

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

3 **Supramaximal Test Verification of Pediatric $\dot{V}O_{2\max}$**

4 Kate M. Sansum¹, Max E. Weston¹, Bert Bond¹, Emma J. Cockcroft², Amy O'Connor¹,
5 Owen W. Tomlinson¹, Craig A. Williams¹ & Alan R. Barker^{1*}

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

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

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

11 **Abstract**

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

23 Supramaximal test time to exhaustion predicted supramaximal test $\dot{V}O_{2\max}$ verification
24 ($P=0.040$). Primary and secondary objective criteria had insufficient sensitivity (7.1-
25 24.1%) and specificity (50-100%) to verify $\dot{V}O_{2\max}$. **Conclusion:** The utility of
26 supramaximal testing to verify $\dot{V}O_{2\max}$ is not affected by sex, body mass or CRF status.
27 Supramaximal testing should replace secondary objective criteria to verify $\dot{V}O_{2\max}$.

28

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

31

32

33

34

35

36

37

38

39

40

41

42

43

44 **Introduction**

45 Maximal oxygen uptake ($\dot{V}O_{2\max}$), typically expressed in relation to a measure of body
46 size, is the “gold-standard” measure of cardiorespiratory fitness (CRF) (10). A valid
47 measurement of $\dot{V}O_{2\max}$ is important in children and adolescents as a high CRF in youth
48 is associated with a lower risk of cardiovascular disease in youth (24), a reduced risk of
49 myocardial infarction (17) and all-cause mortality (18) in adult life. Traditionally, the
50 presence of a plateau of $\dot{V}O_2$ at, or close to exhaustion, during incremental exercise has
51 been used as the primary criterion for attainment of $\dot{V}O_{2\max}$ (36). However, as only
52 between 10-50% of children display a plateau across different testing protocols (5, 29,
53 30, 32), with the reasons behind this still being unclear, the term $\dot{V}O_{2\text{peak}}$ is routinely used
54 to denote the highest $\dot{V}O_2$ recorded without a plateau (1). Secondary objective criteria
55 (e.g. respiratory exchange ratio (RER) and maximal heart rate (HR_{\max}) thresholds) (3, 33)
56 are therefore often used to verify that the $\dot{V}O_{2\text{peak}}$ attained was a “true” $\dot{V}O_{2\max}$ for children
57 and adolescents but significantly underestimate $\dot{V}O_{2\max}$ (5). While, the validity of
58 secondary objective criteria has recently been challenged (5, 26), their use is still
59 commonplace in contemporary pediatric research (e.g. 15, 28).

60 To overcome the validity issues with the primary and secondary criteria, the 1990s saw
61 the emergence of a supramaximal test to verify that the $\dot{V}O_{2\text{peak}}$ that had been achieved in
62 the incremental test is a “true” $\dot{V}O_{2\max}$. This requires participants to exercise at a power
63 output greater than the maximal power output achieved during the incremental test (3, 33)
64 and is a variation of the original protocol proposed by Taylor *et al.* (36). However, the
65 supramaximal verification tests were initially conducted on separate days (3, 33), which

66 may not be feasible, due to logistical requirements of supplementary laboratory visits.
67 Recently, it has been shown that children can successfully perform the supramaximal
68 verification test on the same day as the incremental test, following a short rest of 10-15
69 min following the incremental test (5, 7, 31), and is now the recommended protocol for
70 $\dot{V}O_{2max}$ determination (6, 27).

