Aadland, E., Anderssen, S. A., Andersen, L. B., Resaland, G. K., Kolle, E., Steene-Johannessen, J. (2018). Aerobic fitness thresholds to define poor cardiometabolic health in children and youth. *Scandinavian Journal of Medicine & Science in Sports, 29*, s.240-250.

Aerobic fitness thresholds to define poor cardiometabolic health in children and youth

Eivind Aadland PhD,[1] Sigmund Alfred Anderssen PhD,[1,2] Lars Bo Andersen PhD,[1,2] Geir Kåre Resaland PhD,[1] Elin Kolle PhD,[2] Jostein Steene-Johannessen PhD[2]

[1]*Western Norway University of Applied Sciences, Faculty of Education, Arts and Sports, Department of Sport, Food and Natural Sciences, Campus Sogndal, Box 133, 6851 Sogndal, Norway*

[2]*Norwegian School of Sport Sciences, Department of Sports Medicine, Box 4014 Ullevål Stadion, 0806 Oslo, Norway*

Running head: Aerobic fitness thresholds in children/youth

Corresponding author

Eivind Aadland

Department of Sport, Food and Natural Sciences, Faculty of Faculty of Education, Arts and Sports, Western Norway University of Applied Sciences, Campus Sogndal, Box 133, 6851 Sogndal, Norway. Phone: +47 5767 6086; Email: eivind.aadland@hvl.no

Word count main text: 4190; word count abstract: 245

Abstract

Aerobic fitness is an apparent candidate for screening children and youth for poor cardiometabolic health and future risk of cardiovascular disease (CVD). Yet, age- and sex-specific cut points for children and youth determined using a maximal protocol and directly measured peak oxygen consumption (VO_{2peak}) does not exist. We used a nationally representative sample of 1462 Norwegian children and youth (788 boys and 674 girls aged 8.7–10.4 years and 14.7–16.7 years) who in 2005–2006 performed a maximal cycle ergometer test with direct measurement of VO_{2peak}, along with measurement of several other risk factors for CVD (systolic blood pressure, waist circumference:height ratio, total:high density lipoprotein cholesterol ratio, triglycerides, Homeostasis Model Assessment for Insulin Resistance). Based on the proportion of children having clustering (least favorable quartile) of 6 (1.6%), ≥ 5 (5.2%), and ≥ 4 (10.6%) CVD risk factors, we established the 2nd, 5th, and 10th percentile cut points for VO_{2peak} (ml/kg/min) for children and youth aged 8–18 years. Classification accuracy was determined using the Kappa coefficient (k), sensitivity and specificity. For boys, the 2nd, 5th, and 10th percentile VO_{2peak} cut points were 33.6–36.4, 36.3–39.8, and 38.7–43.0 ml/kg/min, respectively. For girls, the corresponding cut points were 29.7–29.1, 32.4–31.4, and 34.8–33.5 ml/kg/min, respectively. Together with BMI, but without more invasive measures of traditional risk factors for CVD, these cut points can be used to screen schoolchildren for poor cardiometabolic health with moderate discriminating ability (k ≤ 0.53).

Keywords: maximal oxygen consumption; cardiovascular disease; schoolchildren; screening; prevention

Introduction

There is consistent evidence that increased aerobic fitness relates to better cardiometabolic health in children and youth[1-3], and that this relation persists into adulthood [4 5]. Consequently, aerobic fitness is a candidate for screening and early intervention to prevent future metabolic disturbances and cardiovascular disease (CVD). Aerobic fitness [6], especially in combination with fatness [7 8], can be used to identify at-risk individuals, without requiring invasive measures of traditional CVD risk factors (blood samples). However, for measures of aerobic fitness to be used as an indicator of cardiometabolic risk and be decisive in terms of initiating preventive initiatives to increase physical activity, cut points to define children with poor cardiometabolic health are needed.

Ruiz et al [6] reviewed the evidence for VO_{2peak} (ml/kg/min) cut points established to indicate metabolic risk in children and youth aged 8 to 19 years in 2016. Based on seven included studies, they concluded that VO_{2peak} levels lower than 41.8 to 47.0 ml/kg/min in boys and lower than 34.6 to 39.5 ml/kg/min in girls indicated poor cardiometabolic health. Yet, this study has two major limitations. First, all seven studies included in the systematic review estimated VO_{2peak} based on indirect performance measures, some known to underestimate VO_{2peak} and some having unknown validity. For example, Mesa et al [9], Moreira et al [10], and Ruiz et al [11] used the 20 meter multistage shuttle run test (MSRT) to estimate VO_{2peak} applying the Leger [12] and Matsuzaka [13] equations, which are both shown to underestimate VO_{2peak} in children and adolescents [14 15]. Furthermore, we are not aware of any cross-validation studies of the submaximal treadmill protocol used to estimate VO_{2peak} in the studies based on the US National Health and Nutrition Examination Survey (NHANES) [16 17]. Second, the review by Ruiz et al [6] did not account for age, which imposes a clear limitation as it is well-known that VO_{2peak} per kg body mass increases with age during adolescence in boys, whereas it slightly decreases in girls [18]. Thus, accounting for age is critical to attain useful cut points. This is supported by the findings by Adegboye et al [19], the largest study included in the systematic review by Ruiz et al [6], showing that VO_{2peak} cut points indicating poor cardiometabolic health increased with age in boys (43.6 ml/kg/min in 9-year-olds and 46.0 ml/kg/min in 15-year-olds), but decreased with age in girls (37.4 ml/kg/min in 9-year-olds and 33.0 ml/kg/min in 15-year-olds). Similar findings were shown by Welk et al [17], which is the only study to have established cut points by sex and age between ages 10 and 18 years. These cut points have been adapted by FITNESSGRAM [20] but a limitation of the study by Welk et al [17] is reliance on a submaximal test of VO_{2peak}.

