Validity of the Supramaximal Test to Verify Maximal Oxygen Uptake in Children and Adolescents

Supramaximal Test Verification of Pediatric V̇O2max

Kate M. Sansum[1], Max E. Weston[1], Bert Bond[1], Emma J. Cockcroft[2], Amy O'Connor[1], Owen W. Tomlinson[1], Craig A. Williams[1] & Alan R. Barker[1*]

[1]*Children's Health and Exercise Research Centre, Sport and Health Sciences, University of Exeter, Exeter, Devon, UK.*

[2]*NIHR Collaboration for Leadership in Applied Health Research and Care South West Peninsula (PenCLAHRC), University of Exeter Medical School, Exeter, UK.*

*****Corresponding author – A.R.Barker@exeter.ac.uk

Abstract

***Purpose*:** This study had two objectives: 1) to examine whether the validity of the supramaximal verification test for maximal oxygen uptake ($\dot{V}O_{2max}$) differs in children and adolescents when stratified for sex, body mass and cardiorespiratory fitness (CRF); and 2) to assess sensitivity and specificity of primary and secondary objective criteria from the incremental test to verify $\dot{V}O_{2max}$. ***Methods*:** 128 children and adolescents (76 males, 52 females; 9.3-17.4 y) performed a ramp-incremental test to exhaustion on a cycle ergometer followed by a supramaximal test to verify $\dot{V}O_{2max}$. ***Results*:** Supramaximal tests verified $\dot{V}O_{2max}$ in 88% of participants. Group incremental test peak $\dot{V}O_2$ was greater than the supramaximal test (2.27 ± 0.65 L·min^{-1} and 2.17 ± 0.63 L·min^{-1}; P<0.001), although were correlated (r =0.94; P<0.001). No differences were found in $\dot{V}O_2$ plateau attainment or supramaximal test verification between sexes, body mass or CRF statuses (all P>0.18).

Supramaximal test time to exhaustion predicted supramaximal test $\dot{V}O_{2max}$ verification (*P*=0.040). Primary and secondary objective criteria had insufficient sensitivity (7.1-24.1%) and specificity (50-100%) to verify $\dot{V}O_{2max}$. ***Conclusion*:** The utility of supramaximal testing to verify $\dot{V}O_{2max}$ is not affected by sex, body mass or CRF status. Supramaximal testing should replace secondary objective criteria to verify $\dot{V}O_{2max}$.

Key Words: cardiorespiratory fitness, youth, verification test, maximal oxygen uptake, cycle ergometer

Introduction

Maximal oxygen uptake ($\dot{V}O_{2max}$), typically expressed in relation to a measure of body size, is the "gold-standard" measure of cardiorespiratory fitness (CRF) (10). A valid measurement of $\dot{V}O_{2max}$ is important in children and adolescents as a high CRF in youth is associated with a lower risk of cardiovascular disease in youth (24), a reduced risk of myocardial infarction (17) and all-cause mortality (18) in adult life. Traditionally, the presence of a plateau of $\dot{V}O_2$ at, or close to exhaustion, during incremental exercise has been used as the primary criterion for attainment of $\dot{V}O_{2max}$ (36). However, as only between 10-50% of children display a plateau across different testing protocols (5, 29, 30, 32), with the reasons behind this still being unclear, the term $\dot{V}O_{2peak}$ is routinely used to denote the highest $\dot{V}O_2$ recorded without a plateau (1). Secondary objective criteria (e.g. respiratory exchange ratio (RER) and maximal heart rate (HR_{max}) thresholds) (3, 33) are therefore often used to verify that the $\dot{V}O_{2peak}$ attained was a "true" $\dot{V}O_{2max}$ for children and adolescents but significantly underestimate $\dot{V}O_{2max}$ (5). While, the validity of secondary objective criteria has recently been challenged (5, 26), their use is still commonplace in contemporary pediatric research (e.g. 15, 28).

To overcome the validity issues with the primary and secondary criteria, the 1990s saw the emergence of a supramaximal test to verify that the $\dot{V}O_{2peak}$ that had been achieved in the incremental test is a "true" $\dot{V}O_{2max}$. This requires participants to exercise at a power output greater than the maximal power output achieved during the incremental test (3, 33) and is a variation of the original protocol proposed by Taylor *et al.* (36). However, the supramaximal verification tests were initially conducted on separate days (3, 33), which may not be feasible, due to logistical requirements of supplementary laboratory visits. Recently, it has been shown that children can successfully perform the supramaximal verification test on the same day as the incremental test, following a short rest of 10-15 min following the incremental test (5, 7, 31), and is now the recommended protocol for $\dot{V}O_{2max}$ determination (6, 27).

While the supramaximal verification test is an elegant solution to determine $\dot{V}O_{2max}$ in children and adolescents, not all participants have their $\dot{V}O_{2max}$ confirmed in the supramaximal test. Between 8-26% of participants have been reported as not having $\dot{V}O_{2max}$ verified (5, 31), but a recent paper reported a non-verification rate of 100% of children (7). Previous studies have suggested that non-verification may be related to factors such as sex, body mass, CRF and/or maturity status (4, 7, 9) but the current pediatric literature is based on small sample sizes (n = 9-40) (3, 5, 31, 33), which limits examination of these variables on $\dot{V}O_{2max}$ verification. Consequently, male and female data have been combined for analysis (5), despite known sex-differences in $\dot{V}O_{2max}$ and body composition (2, 3). Recently it has been shown that adults with low CRF were less likely to have their $\dot{V}O_{2max}$ confirmed in the supramaximal test than those with a higher CRF status (4), but this has not been investigated in children and adolescents. Few studies (4, 31) have compared individual $\dot{V}O_{2peak}$ values from the incremental test to the supramaximal test, instead comparing the group means (3, 33) which may be misleading as $\dot{V}O_{2max}$ testing is conducted on an individual basis. The effect of body mass status on the verification of $\dot{V}O_{2max}$ using supramaximal exercise has only been studied by Bhammar *et al.* (7) who found both obese (n = 9) and non-obese (n = 9) children to have a significantly greater $\dot{V}O_{2peak}$ in the supramaximal test. Brown *et al.* (9) have also suggested that maturity status may influence plateau attainment during an incremental test, with 23.8% of men achieving a plateau compared with 12.5% of boys, but this has not been investigated in the context of the supramaximal verification test. Finally, the effect of the time to exhaustion (TTE) in the supramaximal test on the utility of the supramaximal verification test is worthy of consideration. TTE in studies where between 74-92% of children had $\dot{V}O_{2max}$ verified is reported to be between 60 to 90 s in duration (5, 31). However, Bhammar *et al.* (7) reported that no children had their $\dot{V}O_{2max}$ verified by the supramaximal verification test and reported a TTE in excess of 125 s. Conversely, a short TTE (e.g. of less than 60 s) could indicate that fatigue is reached before attainment of $\dot{V}O_{2max}$, possibly due to insufficient effort or because the intensity was too high for the $\dot{V}O_2$ kinetic response to attain $\dot{V}O_{2max}$ (16).