71 While the supramaximal verification test is an elegant solution to determine $\dot{V}O_{2max}$ in
72 children and adolescents, not all participants have their $\dot{V}O_{2max}$ confirmed in the
73 supramaximal test. Between 8-26% of participants have been reported as not having
74 $\dot{V}O_{2max}$ verified (5, 31), but a recent paper reported a non-verification rate of 100% of
75 children (7). Previous studies have suggested that non-verification may be related to
76 factors such as sex, body mass, CRF and/or maturity status (4, 7, 9) but the current
77 pediatric literature is based on small sample sizes ($n = 9-40$) (3, 5, 31, 33), which limits
78 examination of these variables on $\dot{V}O_{2max}$ verification. Consequently, male and female
79 data have been combined for analysis (5), despite known sex-differences in $\dot{V}O_{2max}$ and
80 body composition (2, 3). Recently it has been shown that adults with low CRF were less
81 likely to have their $\dot{V}O_{2max}$ confirmed in the supramaximal test than those with a higher
82 CRF status (4), but this has not been investigated in children and adolescents. Few studies
83 (4, 31) have compared individual $\dot{V}O_{2peak}$ values from the incremental test to the
84 supramaximal test, instead comparing the group means (3, 33) which may be misleading
85 as $\dot{V}O_{2max}$ testing is conducted on an individual basis. The effect of body mass status on
86 the verification of $\dot{V}O_{2max}$ using supramaximal exercise has only been studied by
87 Bhammar *et al.* (7) who found both obese ($n = 9$) and non-obese ($n = 9$) children to have
88 a significantly greater $\dot{V}O_{2peak}$ in the supramaximal test. Brown *et al.* (9) have also
89 suggested that maturity status may influence plateau attainment during an incremental

90 test, with 23.8% of men achieving a plateau compared with 12.5% of boys, but this has
91 not been investigated in the context of the supramaximal verification test. Finally, the
92 effect of the time to exhaustion (TTE) in the supramaximal test on the utility of the
93 supramaximal verification test is worthy of consideration. TTE in studies where between
94 74-92% of children had $\dot{V}O_{2max}$ verified is reported to be between 60 to 90 s in duration
95 (5, 31). However, Bhammar *et al.* (7) reported that no children had their $\dot{V}O_{2max}$ verified
96 by the supramaximal verification test and reported a TTE in excess of 125 s. Conversely,
97 a short TTE (e.g. of less than 60 s) could indicate that fatigue is reached before attainment
98 of $\dot{V}O_{2max}$, possibly due to insufficient effort or because the intensity was too high for the
99 $\dot{V}O_2$ kinetic response to attain $\dot{V}O_{2max}$ (16).

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

108 **Methods**

109 Participants:

110 Existing data from our laboratory were pooled and retrospectively analysed to produce a
111 sample of 128 healthy children and adolescents. Only data from 13 participants that form
112 the final sample have previously been published elsewhere to examine the validity of the

113 supramaximal test (5). All data were collected as part of studies which originally were
114 granted ethics approval by institutional and NHS ethics committees (where relevant).
115 Inclusion criteria for this study were: 1) 8-<18 years old and; 2) $\dot{V}O_{2\max}$ assessed using a
116 combined incremental and supramaximal test protocol, conducted on the same day; 3)
117 ostensibly healthy participants and; 4) cycling modality. All children and their parent(s)
118 or guardian(s) gave informed assent and consent, respectively, to participate in the
119 original studies.

120 Anthropometry:

121 Body mass (Seca 877, Seca Ltd, Birmingham, UK) was measured to the nearest 0.1 kg
122 and stature (Harpenden, Holtain Ltd, Crymych, UK) was measured to the nearest 0.01 m.
123 Body mass index (BMI) was calculated, and age-appropriate criteria were used to classify
124 participants into non-overweight and overweight/obese categories (12). Maturity
125 (somatic) offset from the age of peak height velocity (APHV) was calculated through the
126 equations by Moore *et al.* (21), which have been validated in two external samples where
127 90% of predictions are within ± 1 year. Pre-peak height velocity (PHV) children were
128 defined as >-1 year from PHV, circa-PHV children were -1 to 1 year from PHV and post-
129 PHV children were >1 year from PHV.