To our knowledge, aerobic fitness cut points established using a large sample of children performing maximal tests with direct measurement of VO_{2peak} does not exist besides the study by Adegboye et al [19], which included directly measured VO_{2peak} from the *Physical Activity among Norwegian Children*

Study (PANCS) [18 21], as a part of their study sample. However, Adegboye et al [19] did not present age-specific cut points beyond the age groups of 9- and 15-year-olds. Thus, the aim of this paper was to define age- and sex-specific cut points for VO_{2peak} (ml/kg/min) indicative of poor cardiometabolic health in children and youth aged 8 to 18 years old solely from the PANCS study.

Methods

Participants

We included a nationally representative sample of Norwegian children and adolescents aged 8.7 to 10.4 years ("9-year-olds") and 14.7 to 16.7 years ("15-year-olds") from PANCS [18 21], conducted during 2005–2006. Statistics Norway selected the cohort by cluster sampling with schools as the primary unit. When a school agreed to participate, we invited all children in grade 4 and grade 10. Thus, we invited a total of 2818 children and adolescents, of whom 2299 accepted and participated in the study. Of these, 1462 children (788 boys and 674 girls) provided data for the present analyses.

Procedures and methods conform to ethical guidelines defined by the World Medical Association's Declaration of Helsinki and its subsequent revisions. The Regional Committee for Medical and Health Research Ethics and Norwegian Social Sciences Data Services approved the study. We obtained written informed consent from each participant's parents or legal guardian prior to testing.

Procedures

We have previously published detailed descriptions of the PANCS studies [18 21] and therefore provide only a brief overview of the relevant procedures herein.

VO_{2peak} (ml/min/kg) was directly assessed using a portable MetaMax III X oxygen analyzer (Cortex Biophysics, Leipzig, Germany) during a progressive cycle test until volitional exhaustion using an electronically braked cycle ergometer (Ergomedic 839E; Monark, Varberg, Sweden). Initial and incremental work rates were 20 Watt (W) for 9-year-olds weighing <30 kg, 25 W for 9-year-olds weighing ≥30 kg, 40 W for 15-year-old girls and 50 W for 15-year-old boys. Work rate increased every third minute until exhaustion. We recorded heart rate throughout the test using a heart rate monitor (Polar Vantage, Finland) and VO_2, respiratory exchange ratio, and ventilation every 10 seconds during the last minutes of the test. We defined VO_{2peak} as the mean of the three highest consecutive measurements. Every morning we calibrated the analyzer against known gas mixtures, and barometric pressure against values from the local weather station. The cycle ergometer was electronically calibrated every morning and mechanically calibrated after being moved. The primary criterion of an acceptable test (maximal effort) was that the participants demonstrated clear signs of intense effort and clear symptoms of fatigue (e.g., facial flushing or difficulties in maintaining pedaling frequency). In addition, attainment of one objective criterion (heart rate ≥ 185 beats/minute or a respiratory exchange ratio ≥ 0.99) was needed to accept the test as valid. Body mass and height was measured using standardized procedures (Seca 770, SECA GmbH, Hamburg, Germany) with children wearing light clothing. Body mass index (BMI) ($kg \cdot m^{-2}$) was calculated, and individuals' BMI statuses classified according to the BMI criteria by Cole et al [22]. Waist circumference (WC) was measured with a metal anthropometric tape, taken midway between the lower rib and the iliac crest at the end of a normal expiration. After being seated for a minimum of 5 minutes, we measured systolic blood pressure (SBP) five times at 2-minute intervals using an automated blood pressure monitor (Omega™ Noninvasive blood pressure monitor; Invivo research, Inc., Orlando, FL). We used the mean of the last three measurements for analysis. Following an overnight fast, venous blood samples were collected between 08:00 and 10:00 a.m. The samples were analyzed for total cholesterol (TC), triglycerides (TG), high-density lipoprotein (HDL) cholesterol, and glucose at the Central Laboratory of Ullevaal University Hospital (Oslo, Norway) using Routine enzymatic colorimetric assays from Roche Diagnostics and performed on a Cobas Integra analyser (F. Hoffmann-La Roche Ltd, Basel, Switzerland). Insulin was analyzed at the Aker University Hospital (Oslo, Norway) by fluoroimmunoassay using an automatic immunoassay system (AutoDELFIAHInsulin; PerkinElmer, Turku, Finland). We calculated the TC:HDL ratio and Homeostasis Model Assessment for Insulin Resistance (HOMA) (glucose (mmol/L) * insulin (pmol/L) / 22.5) [23].

Statistical analyses

Participants' characteristics were reported as frequencies, means and standard deviations (SD), or medians and interquartile ranges (IQR) if variables were skewed. We tested for differences in characteristics between included and excluded individuals using an independent samples t-test (normally distributed variables) or a Mann-Whitney test (skewed variables), after correcting for age and sex (i.e., using residuals from a linear regression model).