The purpose of the current study was to extend previous work in this area (5) and further examine the validity of testing procedures to determine $\dot{V}O_{2max}$ in a large sample of healthy children and adolescents. Specifically, our aims were to: 1) examine whether the validity of the supramaximal verification test differs in children and adolescents when stratified for sex, body mass and CRF status; and 2) assess the sensitivity and specificity of primary (i.e. plateau) and secondary (i.e. RER and HR thresholds) objective criteria to verify the $\dot{V}O_{2peak}$ attained in the incremental test as $\dot{V}O_{2max}$ when compared to a supramaximal confirmed $\dot{V}O_{2max}$ measurement.

Methods

Participants:

Existing data from our laboratory were pooled and retrospectively analysed to produce a sample of 128 healthy children and adolescents. Only data from 13 participants that form the final sample have previously been published elsewhere to examine the validity of the supramaximal test (5). All data were collected as part of studies which originally were granted ethics approval by institutional and NHS ethics committees (where relevant). Inclusion criteria for this study were: 1) 8-<18 years old and; 2) $\dot{V}O_{2max}$ assessed using a combined incremental and supramaximal test protocol, conducted on the same day; 3) ostensibly healthy participants and; 4) cycling modality. All children and their parent(s) or guardian(s) gave informed assent and consent, respectively, to participate in the original studies.

Anthropometry:

Body mass (Seca 877, Seca Ltd, Birmingham, UK) was measured to the nearest 0.1 kg and stature (Harpenden, Holtain Ltd, Crymych, UK) was measured to the nearest 0.01 m. Body mass index (BMI) was calculated, and age-appropriate criteria were used to classify participants into non-overweight and overweight/obese categories (12). Maturity (somatic) offset from the age of peak height velocity (APHV) was calculated through the equations by Moore *et al.* (21), which have been validated in two external samples where 90% of predictions are within ± 1 year. Pre-peak height velocity (PHV) children were defined as >-1 year from PHV, circa-PHV children were -1 to 1 year from PHV and post-PHV children were >1 year from PHV.

Incremental and supramaximal test protocols:

A combined incremental-ramp and supramaximal test to exhaustion was used to determine $\dot{V}O_{2max}$ (5). Participants were instructed to cycle on an electronically braked ergometer (Lode Excalibur, Groningen, The Netherlands) at a constant self-selected cadence between 70 and 90 revolutions per minute throughout the tests. Participants cycled for ~ 3 min (range 1 to 3 min) at 20 watts (W) to warm up before immediately commencing the incremental-ramp protocol where the power output increased by 10-30 W·min^{-1}, depending on the participants' age and body size, to attempt to elicit exhaustion between 8 and 12 min. Exhaustion was defined as a decrease in cadence below 60 revolutions per minute for 5 consecutive seconds, despite strong verbal encouragement. This was followed by 3 min 30 s (range 0 to 10 min) cool down cycling at 20 W. A rest period of ~ 25 min (range = 5 to 84 min) followed before the commencement of the supramaximal test, which began with a warm up of 3 min at 20 W. The resistance was then increased in a "step" transition to either ~105% (*n* = 117) or ~110% (*n* = 11) of the peak power achieved in the incremental test and participants were required to cycle to exhaustion. Following the test, the participant completed a cool down cycling at 20 W. The measurement of $\dot{V}O_{2peak}$ from the ramp-incremental test to exhaustion has a coefficient of variation of 4.1% (37).

Gas collection and analysis:

Pulmonary gas exchange and heart rate (HR) were measured using online systems (Cortex Metalyzer III B, Leipzig, Germany: *n* =106; EX671; Morgan Medical, Kent, UK, combined with mass spectrometry and a turbine flow meter VMM-401; Interdace Associates, Laguna Niguel, California, USA: *n* =13; and Medgraphics Cardiorespiratory Diagnostics, Express Series, Gloucester, UK: *n* =9). All systems were appropriately calibrated for gas and volume before each test as per manufacturers' recommendations. $\dot{V}O_{2max}$ was accepted as the highest 10-15 second average of $\dot{V}O_2$ recorded in either the incremental or supramaximal tests (5). To control for body-size, both the ratio standard and allometric (via log-linear regression, (38)) models were used to scale $\dot{V}O_{2max}$ for body mass. Although allometric procedures are superior for scaling $\dot{V}O_2$ (39), normative data are unavailable to classify the children and adolescents into CRF groups. Therefore, the ratio standard method to scale for body mass was used to group participants into CRF statuses of low, average and high CRF based on age and sex related normative values (8). Low CRF was defined as > 1 standard deviation (SD) below the age and sex specific mean normative value, average CRF was defined as falling within 1 SD either side of the age and sex specific mean normative value, and high CRF was defined as > 1 SD above the age and sex specific mean normative value.

Criteria and $\dot{V}O_2$ profile classification during incremental exercise:

The methods proposed by Day *et al.* (13) were used to define a plateau and classify the $\dot{V}O_2$ responses during incremental exercise into a linear, acceleration or deceleration profile using GraphPad Prism (GraphPad Software Inc., San Diego, California, USA). A linear regression of the $\dot{V}O_2$-intensity relationship was plotted over the 'linear' portion of the $\dot{V}O_2$ profile, where the data points from the first 2 minutes and the last 3 minutes of exercise were excluded. The linear function was then extrapolated and compared to the residuals to analyse the $\dot{V}O_2$ profile at exhaustion for an accelerated, decelerated (i.e. plateau) or linear response. An accelerated profile required the positive residual to be ≥5% of the peak power projected $\dot{V}O_2$ whereas a decelerated profile required the negative residual to be ≥5% of the projected $\dot{V}O_2$. A linear response was classified by residuals that were <5% of the peak power projected $\dot{V}O_2$, either side of the extrapolated line.

The secondary objective criteria to verify $\dot{V}O_{2max}$ were selected from the pediatric literature (1, 3, 5, 7, 14) and included: RER ≥ 1, RER ≥ 1.1, HR_{max} > 85% of the age-predicted maximum (calculated using 220 minus age), HR_{max} > 95% age-predicted maximum and HR_{max} > 195 beats$\cdot$min^{-1}. HR data are not available for 18 participants, and therefore were excluded from the HR criteria analyses.

