-related validity of field-based fitness tests in youth: a systematic review. British Journal of Sports Medicine. 2009.

Eisenmann JC. Aerobic fitness, fatness and the metabolic syndrome in children and adolescents. Acta Paediatr. 2007: 96: 1723–29.

Eisenmann JC. On the use of a continuous metabolic syndrome score in pediatric research. Cardiovasc Diabetol. 2008: 7: 17.

Eisenmann JC, Welk GJ, Ihmels M, Dollman J. Fatness, Fitness, and Cardiovascular Disease Risk Factors in Children and Adolescents. Medicine & Science in Sports & Exercise. 2007: 39: 1251–1256.

Friedewald WT, Levy RI & Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholestrol in plasma, without use of the preparative ultracentrifuge. Clin Chem. 1972: 18(6): 499–502.

Gonçalves R, Szmuchrowski L, Prado L, Couto B, Machado J, Damasceno V, Lamounier J. Selected anthropometric variables and aerobic fitness as predictors of cardiovascular disease risk in children. Biology of Sport. 2015: 32: 255.

Gurka MJ, Ice CL, Sun SS, DeBoer MD. A confirmatory factor analysis of the metabolic syndrome in adolescents: an examination of sex and racial/ethnic differences. Cardiovascular Diabetology. 2012: 11: 128.

Khoury M, Manlhiot C, McCrindle BW. Role of the waist/height ratio in the cardiometabolic risk assessment of children classified by body mass index. Journal of the American College of Cardiology. 2013: 62: 742–751.

Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet. 2013: 380: 2095–2128.

Maffeis C, Banzato C, Talamini G, of the Italian OSG. Waist-to-height ratio, a useful index to identify high metabolic risk in overweight children. The Journal of Pediatrics. 2008: 152: 207–213. e202.

Matthews D, Hosker J, Rudenski A, Naylor B, Treacher D, Turner R. Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia. 1985: 28: 412–419.

Morrison JA, Friedman LA, Wang P, Glueck CJ. Metabolic syndrome in childhood predicts adult metabolic syndrome and type 2 diabetes mellitus 25 to 30 years later. The Journal of Pediatrics. 2008: 152: 201–206.

Raitakari OT, Juonala M, Kähönen M, Taittonen L, Laitinen T, Mäki-Torkko N, Järvisalo MJ, Uhari M, Jokinen E, Rönnemaa T. Cardiovascular risk factors in childhood and carotid artery intima-media thickness in adulthood: the Cardiovascular Risk in Young Finns Study. JAMA. 2003: 290: 2277–2283.

Resaland GK, Moe VF, Aadland E, Steene-Johannessen J, Glosvik Ø, Andersen JR, Kvalheim OM, McKay HA, Anderssen SA. Active Smarter Kids (ASK): Rationale and design of a cluster-randomized controlled trial investigating the effects of daily physical activity on children's academic performance and risk factors for non-communicable diseases. BMC Public Health. 2015: 15.

Rizzo NS, Ruiz JR, Hurtig-Wennlof A, Ortega FB, Sjostrom M. Relationship of physical activity, fitness, and fatness with clustered metabolic risk in children and adolescents: the European youth heart study. J Pediatr. 2007: 150: 388–394.

Ruiz JR, Ortega FB, Rizzo NS, Villa I, Hurtig-Wennlof A, Oja L, Sjostrom M. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. Pediatr Res. 2007: 61: 350–355.

Sardinha LB, Santos DA, Silva AM, Grøntved A, Andersen LB, Ekelund U. A Comparison between BMI, Waist Circumference, and Waist-To-Height Ratio for Identifying Cardio-Metabolic Risk in Children and Adolescents. PloS One. 2016: 11: e0149351.

Sasayama K, Ochi E, Adachi M. Importance of Both Fatness and Aerobic Fitness on Metabolic Syndrome Risk in Japanese Children. PloS One. 2015: 10(5): e0127400.

Savva S, Tornaritis M, Savva M, Kourides Y, Panagi A, Silikiotou N, Georgiou C, Kafatos A. Waist circumference and waist-to-height ratio are better predictors of cardiovascular disease risk factors in children than body mass index. International journal of obesity and related metabolic disorders: Journal of the International Association for the Study of Obesity. 2000: 24: 1453–1458.

Steene-Johannessen J, Anderssen SA, Kolle E, Andersen LB. Low muscle fitness is associated with metabolic risk in youth. Medicine and Science in Sports and Exercise. 2009: 41: 1361–1367.

Tanner JM. Growth at adolescence. Oxford, United Kingdom: Blackwell; 1962.

Welk GJ, Laurson KR, Eisenmann JC, Cureton KJ. Development of youth aerobic-capacity standards using receiver operating characteristic curves. American Journal of Preventive Medicine. 2011: 41: S111–S116.