Similar to a previous study [8], we used clustering of CVD risk factors as the criterion measure for defining children with poor cardiometabolic health. Clustering was defined as present in an individual if their measurements were within the least favourable quartile (at risk) for a greater than expected number of the six risk factors: SBP, TG, TC:HDL ratio, HOMA, WC:height ratio, and VO_{2peak}, given independence of the risk factors. The observed proportion that had clustering was compared to the expected proportion given independence among the risk factors using the binominal formula ($[n! * p^r * (1 - p)^{n-r}]/[r! * (n - r)!]$), 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 0.1780, 0.3560, 0.2966, 0.1318, 0.0330, 0.0044, and 0.0002, respectively. We then calculated the Odds Ratios (OR) and 95% confidence intervals (CI) for the observed vs. expected proportions having clustering of 3, 4, 5, and 6 risk factors as $exp(Ln[OR] \pm 1.96 * SE[Ln(OR)])$, where $SE[Ln(OR)] = \sqrt{([1/n_i] + [1/N - n_i])}$, and N is the total number of children and n_i is the number of children with the specific number of risk factors.

We used linear regression to determine the relation between age and VO_{2peak} for boys and girls separately as this relation was gender-specific (sex*age interaction: $p < .001$). Based on these equations we calculated the mean VO_{2peak} values for children aged 8 to 18. As the SD increased with age in boys and decreased with age in girls, we calculated age- and sex-specific SDs by establishing the slope with age in boys and girls. We thereafter calculated the cut points for VO_{2peak} being below the 2nd, 5th, and 10th percentile using one-tailed z-scores of -2.05, -1.65, and -1.28, respectively. These proportions correspond to participants defined as at-risk having 6 (1.6%), ≥ 5 (5.2%), and ≥ 4 (10.6%) risk factors (upper quartile), for which observed clustering of risk factors was significantly higher than expected.

The probability of being at risk using clustering (upper quartile) of ≥ 4, ≥ 5, or 6 risk factors as the criterion measure was estimated by logistic regression (OR, 95% CI) for 3 different classifications based on VO_{2peak} (< 10th, < 5th, and < 2nd percentile for VO_{2peak}), 2 different classifications based on BMI (overweight or obese, and obese), and the 6 combined classifications (3 for VO_{2peak} *2 for BMI). Classification accuracy was reported as sensitivity (the true positive/total positive rate), specificity (the true negative/total negative rate), and Kohen's kappa (k). Kohen's kappa was interpreted as fair for k = 0.21–0.40 and moderate for k = 0.41–0.60 [24].

Analyses were performed using IBM SPSS v. 23 (IBM Corporation, Software Group, Somers, New York, USA). A p-value < .05 indicated statistically significant findings.

Results

Participants' characteristics

Of the 1306 9-year-olds and 993 15-year-olds included in the PANCS study, 843 (65%) 9-year-olds and 619 (62%) 15-year-olds provided valid data on all relevant variables and were included in the present analyses (Table 1). Individuals included (n = 1462) and excluded from the analyses (n = 837; n = 291–837 for different variables in the attritional analyses) was similar with respect to proportion of boys and girls, age, body mass, BMI, WC, SBP, and parents' level of education ($p > .170$). The proportion of individuals with Caucasian origin was higher among those included (91 vs. 88%, $p = .045$). Height (mean (SD) 153.1 (17.5) vs. 152.2 (17.3) cm, $p < .001$), WC:height ratio (mean (SD) 0.441 (0.047) vs. 0.447 (0.052), $p = .007$), TC:HDL ratio (median (IQR) 2.54 (0.76) vs. 2.62 (0.83), $p < .001$), TG (median (IQR) 0.63 (0.34) vs. 0.67 (0.39) mmol/l, $p < .001$), and HOMA (median (IQR) 7.86 (6.54) vs. 8.63 (8.65), $p = .003$) differed significantly in favour of the included children. Absolute VO_{2peak} was higher in the excluded individuals (median (IQR) 1.81 (1.17) vs. 1.78 (1.19) l/min, $p = .004$), whereas VO_{2peak} relative to body mass tended to be higher in the included individuals (mean (SD) 46.5 (8.08) vs. 45.9 (8.2) ml/kg/min, $p = .064$).

The observed number of individuals with clustering of risk factors was higher than expected for 4 or more risk factors. 71 (4.9%) individuals had clustering of 4 risk factors (OR = 1.48, 95% CI 1.17–1.88), 49 (3.4%) individuals had clustering of 5 risk factors (OR = 7.73, 95% CI 5.81–10.27), and 23 (1.6%) individuals had clustering of 6 risk factors (OR = 80.0, 95% CI 53.0–121). Thus, we established VO_{2peak} cut points for the proportion of individuals having clustering of ≥ 4 risk factors (n = 143, 10.6%, OR = 2.82, 95% CI 2.37–3.35), ≥ 5 risk factors (n = 75, 5.2%, OR = 11.30, 95% CI 8.92–14.33), and 6 risk factors (n = 23, 1.6%, OR = 80.0, 95% CI 53.0–121). Proportions were similar across sex and age groups ($p = .579$, Table 1).