Criteria for verification of $\dot{V}O_{2max}$ using the supramaximal test:

As used by Barker *et al.* (5), $\dot{V}O_{2max}$ was considered verified by the supramaximal test if the $\dot{V}O_2$ increased by <5% compared to the $\dot{V}O_{2peak}$ attained in the incremental test to account for the typical within-participant error of measurement for $\dot{V}O_{2max}$ (25, 37).

Statistical analyses:

Data were analysed using IBM SPSS (v24, Armonk, NY, USA) and presented as mean ± standard deviation (SD), unless otherwise stated. Statistical significance was accepted at an alpha of 0.05 and data were checked for normality using the Shapiro-Wilk test. Data were log transformed when the normality assumption was violated. Independent t-tests were conducted to examine mean differences in participant characteristics between sex and between body mass statuses within each sex. Chi-squared analyses were performed to test for significant differences in the percentages of males compared to females, overweight compared to non-overweight, and different CRF status' that achieved a plateau during the incremental test and had their $\dot{V}O_{2max}$ verified with the supramaximal test. Paired t-tests and effect sizes (ES) using Cohen's d thresholds (< 0.2 trivial, 0.2 = small. 0.5 = medium, 0.8 = large) (11) were used to compare the $\dot{V}O_{2peak}$ values from the incremental and supramaximal tests for the whole sample, and when stratified for sex, body mass and CRF status. The relationship between the $\dot{V}O_{2peak}$ recorded in the incremental and supramaximal tests was assessed using Pearson's product moment correlation coefficients. Bland-Altman (20) analyses were used to show the absolute ($L \cdot min^{-1}$) and relative (%) level of agreement in the $\dot{V}O_{2max}$ recorded via the incremental and supramaximal tests with 95% limits of agreement (95% LoA) for the whole group combined, and based on sex, body mass and CRF status. Checks for proportional bias were undertaken using Pearson's correlation and satisfied for all Bland-Altman plots.

Separate logistic regression analyses were run to identify 1) significant predictors of plateau attainment in the incremental test and; 2) verification of $\dot{V}O_{2max}$ through the use of the supramaximal test. The variables tested in both models were age (years), sex, body mass status (overweight/obese or non-overweight), APHV (pre-, circa-, post-PHV), $\dot{V}O_{2max}$ expressed using the ratio standard and allometric methods, CRF status (low, average or high) and incremental test TTE (s). The supramaximal test TTE (s), supramaximal test intensity (% of peak power attained in the incremental test) and rest period between the incremental and supramaximal test (s) were also included for predicting $\dot{V}O_{2max}$ verification using the supramaximal test. Variables were entered using the backward stepwise (likelihood ratio) method.

Primary and secondary objective criteria from the incremental test were assessed for their sensitivity (ability to correctly identify attainment of "true" $\dot{V}O_{2max}$) and specificity (ability to correctly identify non-attainment of "true" $\dot{V}O_{2max}$) to verify $\dot{V}O_{2peak}$ in the incremental test as $\dot{V}O_{2max}$ when compared to the supramaximal test verification method. Each of the criteria was also assessed for their positive and negative predictive value i.e. the likelihood that a positive or negative result from the criteria for attainment of "true" $\dot{V}O_{2max}$ in the incremental test is the correct result. The equations below (19) were used to calculate sensitivity, specificity and the positive and negative predictive value for each criteria. The ability of each criterion to confirm $\dot{V}O_{2max}$ was assessed, allowing for 5% error at an individual level and compared to whether the supramaximal test was able to verify $\dot{V}O_{2max}$. Receiver operator characteristic (ROC) curves (1-specificity vs sensitivity) were also used to calculate the area under the curve (AUC).

Sensitivity = True positives / (True positives + False negatives)

Specificity = True negatives / (True negatives + False positives)

Positive predictive value = True positives / (True positives + False positives)

Negative predictive value = True negatives / (True negatives + False negatives)

Results

Table 1 presents the participants' characteristics and physiological responses to the incremental and supramaximal tests by sex and body mass status. The males had a greater ramp test TTE (*P*<0.001), ramp and supramaximal test absolute $\dot{V}O_{2peak}$ (both *P*<0.001), supramaximal test RER_{peak} (*P*=0.010), ratio standard $\dot{V}O_{2max}$ (*P*<0.001) and allometrically scaled $\dot{V}O_{2max}$ (P<0.001) than the females. By contrast, the females had higher BMI (*P*=0.032), APHV (*P*=0.022) and ramp test RER_{peak} (*P*<0.001) than the males. Overweight males had a higher age (*P*=0.018), stature (*P*=0.025), body mass (P<0.001), BMI (*P*<0.001), APHV (*P*=0.012), ramp test absolute $\dot{V}O_{2peak}$ (*P*=0.007), supramaximal test TTE (*P*=0.003) and supramaximal test absolute $\dot{V}O_{2peak}$ (*P*=0.001), and lower ratio standard $\dot{V}O_{2max}$ (*P*=0.002) compared with non-overweight males. Furthermore, overweight females had greater body mass (*P*=0.004) and BMI (*P*<0.001) and lower ratio standard $\dot{V}O_{2max}$ (*P*<0.001) and lower allometrically scaled $\dot{V}O_{2max}$ (P=0.004) than non-overweight females. The mean ratio standard $\dot{V}O_{2max}$ was greater in the non-overweight children and adolescents compared with the overweight children and adolescents (46 ± 10 mL·kg^{-1}·min^{-1} vs. 36 ± 8 mL·kg^{-1}·min^{-1}; *P*<0.001).

***** Insert Table 1 *****

Incremental and supramaximal responses: whole group

Analysis of individual participant $\dot{V}O_2$-intensity profiles revealed 27% (*n*= 35) of participants demonstrated a $\dot{V}O_2$ plateau during the incremental test, 14% (*n*=18) had an accelerated $\dot{V}O_2$ profile and 59% (n=75) had a linear $\dot{V}O_2$ profile. When comparing the $\dot{V}O_{2peak}$ values obtained from the incremental and supramaximal test on an individual basis, 88% (*n*=112) of children and adolescents had their $\dot{V}O_{2peak}$ in the incremental test verified as their "true" $\dot{V}O_{2max}$. For the remaining 12% who did not have $\dot{V}O_{2max}$ verified, the $\dot{V}O_{2peak}$ recorded was between 6 and 23% greater than that recorded in the incremental test.