Tables

Table 1. *Physical characteristics and CVD risk factors in the children*			
	Boys (n = 466)	**Girls (n = 445)**	**p-value**
Age (years)	10.2 (0.3)	10.2 (0.3)	0.852
Height (cm)	143.1 (6.7)	142.5 (6.7)	0.152
Weight (kg)[†]	35.5 (31.6–40.9)	35.3 (31.5–40.9)	0.983
BMI (kg/m^2)[†]	17.2 (15.8–19.4)	17.3 (15.9–19.5)	0.338
Skinfold (mm)[†]	35.0 (26.8–53.7)	51.0 (36.1–72.7)	< 0.001*
Waist Circumference (cm)[†]	60.8 (57.3–65.8)	59.5 (56.0–65.2)	0.017
Prepubertal (Tanner stage 1–2)[‡]	89.7 (418)	88.3 (393)	0.504
Systolic BP (mmHg)	105.3 (8.2)	105.3 (8.5)	0.970
Diastolic BP (mmHg)	57.3 (6.0)	58.1 (6.3)	0.057
Andersen (m)[†]	931 (860–1000)	875 (815–926)	< 0.001*
VO$_{2peak}$ (ml/min^{-1}/kg^{-1})[†]	56.6 (51.5–60.6)	49.8 (45.8–52.9)	< 0.001*
Glucose (mmol/L)	5.02 (0.31)	4.90 (0.33)	< 0.001*
Insulin (pmol/L)	50.04 (24.87)	60.87 (33.52)	< 0.001*
Cholestrol (mmol/L)	4.45 (0.72)	4.45 (0.67)	0.897
HDL (mmol/L)	1.62 (0.33)	1.55 (0.34)	< 0.001*
LDL (mmol/L)	2.49 (0.67)	2.52 (0.62)	0.453
Triglycerid (mmol/L)[†]	0.65 (0.52–0.83)	0.73 (0.57–0.96)	< 0.001*
Insulin resistance (HOMA score)[†]	10.04 (7.16–13.77)	11.60 (8.19–16.89)	< 0.001*
Clustered CVD risk (> 4 risk factors)[‡]	9.0 (42)	14.6 (65)	< 0.009*
Values are presented as mean (standard deviation). BMI = body mass index, BP = blood pressure, CRF = cardiorespiratory fitness, HDL = high-density lipoprotein, LDL = low-density lipoprotein. *Significant p-value at < 0.05 level. †Median (25 – 75 inter quartile). ‡Percent (n).			

Table 2. *Pearson correlation coefficients (r) between the noninvasive scores and the comprehensive CVD risk factor score.*

	Comprehensive score	
	Boys (n = 231)	Girls (n = 224)
	r*	*r**
Noninvasive score 1		
a) CRF and WHtR	0.77	0.78
b) CRF, WHtR and SBP	0.84	0.86
Noninvasive score 2		
a) CRF and BMI	0.77	0.79
b) CRF, BMI and SBP	0.83	0.84
Noninvasive score 3		
a) CRF and skinfold	0.75	0.76
b) CRF, skinfold and SBP	0.82	0.84

Values are presented as Pearson's correlation coefficient *r*. CRF = cardiorespiratory fitness (the Andersen test), WHtR = waist-to-height ratio, SBP = systolic blood pressure, BMI = body mass index.
*All significant p-values at < 0.001 level.

Table 3. *Classification accuracy in boys and girls for noninvasive scores to predict clustering of CVD risk factors. Sensitivity and specificity for a threshold point of 0.77.*

Noninvasive score	AUC (95% CI)	Sensitivity (%)	Specificity (%)	Kappa (*k*)	Correctly classified (%)
Boys					
1a) CRF and WHtR	0.93 (0.89–0.97)	46	97	0.53	92
1b) CRF, WHtR and SBP	0.97 (0.94–0.99)	65	98	0.69	94
2a) CRF and BMI	0.94 (0.91–0.97)	46	98	0.53	92
2b) CRF, BMI and SBP	0.97 (0.95–0.99)	62	98	0.65	94
3a) CRF and skinfold	0.94 (0.91–0.97)	42	98	0.51	92
3b) CRF, skinfold and SBP	0.97 (0.96–0.99)	65	98	0.69	94
Girls					
1a) CRF and WHtR	0.94 (0.91–0.98)	57	97	0.64	92
1b) CRF, WHtR and SBP	0.97 (0.95–0.99)	63	98	0.66	93
2a) CRF and BMI	0.94 (0.91–0.98)	50	96	0.52	90
2b) CRF, BMI and SBP	0.97 (0.94–0.99)	53	99	0.61	92
3a) CRF and skinfold	0.93 (0.88–0.97)	47	97	0.49	90
3b) CRF, skinfold and SBP	0.96 (0.94–0.99)	50	98	0.58	92

AUC = area under the curve, CI = 95% confidence interval, *K* = Cohen's kappa coefficient, CRF = cardiorespiratory fitness (the Andersen test), WHtR = waist-to-height ratio, SBP = systolic blood pressure, BMI = body mass index.

Figure legends

Figure 1. *Bland–Altman plot of noninvasive scores 1a (CRF and WHtR) and 1b (CRF, WHtR and SBP) for boys (A and C) and girls (B and D) compared with the comprehensive score. The full line represents the mean differences (bias) between noninvasive scores and the comprehensive score; the upper and lower short-dashed lines represent the upper and lower 95% limits of agreement (mean differences ± 1.96 SD of the differences).*

Figures

Figure 1