Cut points for VO_{2peak}

Based on the proportions of individuals with clustering of ≥ 4 (10.6%), ≥ 5 (5.2%), and 6 (1.6%) risk factors, we established the 2nd, 5th, and 10th percentile cut points for VO_{2peak} using the respective one-tailed z-scores (Table 2). VO_{2peak} increased by 0.66 ml/kg/min per year for boys (VO_{2peak} = 41.99806 + 0.66173*age, $R^2 = 0.066$, $p < .001$) and decreased by 0.25 ml/kg/min per year (VO_{2peak} = 45.32124 – 0.25030*age, $R^2 = .014$, $p = .002$) for girls. The mean SD was 7.46 and increased by 0.19 ml/kg/min per year in boys, whereas the mean SD was 6.26 and decreased by 0.09 ml/kg/min per year in girls. The resulting cut points for VO_{2peak} increased with age for boys (2nd percentile: 0.28 ml/kg/min per year; 5th percentile: 0.35 ml/kg/min per year; 10th percentile: 0.42 ml/kg/min per year), but decreased with age for girls (2nd percentile: -0.06 ml/kg/min per year; 5th percentile: -0.10 ml/kg/min per year; 10th percentile: -0.13 ml/kg/min per year).

Probability and classification accuracy of cardiometabolic risk for different VO_{2peak} and BMI indices

Table 3 shows the probability of being classified with clustering of risk factors along with sensitivity, specificity, and the Kappa coefficient based on different VO_{2peak} and BMI indices. Overall, 1.6–10.6% of the individuals were defined as at risk based on actual observed clustering of ≥ 4–6 CVD risk factors, whereas the corresponding proportions defined as at risk using the 11 VO_{2peak} and BMI indices of cardiometabolic risk varied from 1.0 to 14.2%. All indices of cardiometabolic risk significantly increased the odds of poor metabolic health as defined by any outcome. The probability of being classified as at risk based on VO_{2peak} varied from OR 14.2 to 33.0 among the outcome criteria. Overweight or obesity (ORs 21.3–44.6), and especially obesity (63.2–91.7), showed stronger associations with metabolic risk than VO_{2peak} for all outcomes. The combined VO_{2peak} and BMI indices increased the odds of cardiometabolic risk beyond the odds of the respective variables separately for some indices, but this pattern was not consistent. Low VO_{2peak} AND overweight or obesity resulted in ORs 20.9–66.3, whereas low VO_{2peak} AND obesity resulted in ORs 21.7–153 across outcomes.

Classification accuracy varied widely across all outcomes and VO_{2peak} and BMI indices (k = 0.15–0.53). Among the VO_{2peak} indices, the 10^{th} percentile was the best index for determining clustering of ≥ 4 risk factors (k = 0.41), the 5^{th} and 10^{th} percentile were similar for determining clustering of ≥ 5 risk factors (k = 0.38–0.39), and the 2^{nd} percentile was the best index for determining clustering of 6 risk factors (k = 0.22). Among the BMI indices, overweight or obese was the best index for determining clustering of ≥ 4 risk factors (k = 0.48), whereas obesity performed best for clustering of ≥ 5 risk factors (k = 0.49) and 6 risk factors (k = 0.39). The highest kappa was found for VO_{2peak} below the 10^{th} percentile AND overweight or obese with clustering of ≥ 4 (k = 0.48) and ≥ 5 (k = 0.53) risk factors. Kappa coefficients were ≤ 0.39 for all indices for clustering of 6 risk factors. Sensitivity and specificity varied with the proportion classified at risk by the indices and outcomes.

Discussion

We are the first to present age- and sex-specific cut points for VO_{2peak} (ml/kg/min) to define children and youth with poor cardiometabolic health based on a large sample and using a maximal test with direct measurement of oxygen consumption. These cut points, together with BMI, can be used to screen children for poor cardiometabolic health.

Ruiz et al [6] recently summarized previous studies that have sought to determine VO_{2peak} cut points to indicate cardiometabolic risk in children and youth. However, these results are of limited usefulness because they did not account for the age-related development of VO_{2peak} per kg body mass[18], which is a key feature during childhood growth and development. Moreover, there are at least two major challenges and inconsistencies when constructing and comparing cut points across previous studies. One relates to the operationalization or definition of cardiometabolic risk or the metabolic syndrome (who should be targeted for intervention?), and the other relates to the proportion of children actually detected by screening, which most often correspond poorly with the target population in previous studies (who is targeted for intervention through the screening?).

Our findings show that classification accuracy for most VO_{2peak} and BMI indices were fair to moderate. Yet, it should be mentioned that the Kappa statistic provides lower values with higher class skew between positive and negative cases [24], which is a challenge in the present analysis where small proportions were defined as at risk across the criterion measures (1.6–10.6%). Nevertheless, sensitivity (the ability to flag those at risk) increased and specificity (the ability to exclude those not at risk) decreased when classifying a greater proportion as at risk. This is a key challenge when screening for risk/disease, and thus the crucial question arises: what proportion of children and youth should we classify as at risk? As there is no universally accepted standard for defining metabolic syndrome or poor cardiometabolic health in children and youth, the proportion of children defined as at risk varies widely between studies. Welk et al [17] defined 6% as at risk based on having ≥ 3 of 5 risk factors (WC, blood pressure, HDL, TG, glucose) using the National Cholesterol Education Program/Adult Treatment PANEL III criteria. Ruiz et al [11] defined 15% as at risk based on achieving ideal levels for ≥ 4 of 7 health behaviors/factors (smoking, BMI, physical activity, diet, glucose, blood pressure, and TG) using cut points suggested by the American Heart Association. Using pooled data from 15794 6–18-year-olds, Andersen et al [16] showed that the prevalence of metabolic syndrome as defined by the IDF criteria was less than 1%, whereas 6.2 and 15.7% were classified as at risk of poor cardiometabolic health when defined as exhibiting clustering (upper quartile) of ≥ 4 and ≥ 3 of 5 risk factors (WC, SBP, inverse HDL, TG, and HOMA), respectively. A common approach to avoid disagreement when applying absolute cut points for CVD risk factors in children and youth has been to define individuals scoring < -1 SD (16%) on a continuous composite score (sum of z-scores) as at risk [10 16 19 25]. We used clustering of 6 (1.6% of individuals), ≥ 5 (5.2% of children), and ≥ 4 (10.6% of children) of 6 risk factors (WC:height ratio, SBP, TC:HDL ratio, TG, HOMA and inverse VO_{2peak}) as a basis for defining the 2nd, 5th, and 10th percentile VO_{2peak} cut points. Thus, similar to Andersen et al [8], we used a biological basis for defining at-risk individuals, by determining the proportions exhibiting clustering of more risk factors than expected given independence among risk factors. Nonetheless, to facilitate comparability among studies we have provided age- and sex-specific mean VO_{2peak} values along with SDs, allowing for calculation of any percentile in the present sample.