For the entire sample, the $\dot{V}O_{2peak}$ recorded in the incremental test (2.27 ± 0.65 $L{\cdot}min^{-1}$) was higher than the supramaximal test (2.17 ± 0.63 $L{\cdot}min^{-1}$; $P<0.001$; ES = 0.15), and the two were correlated (r=0.94; $P<0.001$).

Figure 1 shows the absolute (1A) and relative (1D) differences in the incremental and supramaximal test $\dot{V}O_{2peak}$ for the whole group combined. Mean absolute and relative bias was -0.10 $L{\cdot}min^{-1}$ and -4.6% with LoA as -0.52 to 0.32 $L{\cdot}min^{-1}$ and -22 to 13%, respectively.

***** Insert Figure 1 *****

Incremental and supramaximal responses: influence of sex

There were no differences in the proportion of plateau observations during the incremental test between males and females (29%; *n* = 22 vs. 25%; *n* = 13; P=0.62). Similarly, no differences were found in the proportion of supramaximal tests that verified

$\dot{V}O_{2max}$ between males and females (89%; n = 68 vs. 85%; n = 44; P=0.41).

The mean absolute $\dot{V}O_{2peak}$ of the males recorded in the incremental test (2.48 ± 0.73 L·min^{-1}) was greater than the supramaximal test (2.36 ± 0.72 L·min^{-1}; P<0.001; ES = 0.17). Likewise, the mean absolute $\dot{V}O_{2peak}$ recorded for the females in the incremental test was higher than in the supramaximal test (1.96 ± 0.31 L·min^{-1} and 1.89 ± 0.34 L·min^{-1}, respectively; P=0.007; ES = 0.22). Incremental and supramaximal test $\dot{V}O_{2peak}$ values were correlated (males r=0.95 and females r=0.85; both P<0.001).

Figure 1 depicts the Bland-Altman plots for absolute (1B and 1C) and percentage (1E and 1F) difference in $\dot{V}O_{2peak}$ recorded between the incremental and supramaximal tests for each sex. Mean absolute and relative bias for the males was -0.12 L·min^{-1} and -5.4% with LoA as -0.58 to 0.33 L·min^{-1} and -23 to 13%, respectively. The absolute and relative mean bias for the females was -0.06 L·min^{-1} and -3.5% with LoA as -0.41 to 0.29 L·min^{-1} and -21 to 14%, respectively.

Incremental and supramaximal responses: influence of body mass status

No difference was found between the proportion of $\dot{V}O_2$ plateau observations between the non-overweight and overweight children and adolescents (26%; n = 28 vs. 37%; n = 7; P=0.31). Furthermore, no difference was found between the proportion of non-overweight compared with overweight children and adolescents who had their $\dot{V}O_{2max}$ verified in the supramaximal test (89%; n = 97 vs. 79%; n = 15; P=0.22).

The mean absolute $\dot{V}O_{2peak}$ for non-overweight children and adolescents was greater in the incremental test compared with the supramaximal test (2.27 ± 0.63 L·min^{-1} vs. 2.13 ± 0.54 L·min^{-1}; P<0.001; ES = 0.24). In contrast, the mean absolute $\dot{V}O_{2peak}$ was not different for the overweight children and adolescents between the incremental and supramaximal tests (2.57 ± 0.79 L·min^{-1} vs. 2.54 ± 0.81 L·min^{-1}; *P*=0.65; ES = 0.04). The $\dot{V}O_{2peak}$ recorded in the incremental and supramaximal tests were correlated for each group (non-overweight children and adolescents r=0.94, overweight children and adolescents r=0.95; both *P*<0.001).

Figure 2 displays the absolute (2A & 2B) and percentage (2C & 2D) differences in the $\dot{V}O_{2peak}$ from the incremental and supramaximal tests for non-overweight and overweight children and adolescents. The absolute and relative mean bias for non-overweight children and adolescents was -0.11 L·min^{-1} and -5.2%, and LoA were -0.51 to 0.29 L·min^{-1} and -22 to 12%, respectively. Mean absolute and relative bias for overweight children and adolescents was -0.03 L·min^{-1} and -1.5% with LoA as -0.54 to 0.49 L·min^{-1} and -21 to 18%, respectively.

***** Insert Figure 2 *****

Incremental and supramaximal responses: influence of CRF status

A $\dot{V}O_2$ plateau was demonstrated by 27% (*n* = 12) of the low CRF group in the incremental test compared with 28% (*n* = 17) of the average CRF group and 26% (*n* = 6) of the high CRF group (all *P*>0.84). Similarly, there were no differences between CRF statuses for supramaximal test verification which occurred for 87% (n = 39) of the low CRF group, 85% (*n* = 51) of the average CRF group and 96% (*n* = 22) of the high CRF group (all *P*>0.18).

Mean absolute $\dot{V}O_{2peak}$ was higher in the incremental compared with the supramaximal test for the low (2.08 ± 0.62 L·min^{-1} vs. 2.02 ± 0.64 L·min^{-1}; *P*=0.007; ES = 0.10), average (2.38 ± 0.55 L·min^{-1} vs. 2.26 ± 0.50 L·min^{-1}; *P*=0.001; ES = 0.23) and high (2.33 ± 0.87

$L\cdot min^{-1}$ vs. 2.21 ± 0.86 $L\cdot min^{-1}$; *P*=0.003; ES = 0.14) CRF groups, respectively. The mean absolute $\dot{V}O_{2peak}$ from incremental and supramaximal testing were correlated for the low (r=0.97), average (r=0.88) and high (r=0.98) CRF groups (*P*<0.001 for all).

Figure 3 displays the absolute (3A, 3B & 3C) and relative (3D, 3E & 3F) differences in $\dot{V}O_{2peak}$ from the incremental and supramaximal tests for low (3A & 3D), average (3B & 3E) and high (3C & 3F) CRF groups. The absolute and relative mean bias for the low CRF group was -0.06 $L\cdot min^{-1}$ and -3.3% with LoA as -0.37 to 0.25 $L\cdot min^{-1}$ and -18 and 12%. Average CRF absolute and relative mean bias was -0.12 $L\cdot min^{-1}$ and -5.1%, and LoA were -0.62 to 0.38 $L\cdot min^{-1}$ and -25 to 14%. For the high CRF group, the absolute and relative mean bias was -0.12 $L\cdot min^{-1}$ and -6.0% with LoA as -0.49 to 0.24 $L\cdot min^{-1}$ and -23 to 11%.

***** Insert Figure 3 *****

Predicting plateau attainment in the incremental test:

Of the variables entered into the model, no variables were predictors for attaining a $\dot{V}O_2$ plateau in the incremental test (*P*>0.30 for all).

Predicting verification of $\dot{V}O_{2max}$ in the supramaximal test:

TTE on the supramaximal test was the only predictor of whether the supramaximal test can verify the $\dot{V}O_{2peak}$ attained in the incremental test as $\dot{V}O_{2max}$ (*P*=0.040; odds ratio = 0.978; 95% confidence limits = 0.958-0.999). A longer supramaximal test TTE decreased the likelihood that $\dot{V}O_{2peak}$ would be verified as $\dot{V}O_{2max}$ (β=-0.022; standard error = 0.011). The regression equation below predicts supramaximal test verification.

Supramaximal test verification (Y) = 4.212 + (Supramaximal Test TTE [s] * -0.022)

Sensitivity and specificity of primary and secondary objective criteria:

Table 2 displays the sensitivity and specificity analysis on the primary and secondary objective criteria compared with the supramaximal test, as well as the positive and negative predictive values. All criteria had low sensitivity (7.1-24.1%) but the majority had high specificity (78.6-100%), apart from the plateau attainment (50%). Both primary and secondary criteria had high positive predictive values (77.1-100%). By contrast, negative predictive values were low for all criteria (8.6-14.7%) excluding the HR_{max} > 195 beats·min^{-1} (92.9%). The AUC were low, ranging from 0.527 to 0.629.

***** Insert Table 2 *****

Discussion

The main findings of the study were: 1) 88% of children and adolescents had their absolute $\dot{V}O_{2max}$ verified in the supramaximal test which had a tendency to result in a ~ 5% decrease in absolute $\dot{V}O_{2peak}$ in most, but not all of the sample; 2) the utility of the supramaximal test to verify $\dot{V}O_{2max}$was similar when stratified for sex, body mass or CRF status; 3) TTE on the supramaximal test was the only significant predictor of $\dot{V}O_{2peak}$ being verified as $\dot{V}O_{2max}$ in the supramaximal test whereas there were no significant predictors of plateau attainment in the ramp test and; 4) primary (plateau) and secondary (HR and RER thresholds) objective criteria from the incremental test have insufficient sensitivity and specificity to validate attainment of $\dot{V}O_{2max}$ in children.

The majority of our participants had their $\dot{V}O_{2max}$ verified in the supramaximal test (88%), which is line with most (5, 31) but not all (7) of the literature. For example, recent findings (7) showed that none of the obese and non-obese children had their $\dot{V}O_{2max}$ verified in the supramaximal test. This difference in findings may be due to methodological differences (e.g. their use of Douglas bags), or the smaller sample size (7). Alternatively, it could be because the participants in Bhammar *et al.*'s study (7) had a supramaximal TTE of greater than two minutes, which is uncommon in both the adult and pediatric literature (4, 5, 31) and not in line with our findings (whole group = 98 ± 23 s, with 11 participants >120 s). In the current study, supramaximal test TTE was a significant negative predictor for the supramaximal test confirming attainment of $\dot{V}O_{2max}$ which could explain the lack of verification of $\dot{V}O_{2max}$ observed in Bhammar *et al.*'s (7) study. The longer TTE may reflect that the incremental test was sub-optimal and terminated early before $\dot{V}O_{2max}$ was attained. However, it should be noted that although statistically significant, the finding that the supramaximal test TTE was a significant predictor of $\dot{V}O_{2max}$ verification is unlikely to be meaningful due to significance level in the logistic regression being P=0.040 and CLs confidence limits = 0.958-0.999. Furthermore, we were not able to identify a cut-off threshold for supramaximal test non-verification based on TTE.

Mean $\dot{V}O_{2peak}$ for the whole group was significantly lower in the supramaximal test, although the effect size was trivial, contradicting the pediatric literature (5, 7), but supporting a recent study in adults (4). A previous paper by Barker *et al.* (5) showed a similar but non-significant (P=0.09) ~ 4% decrease in $\dot{V}O_{2peak}$ in the supramaximal test compared with the incremental test in a small sample of 13 children. Robben *et al.* (31) reported a smaller negative mean bias of -0.02 $L \cdot min^{-1}$ in a sample of 27 healthy children whereas Bhammar *et al.* (7), found a positive mean bias of 0.12 $L \cdot min^{-1}$ for the supramaximal test in a small sample of 9 obese and 9 non-obese children. The reason for finding a significant difference between the $\dot{V}O_{2peak}$ in the incremental and supramaximal test, which is not in line with the literature, could be due to our much larger sample size providing greater statistical power to detect smaller differences between the $\dot{V}O_{2peak}$

recorded in each test. Bland Altman analysis revealed more variation in differences between the incremental and supramaximal tests through wider limits of agreement than previous studies; -0.52 to 0.32 $L\cdot min^{-1}$ for the whole sample compared with -0.09 to 0.33 $L\cdot min^{-1}$ (7) and -0.15 to 0.10 $L\cdot min^{-1}$ (31), respectively. Therefore, when examining the effectiveness of the supramaximal test it is important to include individual participant analysis to prevent misinterpretation due to responses being concealed when analysed on a group mean level.

Both males and females had significantly lower mean absolute $\dot{V}O_{2peak}$ values obtained in the supramaximal test than the incremental test, opposing the findings of Robben *et al.* (31) who found no significant difference between tests for either sex. This could be because of their much lower sample size, providing lower statistical power to detect small differences. However, although the differences were significant, effect sizes showed this difference was trivial for the males and small for the females. Our results also show a significantly lower mean absolute $\dot{V}O_{2peak}$ recorded in the supramaximal test compared with the incremental test for the non-overweight children and adolescents (5.2%) with a small effect size for this difference but not for the overweight children and adolescents (-1.5%) who had a trivial effect size. This contradicts the results of Bhammar et al.'s (7) study who reported $\dot{V}O_{2peak}$ was significantly greater in the verification test in both obese and non-obese children. In addition, this is the first study to assess the effect of CRF status on the $\dot{V}O_2$ response between the incremental and supramaximal test in a pediatric population. We found that there was a significant decrease in $\dot{V}O_{2peak}$ in the supramaximal test than the incremental test for all CRF statuses, although effect sizes showed the difference to be trivial for low and high CRF and small for average CRF. The magnitude of the decrease within the groups was consistent with when we separated the group by sex or body mass status. Astorino and DeRevere (4) demonstrated in adults that CRF may be related to the ability of the supramaximal test to verify "true" $\dot{V}O_{2max}$ with less fit individuals more likely to increase their $\dot{V}O_2$ in the supramaximal test than average or high fit individuals (4 low fit participants vs. 1 moderate/high fit participant respectively). In contrast, we found no significant differences between the different levels of CRF statuses on the percentage of children and adolescents who had their $\dot{V}O_{2max}$ verified through the supramaximal test. Astorino and DeRevere's (4) finding in adults may result from early termination of the initial ramp test in the low fit participants as they had a significantly lower TTE on the incremental test compared with moderate and high fit participants (9.1 ± 1.2 min vs. 10.4 ± 1.0 min and 10.8 ± 1.1 min, respectively). This might be due to lack of motivation from the low-fit participants, possibly due to their unfamiliarity with the demands of maximal intensity exercise (27). Despite these differences, the original findings in the current study indicate that the validity of the supramaximal test to verify $\dot{V}O_{2max}$ does not appear to be influenced by sex, body mass or CRF status in children and adolescents.