A further challenge in the existing literature is that most previous studies have applied Receiver Operating Characteristics (ROC) curve analyses to determine the best cut points for a healthy VO_{2peak}. By definition, when using the ROC approach, the best cut point is selected as the one that balances sensitivity and specificity. Due to the class skew (e.g., 10 % at risk; 90% not at risk) and the moderate relationship between the exposure and the outcome, achieving a balanced true positive and true negative rate leads to a "right-skew" of the cut points, resulting in cut points that define an exaggerated proportion of the sample at risk (Table 4), which leads to a high absolute false positive rate. Thus, although most previous studies have defined metabolic risk as scoring below < -1 SD on a continuous composite score, by definition classifying 16% of individuals at risk, the cut points established have selected 9–22 [19], 21–37 [16], and 44–49% [25] as at risk. Similarly, Welk et al [17] classified 6.3 and 5.9% of boys and girls with the metabolic syndrome, whereas their cut points classified 33–48% as at risk. Also, Ruiz et al [11] classified 13 and 16% of boys and girls with cardiometabolic risk, whereas their cut points classified 30–38% as at risk. As can be seen in Table 4, the cut points suggested by Ruiz et al [25] and Welk et al ("healthy fitness zone") [17] are close to the mean values of the samples, also found by Mesa et al [9]. This implies that most previous cut points are substantially overestimated (too high) in terms of identifying the target population defined to have poor cardiometabolic health. The lower proportion classified as at risk explains the lower sensitivity and higher specificity for most VO_{2peak} and BMI indices in the present study, compared to these previous studies [9 11 16 17 25].

Discriminating between children that are at risk and children not at risk is a challenging task given the trade-off between sensitivity and specificity, which consequently affects the proportion of children that could be targeted for preventive initiatives. The combined VO_{2peak} and BMI indices provided the best classification accuracy, with estimates similar to a previous study investigating the performance of a noninvasive composite score using the Andersen aerobic intermittent running field test and BMI in 10-year-old children (k = 0.52–0.53) [7]. In a class of 30 schoolchildren, assuming 10% are at risk as defined by having clustering of ≥ 4 of 6 CVD risk factors in the present study, defining children as at risk according to the index with the highest kappa (k = 0.48), the *VO_{2peak} 10^{th} percentile AND overweight/obese* index (6.0% of children at classified at risk; sensitivity 61%; specificity 97%) could classify (≈)2 out of 3 children correctly (true positive) and 1 out of 27 children incorrectly (false positive) as at risk. These numbers would obviously change with a stricter or more liberal cut point. Ultimately, the choice of VO_{2peak} cut points (and cardiometabolic risk criteria) must be a comprehensive consideration of the pros and cons of misclassification, including the health risk of being erroneously missed (false negative), the worry and/or stigmatization of being erroneously detected (false positive), and the type of preventive action that can be offered.

Of great importance for screening, VO_{2peak} can be estimated from different field tests, which makes aerobic fitness both a simple and feasible indicator of cardiometabolic risk, especially in a school setting where more invasive measures of traditional CVD risk factors are less practical to obtain. However, estimation of VO_{2peak} from field tests is limited by large prediction errors on the individual level [14 15 26]. Therefore, future studies should determine cut points for performance measures directly using for example the MSRT (laps or speed) [12] or the Andersen test (meters) [27]. In fact, using these tests can improve the prediction of metabolic health in children, as it has been shown that the Andersen test associates more strongly with metabolic health than does VO_{2peak} [28]. Derived from our findings that 10.6 and 5.2% of children across age had clustering of ≥ 4 and 5 risk factors (which were thus used as the criteria for our VO_{2peak} cut points), we suggest the 5th and 10th percentile thresholds as determined from the international MSRT references suggested by Tomkinson et al [29], could indicate increased metabolic risk. Yet, these proportions differ substantially from those suggested by Tomkinson et al, derived from the FITNESSGRAM cut points [17 20], for reasons discussed above.