It has been noted that the manipulation of recovery period and its effect on $\dot{V}O_{2max}$ verification using the supramaximal test is an under researched area (35), and to our knowledge this is the first study to examine this concept in a pediatric group. Whilst it was not a systematically manipulated outcome, the rest period between the incremental and supramaximal test was not a significant predictor of $\dot{V}O_{2max}$ verification in the supramaximal test suggesting there is no effect of the duration of recovery on the measurement of $\dot{V}O_{2max}$ in a healthy pediatric population.

The low plateau attainment in the incremental test in this study (27%) is consistent with the pediatric literature (5, 33) and was consistently found between sex, body mass and CRF statuses, highlighting the need for the supramaximal verification test as the alternative method to identify $\dot{V}O_{2max}$ in children and adolescents. Similarly, Wood *et al.* (40) found overweight adults were no less likely to show a plateau than non-overweight adults. However, early treadmill work by Myers *et al.* (23) led to the suggestion that the occurrence of a $\dot{V}O_2$ plateau might be a random occurrence because, although all participants demonstrated a plateau in the initial incremental test, three of these did not plateau in the subsequent incremental test. The more recent findings provide some evidence to refute the suggestion by Myers *et al.* (23) since if it were a random occurrence, studies would be reporting different attainment levels of a plateau with some reporting lower plateau attainment and others reporting much higher plateau attainment. Our investigation into potential predictors of plateau attainment during the incremental test (e.g. age, sex, CRF status) did not find any significant predictors. Thus, we are unable to offer further explanation as to why plateau attainment is low in children and adolescent during the incremental test. Previous research has suggested that maturation may influence attainment of a $\dot{V}O_2$ plateau because almost double the number of adult males (23.8%) achieved a plateau compared with the boys at Tanner stages 1 or 2 (12.5%) (9), but we did not find maturity (somatic) status to be a significant predictor for attaining a plateau, nor for supramaximal test verification. However, this could be due to differences in sample size as Brown *et al.* (9) only studied 16 young boys, protocol differences since the study was conducted on a treadmill rather than a cycle ergometer or due to Brown *et al.* having a wider range of maturation statuses (comparing children Tanner stages 1-2 to adults Tanner stage 5) than the present study.

Murias *et al.* (22) recently stated that the supramaximal test should not be used as the gold standard in $\dot{V}O_{2max}$ measurement based on their analysis of adult males where no significant differences were found between $\dot{V}O_{2peak}$ observed in the incremental and supramaximal tests. Instead, the authors advocate the use of secondary criteria from the initial incremental test. However, in agreement with Bhammar *et al.* (7), our results do not show any of the primary or secondary objective criteria to have a sufficient level of both sensitivity and specificity to support their use to verify attainment of $\dot{V}O_{2max}$ in the incremental test in children and adolescents. Based on the use of a plateau criterion alone in the incremental test, only 27% of the population would have been deemed to have attained $\dot{V}O_{2max}$, but after the use of the supramaximal test, this increased to 88% regardless of sex, body mass and CRF status. Therefore, it is apparent that attainment of a $\dot{V}O_{2max}$ plateau in the incremental test is not an essential feature for $\dot{V}O_{2max}$ to be identified in children and adolescents. Additionally, the low AUCs from the ROC analyses for the primary and secondary criteria based on the incremental test (all <0.629), further demonstrates their poor ability to accurately validate $\dot{V}O_{2max}$ attainment, which does not support the recent recommendation by Murias *et al.*'s (22). Furthermore, although $\dot{V}O_{2max}$ is typically attained in the incremental test (88% in our sample) in children and adolescents, the attainment of "true" $\dot{V}O_{2max}$ is not certain until the supramaximal test has been performed because secondary objective criteria significantly underestimate $\dot{V}O_{2max}$ (5, 26). It is therefore essential that the supramaximal test is performed to ensure a valid measurement of $\dot{V}O_{2max}$, because even though the secondary objective criteria are often used in combination (15, 25), combining multiple poor methods does not make a good method to verify $\dot{V}O_{2max}$. This may be especially important in clinical groups or unfit populations due to their inexperience with performing maximal intensity exercise (27) and less experienced research teams may have lower validation rates with the supramaximal test. Consequently, our data support previous proposals for pediatric and adults groups (5, 27) that the use of primary and secondary objective criteria from the incremental test should be discontinued in favour of the use of the supramaximal test (5) when determining $\dot{V}O_{2max}$.

The major strength of this study is that, for the first time, the sample has been stratified based on sex, body mass and CRF statuses, made possible by our large sample size of 128 ostensibly healthy children and adolescents. Within the large sample, there was a broad range of CRF statuses (22.6-72.1 $mL \cdot kg^{-1} \cdot min^{-1}$) and maturation status'. However, although the overall sample size was large, it was lacking participants who were classed as overweight, especially for the girls – likely due to a self-selection bias for involvement in exercise studies. A further limitation of this study is that CRF status was determined using the ratio standard scaling for body mass, which may have resulted in misclassification for some participants. However, we are not aware of normative CRF data to classify CRF status using allometric scaling for body mass. Emerging data show that the supramaximal test is equally useful in clinical groups (34) and the variables investigated in this paper should be assessed in clinical and adult populations in case they are significant predictors of a $\dot{V}O_2$ plateau or supramaximal test verification. Additionally, the literature needs to address the issue of the remaining 12% of children who did not have their $\dot{V}O_{2max}$ verified in the supramaximal test, whether that is by investigating the utility of conducting a secondary supramaximal test on the same day or on a separate day, or whether a supplementary incremental and supramaximal test is required.

In conclusion, although only 27% had a plateau after the incremental test, the supramaximal test verified $\dot{V}O_{2max}$ in 88% of children and adolescents and was equally robust when participants were stratified for sex, body mass, maturation and CRF status. TTE on the supramaximal test was the only significant predictor of $\dot{V}O_{2max}$ being verified in the supramaximal test, with a longer TTE suggesting the initial incremental test was prematurely terminated (either by the experimenter or participant). The secondary objective criteria commonly used in the literature failed to have adequate levels of both sensitivity and specificity and their use in research should be discontinued. Results of this study support the use of the supramaximal test to verify $\dot{V}O_{2max}$ in a pediatric population.

References

1. Armstrong N, Welsman JR. Assessment and interpretation of aerobic fitness in children and adolescents. Exercise & Sport Sciences Reviews. 1994;22:435-76.
2. Armstrong N, Kirby BJ, McManus AM, Welsman JR. Aerobic fitness of prepubertal children. Annals of Human Biology. 1995;22(5):427-41.
3. Armstrong N, Welsman J, Winsley R. Is peak VO2 a maximal index of children's aerobic fitness? International Journal of Sports Medicine. 1996;17(5):356-9.
4. Astorino TA, DeRevere J. Efficacy of constant load verification testing to confirm VO2 max attainment. Clinical Physiology & Functional Imaging. 2017;38(4):703-9.
5. Barker AR, Williams CA, Jones AM, Armstrong N. Establishing maximal oxygen uptake in young people during a ramp cycle test to exhaustion. British Journal of Sports Medicine. 2011;45(6):498-503.
6. Barker AR, Williams CA, Tolfrey K, Fawkner S, Sandercock G. The BASES expert statement on measurement and interpretation of aerobic fitness in young people. The Sport & Exercise Scientist. 2013;37.
7. Bhammar DM, Stickford JL, Bernhardt V, Babb TG. Verification of maximal oxygen uptake in obese and nonobese children. Medicine & Science in Sports & Exercise. 2017;49(4):702-10.
8. Bongers BC, van Brussel M, Hulzebos EHJ, Takken T. Pediatric norms for cardiopulmonary exercise testing: In relation to sex and age. Second ed. Den Bosch: BoxPress; 2014.
9. Brown JD, Mahon AD, Plank DM. Attainment of maximal exercise criteria in boys and men. Journal of Sports Medicine & Physical Fitness. 2002;42(2):135-40.
10. Caspersen CJ, Powell KE, Christenson GM. Physical-activity, exercise, and physical-fitness - Definitions and distinctions for health-related research. Public Health Reports. 1985;100(2):126-31.
11. Cohen J. A power primer. Psychological Bulletin. 1992;112(1):155-9.
12. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. British Medical Journal. 2000;320(7244):1240-3.

13. Day JR, Rossiter HB, Coats EM, Skasick A, Whipp BJ. The maximally attainable VO2 during exercise in humans: the peak vs. maximum issue. Journal of Applied Physiology. 2003;95(5):1901-7.
14. Dencker M, Thorsson O, Karlsson MK, Linden C, Eiberg S, Wollmer P, et al. Gender differences and determinants of aerobic fitness in children aged 8-11 years. European Journal of Applied Physiology. 2007;99(1):19-26.
15. Duff DK, De Souza AM, Human DG, Potts JE, Harris KC. A novel treadmill protocol for exercise testing in children: the British Columbia Children's Hospital protocol. BMJ Open Sport & Exercise medicine. 2017;3(1):e000197-e.
16. Hill DW, Poole DC, Smith JC. The relationship between power and the time to achieve VO2max. Medicine & Science in Sports & Exercise. 2002;34(4):709-14.
17. 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.
18. Högström G, Nordström A, Nordström P. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. International Journal of Epidemiology. 2016;45(4):1159-68.
19. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. Continuing Education in Anaesthesia Critical Care & Pain. 2008;8(6):1743-816.
20. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;327(8476):307-10.
21. Moore SA, McKay HA, Macdonald H, Nettlefold L, Baxter-Jones ADG, Cameron N, et al. Enhancing a somatic maturity prediction model. Medicine & Science in Sports & Exercise. 2015;47(8):1755-64.
22. Murias JM, Pogliaghi S, Paterson DH. Measurement of a true VO2max during a ramp incremental test is not confirmed by a verification phase. Frontiers in Physiology. 2018;9(143).
23. Myers J, Walsh D, Buchanan N, Froelicher VF. Can maximal cardiopulmonary capacity be recognised by a plateau in oxygen-uptake. Chest. 1989;96(6):1312-6.
24. Ortega FB, Ruiz JR, Castillo MJ, Sjostrom M. Physical fitness in childhood and adolescence: A powerful marker of health. International Journal of Obesity. 2008;32(1):1-11.
25. Pivarnik JM, Dwyer MC, Lauderdale MA. The reliability of aerobic capacity VO2max testing in adolescent girls. Res Q Exerc Sport. 1996;67(3):345-8.
26. Poole DC, Wilkerson DP, Jones AM. Validity of criteria for establishing maximal VO2 uptake during ramp exercise tests. European Journal of Applied Physiology. 2008;102(4):403-10.
27. Poole DC, Jones AM. Measurement of the maximum oxygen uptake VO2max : VO2peak is no longer acceptable. Journal of Applied Physiology. 2017;122(4):997-1002.
28. Raine L, Drollette E, Kao S-C, Westfall D, Chaddock-Heyman L, Kramer AF, et al. The Associations between Adiposity, Cognitive Function, and Achievement in Children. Medicine & Science in Sports & Exercise. 2018;50(9):1868-74.
29. Rivera-Brown A, Rivera MA, Frontera WR. Reliability of VO2max in Adolescent Runners: A Comparison Between Plateau Achievers and Nonachievers. Pediatric Exercise Science. 1995;7:203-10.
30. Rivera-Brown AM, Rivera MA, Frontera WR. Applicability of Criteria for V̇O2max in Active Adolescents. Pediatric Exercise Science. 1992;4(4):331-9.
31. Robben KE, Poole DC, Harms CA. Maximal oxygen uptake validation in children with expiratory flow limitation. Pediatric Exercise Science. 2013;25(1):84-100.
32. Rowland TW, Cunningham LN. Oxygen uptake plateau during maximal treadmill exercise in children. Chest. 1992;101(2):485-9.
33. Rowland TW. Does peak VO2 reflect VO2max in children - Evidence from supramaximal testing. Medicine & Science in Sports & Exercise. 1993;25(6):689-93.