Perspectives

Given the indisputable health effects of regular physical activity [30 31], one might argue increasing sensitivity at the expense of decreasing specificity by choosing, for example, the 50th VO_{2peak} percentile cut point as indicative of CVD risk, would be a minor problem. The 50th percentile in the present study led to sensitivity of 93.0, 97.2, and 100% (specificity 54.9, 52.7, and 51.0%), respectively, using clustering of ≥ 4, ≥ 5, and 6 CVD risk factors as the outcomes (results not shown). However, targeting half the population with physical activity initiatives to capture most individuals at risk points toward a population strategy rather than a high-risk strategy for prevention [32]. As such, if using a population strategy that targets all individuals, screening would be of minor relevance. Both strategies have their advantages and disadvantages, and both are likely needed in an effort to prevent poor health [32]. Importantly, increased school-based physical activity can significantly increase aerobic fitness and improve cardiometabolic health [33-35], with greater effects in the children having the least favorable cardiometabolic profile [33]. This finding demonstrates that high-risk children can be reached through a school-based population approach to combat low physical activity levels among children and youth [36].

Strengths and limitations

This study has several strengths. Most important is the use of a large, nationally representative sample of children and youth performing a maximal test with direct measurement of oxygen consumption to determine VO_{2peak}. The maximal effort is verified by high peak heart rates and respiratory exchange ratio values across the age and sex groups. Yet, we used a cycle ergometer test, which is known to produce lower values than uphill treadmill running in both children (5% lower) [37] and adults (7% lower) [38 39]. Thus, raising our cut points ≈2–3 ml/kg/min would make them equivalent to values obtained by a treadmill protocol. Further, while most previous studies [6 9-11 16], including the systematic review by Ruiz et al [6], only suggest cut points for age groups, the age- and sex-specific cut points provided herein allow for a more precise and accurate identification of risk at all ages. Moreover, researchers and clinicians can use the present findings to classify schoolchildren of all ages as at risk according to the 2nd, 5th, or 10th percentile, or alternatively, by using the means and SDs reported to calculate any given percentile according to VO_{2peak} based on this representative sample. However, it should be borne in mind that the age range included in this study comprised children aged 8.7 to 10.4 years and 14.7 to 16.7 years, and that trends by age could be non-linear. Additionally, on an individual level, maturation and pubertal development could influence both CVD risk factors and VO_{2peak} beyond chronological age. Yet, we did not account for these factors because age and maturation status (analyzed by Tanner stage and peak height velocity offset) was completely collinear ($r \geq 0.97$) in the current study, due to the study sampling. Moreover, applying this information would reduce the clinical value of the current data, for example if used for screening in a school setting, as this information might not be readily available.

A limitation of our study is the cross-sectional design, which precludes determination of the predictive validity of the suggested VO_{2peak} cut points. Thus, future studies should apply and cross-validate the cut points in new samples and using longitudinal study designs. Furthermore, how to verify a maximal (peak) effort on graded exercise tests in children is a matter of debate [40]. Our secondary criteria for accepting a VO_{2peak} test may seem rather low. However, due to the large individual variation in maximal (peak) values, we used these criteria to avoid excluding participants erroneously. Moreover, our peak heart rate levels, obtained from cycle ergometry, are comparable to other studies in children using both treadmill and cycle ergometry (200-204 beats/min) [41-43], despite cycle ergometry being known to provide lower maximal values than treadmill exercise [40].

Finally, we did not measure the participants' lean body mass. Reporting VO_{2peak} per kg lean body mass has been suggested to be the best expression of aerobic capacity because it theoretically makes sense to express aerobic capacity relative to the maximal work capacity of muscle (not including fat tissue), and because this measure is empirically weaker as related to fatness than VO_{2peak} per kg body mass [44]. Alternatively, allometric scaling (raising body mass to a power function) can be used to reduce the relation to body size and fatness [44]. VO_{2peak} per kg body mass is generally negatively related to indices of fatness [44 45] (r = -0.23 with body mass and r = -0.42 with BMI in the present study after adjustment for age and sex, results not shown), meaning that VO_{2peak} per kg body mass in relation to metabolic health is confounded by fatness. Thus, studies in children have shown that associations with metabolic health indices have been attenuated when expressing VO_{2peak} per kg lean mass versus body mass or when controlling for fatness [43 45 46]. Contrary to these studies, which sought to determine the association between aerobic fitness and diverse indicators of metabolic risk, we present herein cut points that can be used to classify children at risk. Thus, the abovementioned studies and the present study are answering two different questions. One question is whether cardiorespiratory fitness is related to metabolic health *per se*. The other is whether aerobic fitness could predict metabolic risk and thus be used for screening in healthy populations. This is a fundamental distinction, because the first question is a question of etiology (where confounding should be removed), whereas the other is a question of prediction (where confounding is no problem or actually could be seen as a strength). If VO_{2peak} were 100% confounded by a variable strongly related to metabolic health, VO_{2peak} would be an excellent predictor and the relating cut points would consequently do an excellent job in discriminating between those at risk and those not at risk. The present study's research question is one of prediction; we have therefore not attempted to remove possible confounding as in this case it does not make sense to remove information relevant for the outcome.

Conclusion

We present VO_{2peak} (ml/kg/min) cut points to classify children and youth with poor cardiometabolic health without using invasive measures of traditional CVD risk factors. Together with BMI, these cut points can be used to screen schoolchildren for poor cardiometabolic health and thereby inform teachers and health authorities who to target with physical activity initiatives. However, screening deserves thorough consideration of the target population and the approach for intervention, given the moderate discriminating performance for current cardiometabolic risk and uncertain future risk.

Acknowledgements

Financial support was received from the Norwegian Directorate of Health and the Norwegian School of Sport Sciences. The authors would like to thank all of the test personnel for their work during the data collection process and the staff at the Central Laboratory, Ullevaal University Hospital, the Hormon Laboratory Aker University Hospital for performing the blood analyses. We also thank Paul Remy Jones for giving input on the paper.

References

1. Andersen LB, Sardinha LB, Froberg K, et al. Fitness, fatness and clustering of cardiovascular risk factors in children from Denmark, Estonia and Portugal: the European Youth Heart Study. *Int J Pediatr Obes* 2008;3 Suppl 1:58-66. doi: 10.1080/17477160801896366
2. Anderssen SA, Cooper AR, Riddoch C, et al. Low cardiorespiratory fitness is a strong predictor for clustering of cardiovascular disease risk factors in children independent of country, age and sex. *Eur J Cardiovasc Prev & Rehabil* 2007;14(4):526-31. doi: 10.1097/HJR.0b013e328011efc1
3. Ortega FB, Ruiz JR, Castillo MJ, et al. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes* 2008;32(1):1-11. doi: 10.1038/sj.ijo.0803774
4. Hogstrom G, Nordstrom A, Nordstrom P. High aerobic fitness in late adolescence is associated with a reduced risk of myocardial infarction later in life: a nationwide cohort study in men. *Eur Heart J* 2014;35(44):3133-40. doi: 10.1093/eurheartj/eht527
5. Ruiz JR, Castro-Pinero J, Artero EG, et al. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009;43(12):909-23. doi: 10.1136/bjsm.2008.056499
6. Ruiz JR, Cavero-Redondo I, Ortega FB, et al. Cardiorespiratory fitness cut points to avoid cardiovascular disease risk in children and adolescents; what level of fitness should raise a red flag? A systematic review and meta-analysis. *Br J Sports Med* 2016;50(23):1451-58. doi: 10.1136/bjsports-2015-095903
7. Lerum Ø, Aadland E, Andersen LB, et al. Validity of noninvasive composite scores to assess cardiovascular risk in 10-year-old children. *Scand J Med Sci Sports* 2017 doi: 10.1111/sms.12826
8. Andersen LB, Lauersen JB, Brond JC, et al. A new approach to define and diagnose cardiometabolic disorder in children. *J Diabetes Res* 2015:10. doi: 10.1155/2015/539835
9. Mesa JL, Ruiz JR, Ortega FB, et al. Aerobic physical fitness in relation to blood lipids and fasting glycaemia in adolescents: influence of weight status. *Nutr Metab Cardiovasc Dis* 2006;16(4):285-93.
10. Moreira C, Santos R, Ruiz JR, et al. Comparison of different VO_{2max} equations in the ability to discriminate the metabolic risk in Portuguese adolescents. *J Sci Med Sport* 2011;14(1):79-84. doi: 10.1016/j.jsams.2010.07.003
11. Ruiz JR, Huybrechts I, Cuenca-Garcia M, et al. Cardiorespiratory fitness and ideal cardiovascular health in European adolescents. *Heart* 2015;101(10):766-73. doi: 10.1136/heartjnl-2014-306750
12. Leger LA, Mercier D, Gadoury C, et al. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci* 1988;6(2):93-101. doi: 10.1080/02640418808729800
13. Matsuzaka A, Takahashi Y, Yamazoe M. Validity of the multistage 20-m shuttle-run test for Japanese children, adolescents, and adults. *Pediatr Exerc Sci* 2004;16(2):113-25.
14. Batista MB, Cyrino ES, Arruda M, et al. Validity of equations for estimating VO_{2peak} from the 20-m shuttle run test in adolescents aged 11-13 years. *J Strength Cond Res* 2013;27(10):2774-81. doi: 10.1519/JSC.0b013e3182815724
15. Melo X, Santa-Clara H, Almeida JP, et al. Comparing several equations that predict peak VO2 using the 20-m multistage-shuttle run-test in 8-10-year-old children. *Eur J Appl Physiol* 2011;111(5):839-49. doi: 10.1007/s00421-010-1708-z
16. Lobelo F, Pate RR, Dowda M, et al. Validity of cardiorespiratory fitness criterion-referenced standards for adolescents. *Med Sci Sports Exerc* 2009;41(6):1222-29. doi: 10.1249/MSS.0b013e318195d491
17. Welk GJ, Laurson KR, Eisenmann JC, et al. Development of youth aerobic-capacity standards using receiver operating characteristic curves. *Am J Prev Med* 2011;41(4):S111-S16. doi: 10.1016/j.amepre.2011.07.007
18. Kolle E, Steene-Johannessen J, Andersen LB, et al. Objectively assessed physical activity and aerobic fitness in a population-based sample of Norwegian 9- and 15-year-olds. *Scand J Med Sci Sports* 2010;20(1):1-7. doi: 10.1111/j.1600-0838.2009.00892.x