34. Saynor ZL, Barker AR, Oades PJ, Williams CA. A protocol to determine valid VO2max in young cystic fibrosis patients. Journal of Science & Medicine in Sport. 2013;16(6):539-44.
35. Schaun GZ. The maximal oxygen uptake verification phase: A light at the end of the tunnel? Sports Medicine - Open. 2017;3(1):44-.
36. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of cardio-respiratory performance. Journal of Applied Physiology. 1955;8(1):73-80.
37. Welsman J, Bywater K, Farr C, Welford D, Armstrong N. Reliability of peak VO2 and maximal cardiac output assessed using thoracic bioimpedance in children. European Journal of Applied Physiology. 2005;94(3):228-34.
38. Welsman JR, Armstrong N, Nevill AM, Winter EM, Kirby BJ. Scaling peak VO2 for differences in body size. Medicine and Science in Sports and Exercise. 1996;28(2):259-65.
39. Welsman JR, Armstrong N. Statistical techniques for interpreting body size-related exercise performance during growth. Pediatric Exercise Science. 2000;12(2):112-27.
40. Wood RE, Hills AP, Hunter GR, King NA, Byrne NM. VO2max in overweight and obese adults: Do they meet the threshold criteria? Medicine & Science in Sports & Exercise. 2010;42(3):470-7.

Table 1 – Participant characteristics and physiological responses to the incremental and supramaximal tests

	Males			Females		
	Overall (*n* =76)	Non-overweight (*n* =65)	Overweight (*n* =11)	Overall (*n* =52)	Non-overweight (*n*=44)	Overweight (*n* =8)
Age (y)	13.3 ± 1.9 [b]	13.1 ± 1.9 [b]**	14.6 ± 1.1 [b]	13.9 ± 1.6 [b]	14.0 ± 1.6 [b]	13.7 ± 1.5 [b]
Stature (m)	1.61 ± 0.15 [b]	1.60 ± 0.16 [b]**	1.71 ± 0.08 [b]	1.60 ± 0.10 [b]	1.61 ± 0.09 [b]	1.60 ± 0.12 [b]
Body mass (kg)	52.9 ± 16.7 [b]	48.7 ± 13.4 [b]**	78.0 ± 11.3 [b]	54.2 ± 10.3 [b]	52.2 ± 8.0 [b]***	65.5 ± 14.6 [b]
BMI ($kg\cdot m^{-2}$)	19.8 ± 3.6 [b]*	18.7 ± 2.1 **	26.7 ± 3.0	20.9 ± 2.8 [b]	20.1 ± 1.9 [b]***	25.4 ± 3.1 [b]
APHV (y)	0.9 ± 1.7*	0.8 ± 1.7**	2.0 ± 1.0	1.8 ± 1.4	1.8 ± 1.4	1.6 ± 1.5
CRF status (% low, average, high)	29, 45, 26	22, 48, 31	73, 27, 0	44, 50, 6	39, 55, 7	75, 25, 0
Ratio standard $\dot{V}O_{2max}$ ($mL\cdot kg^{-1}\cdot min^{-1}$)	49 ± 10 [b]*	50 ± 10 **	40 ± 7	38 ± 7 [b]	38 ± 6 [b]***	31 ± 5 [b]
Allometrically scaled $\dot{V}O_{2max}$ ($mL\cdot kg^{-0.66}\cdot min^{-1}$)	187 ± 33 [b]*	189 ± 34	181 ± 31	147 ± 23 [b]	150 ± 22 [b]***	128 ± 17 [b]
Peak ramp $\dot{V}O_2$ ($L\cdot min^{-1}$)	2.48 ± 0.73 [b]*	2.39 ± 0.70 [b]**	3.03 ± 0.68 [b]	1.96 ± 0.31 [b]	1.96 ± 0.30	1.94 ± 0.40

Peak supramaximal $\dot{V}O_2$ (L·min^{-1})	2.36 ± 0.72 [b]*	2.24 ± 0.66 [b]**	3.04 ± 0.70 [b]	1.89 ± 0.34 [b]	1.90 ± 0.35 [b]	1.86 ± 0.31 [b]
Peak ramp HR (beats·min^{-1})	193 ± 10[a]	194 ± 10[a]	191 ± 10[a]	194 ± 7[a]	195 ± 7[a]	192 ± 6
Peak supramaximal HR (beats·min^{-1})	187 ± 11[a]	186 ± 11[a]	188 ± 9[a]	190 ± 8[a]	191 ± 8[a]	186 ± 7
Peak ramp RER	1.19 ± 0.10 *	1.19 ± 0.10 [b]	1.19 ± 0.08 [b]	1.26 ± 0.09	1.27 ± 0.09	1.24 ± 0.10
Peak supramaximal RER	1.18 ± 0.12 *	1.18 ± 0.12	1.21 ± 0.11	1.24 ± 0.14	1.25 ± 0.14 [b]	1.19 ± 0.13 [b]
Ramp TTE (s)	568 ± 126 [b]*	570 ± 125	557 ± 139	483 ± 110 [b]	489 ± 113 [b]	452 ± 96 [b]
Supramaximal TTE (s)	98 ± 25 [b]	94 ± 21 [b]**	119 ± 35 [b]	99 ± 20 [b]	100 ± 19	91 ± 23

Data presented as mean ± standard deviation (SD). BMI = body mass index. APHV = age from peak height velocity. CRF = cardiorespiratory fitness. HR = heart rate. RER = respiratory exchange ratio. TTE = time to exhaustion. [a] = denotes incomplete data. [b] = denotes data log transformed for t-test analysis. * = significant difference of males compared with females. ** = significant difference of non-overweight males compared with overweight males. *** = significant difference of non-overweight females compared with overweight females.

616 **Table 2** – Sensitivity and specificity analysis of primary and secondary objective
617 criteria to verify $\dot{V}O_{2max}$

	Plateau achieved in incremental test?	RER>1.0	RER>1.1	HR_{max}> 85% age predicted maximum	HR_{max}> 195 beats·min^{-1}	HR_{max}> 95% age predicted maximum
Sensitivity (%)	24.1	7.1	17.9	10.4	14.6	15.6
Specificity (%)	50.0	100.0	87.5	78.6	92.9	100.0
PPV (%)	77.1	100.0	90.9	76.9	93.3	100.0
NPV (%)	8.6	13.3	13.2	11.3	92.9	14.7
AUC	0.629	0.536	0.527	0.555	0.537	0.578

618 PPV = positive predictive value, NPV = negative predictive value, AUC = area under
619 receiver operator characteristic curve.