19. Adegboye ARA, Anderssen SA, Froberg K, et al. Recommended aerobic fitness level for metabolic health in children and adolescents: a study of diagnostic accuracy. *Br J Sports Med* 2011;45(9):722-28. doi: 10.1136/bjsm.2009.068346
20. Plowman SA, Meredith MD. Fitnessgram/Activitygram reference guide (4th Edition). Dallas: The Cooper Institute. 2013.
21. Steene-Johannessen J, Kolle E, Anderssen SA, et al. Cardiovascular disease risk factors in a population-based sample of Norwegian children and adolescents. *Scand J Clin Lab Invest* 2009;69(3):380-6. doi: 10.1080/00365510802691771
22. Cole TJ, Bellizzi MC, Flegal KM, et al. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320(7244):1240-43. doi: 10.1136/bmj.320.7244.1240
23. Matthews DR, Hosker JP, Rudenski AS, et al. Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 1985;28(7):412-19. doi: 10.1007/bf00280883
24. Altman D. Practical statistics for medical research. Florida: CRC Press 1991.
25. Ruiz JR, Ortega FB, Rizzo NS, et al. High cardiovascular fitness is associated with low metabolic risk score in children: The European Youth Heart Study. *Pediatr Res* 2007;61(3):350-55. doi: 10.1203/pdr.0b013e318030d1bd
26. Aadland E, Andersen LB, Lerum Ø, et al. The Andersen aerobic fitness test: New peak oxygen consumption prediction equations in 10 and 16-year olds. *Scand J Med Sci Sports* 2017 doi: 10.1111/sms.12985
27. Andersen LB, Andersen TE, Andersen E, et al. An intermittent running test to estimate maximal oxygen uptake: the Andersen test. *J Sports Med Phys Fitness* 2008;48(4):434-37.
28. Aadland E, Kvalheim OM, Rajalahti T, et al. Aerobic fitness and metabolic health in children: A clinical validation of directly measured maximal oxygen consumption versus performance measures as markers of health. *Prev Med Rep* 2017 doi: 10.1016/j.pmedr.2017.05.001
29. Tomkinson GR, Lang JJ, Tremblay MS, et al. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. *Br J Sports Med* 2017;51(21):1545-+. doi: 10.1136/bjsports-2016-095987
30. Poitras VJ, Gray CE, Borghese MM, et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Appl Physiol Nutr Metab* 2016;41(6):S197-S239. doi: 10.1139/apnm-2015-0663
31. Janssen I, LeBlanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int J Behav Nutr Phys Act* 2010;7 doi: 10.1186/1479-5868-7-40
32. Rose G. Sick individuals and sick populations. 1985. *Bull World Health Organ* 2001;79(10):990-96.
33. Resaland GK, Aadland E, Nilsen AKO, et al. The effect of a two-year school-based daily physical activity intervention on a clustered CVD risk factor score-The Sogndal school-intervention study. *Scand J Med Sci Sports* 2017 doi: 10.1111/sms.12955
34. Kriemler. Effect of school based physical activity programme (KISS) on fitness and adiposity in primary schoolchildren: cluster randomised controlled trial (vol 340, pg c785, 2010). *BMJ* 2010;340 doi: 10.1136/bmj.c2968
35. Minatto G, Barbosa VC, Berria J, et al. School-based interventions to improve cardiorespiratory fitness in adolescents: systematic review with meta-analysis. *Sports Med* 2016;46(9):1273-92. doi: 10.1007/s40279-016-0480-6
36. Cooper AR, Goodman A, Page AS, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). *Int J Behav Nutr Phys Act* 2015;12 doi: 10.1186/s12966-015-0274-5
37. Mamen A, Resaland GK, Mo DA, et al. Comparison of peak oxygen uptake in boys exercising on treadmill and cycle ergometers. *Acta Kinesiologiae Universitatis Tartuensis* 2007;12:130-30.
38. Hermansen L, Saltin B. Oxygen uptake during maximal treadmill and bicycle exercise. *J Appl Physiol* 1969;26(1):31-7.

39. Basset FA, Boulay MR. Specificity of treadmill and cycle ergometer tests in triathletes, runners and cyclists. *Eur J Appl Phys* 2000;81(3):214-21. doi: 10.1007/s004210050033
40. Armstrong N. Understanding the role of aerobic fitness in relation to young people's health and well-being. *Phys Ther Rev* 2017;22(3-4):133-38. doi: 10.1080/10833196.2017.1287647
41. Resaland GK, Andersen LB, Mamen A, et al. Effects of a 2-year school-based daily physical activity intervention on cardiorespiratory fitness: the Sogndal school-intervention study. *Scand J Med Sci Sports* 2011;21(2):302-09. doi: 10.1111/j.1600-0838.2009.01028.x
42. Lintu N, Viitasalo A, Tompuri T, et al. Cardiorespiratory fitness, respiratory function and hemodynamic responses to maximal cycle ergometer exercise test in girls and boys aged 9-11 years: the PANIC Study. *Eur J Appl Physiol* 2015;115(2):235-43. doi: 10.1007/s00421-014-3013-8
43. Shaibi GQ, Cruz ML, Ball GDC, et al. Cardiovascular fitness and the metabolic syndrome in overweight Latino youths. *Med Sci Sports Exerc* 2005;37(6):922-28. doi: 10.1029/01.mss.0000170472.75214.53
44. Loftin M, Sothern M, Abe T, et al. Expression of VO_{2peak} in children and youth, with special reference to allometric scaling. *Sports Med* 2016;46(10):1451-60. doi: 10.1007/s40279-016-0536-7
45. McMurray RG, Hosick PA, Bugge A. Importance of proper scaling of aerobic power when relating to cardiometabolic risk factors in children. *Ann Hum Biol* 2011;38(5):647-54. doi: 10.3109/03014460.2011.598561
46. Ahn B, McMurray R, Harrell J. Scaling of VO_{2max} and its relationship with insulin resistance in children. *Pediatr Exerc Sci* 2013;25(1):43-51. doi: 10.1123/pes.25.1.43