130 Incremental and supramaximal test protocols:

131 A combined incremental-ramp and supramaximal test to exhaustion was used to
132 determine $\dot{V}O_{2\max}$ (5). Participants were instructed to cycle on an electronically braked
133 ergometer (Lode Excalibur, Groningen, The Netherlands) at a constant self-selected
134 cadence between 70 and 90 revolutions per minute throughout the tests. Participants
135 cycled for ~ 3 min (range 1 to 3 min) at 20 watts (W) to warm up before immediately

136 commencing the incremental-ramp protocol where the power output increased by 10-30
137 W·min⁻¹, depending on the participants' age and body size, to attempt to elicit exhaustion
138 between 8 and 12 min. Exhaustion was defined as a decrease in cadence below 60
139 revolutions per minute for 5 consecutive seconds, despite strong verbal encouragement.
140 This was followed by 3 min 30 s (range 0 to 10 min) cool down cycling at 20 W. A rest
141 period of ~ 25 min (range = 5 to 84 min) followed before the commencement of the
142 supramaximal test, which began with a warm up of 3 min at 20 W. The resistance was
143 then increased in a "step" transition to either ~105% (*n* = 117) or ~110% (*n* = 11) of the
144 peak power achieved in the incremental test and participants were required to cycle to
145 exhaustion. Following the test, the participant completed a cool down cycling at 20 W.
146 The measurement of $\dot{V}O_{2\text{peak}}$ from the ramp-incremental test to exhaustion has a
147 coefficient of variation of 4.1% (37).

148 *Gas collection and analysis:*

149 Pulmonary gas exchange and heart rate (HR) were measured using online systems (Cortex
150 Metalyzer III B, Leipzig, Germany: *n* =106; EX671; Morgan Medical, Kent, UK,
151 combined with mass spectrometry and a turbine flow meter VMM-401; Interdace
152 Associates, Laguna Niguel, California, USA: *n* =13; and Medgraphics Cardiorespiratory
153 Diagnostics, Express Series, Gloucester, UK: *n* =9). All systems were appropriately
154 calibrated for gas and volume before each test as per manufacturers' recommendations.
155 $\dot{V}O_{2\text{max}}$ was accepted as the highest 10-15 second average of $\dot{V}O_2$ recorded in either the
156 incremental or supramaximal tests (5). To control for body-size, both the ratio standard
157 and allometric (via log-linear regression, (38)) models were used to scale $\dot{V}O_{2\text{max}}$ for body
158 mass. Although allometric procedures are superior for scaling $\dot{V}O_2$ (39), normative data
159 are unavailable to classify the children and adolescents into CRF groups. Therefore, the

160 ratio standard method to scale for body mass was used to group participants into CRF
161 statuses of low, average and high CRF based on age and sex related normative values (8).
162 Low CRF was defined as > 1 standard deviation (SD) below the age and sex specific
163 mean normative value, average CRF was defined as falling within 1 SD either side of the
164 age and sex specific mean normative value, and high CRF was defined as > 1 SD above
165 the age and sex specific mean normative value.

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

167 The methods proposed by Day *et al.* (13) were used to define a plateau and classify the
168 $\dot{V}O_2$ responses during incremental exercise into a linear, acceleration or deceleration
169 profile using GraphPad Prism (GraphPad Software Inc., San Diego, California, USA). A
170 linear regression of the $\dot{V}O_2$ -intensity relationship was plotted over the 'linear' portion of
171 the $\dot{V}O_2$ profile, where the data points from the first 2 minutes and the last 3 minutes of
172 exercise were excluded. The linear function was then extrapolated and compared to the
173 residuals to analyse the $\dot{V}O_2$ profile at exhaustion for an accelerated, decelerated (i.e.
174 plateau) or linear response. An accelerated profile required the positive residual to be
175 $\geq 5\%$ of the peak power projected $\dot{V}O_2$ whereas a decelerated profile required the negative
176 residual to be $\geq 5\%$ of the projected $\dot{V}O_2$. A linear response was classified by residuals
177 that were $< 5\%$ of the peak power projected $\dot{V}O_2$, either side of the extrapolated line.

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

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

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

187 Statistical analyses:

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

205 combined, and based on sex, body mass and CRF status. Checks for proportional bias
206 were undertaken using Pearson's correlation and satisfied for all Bland-Altman plots.

207 Separate logistic regression analyses were run to identify 1) significant predictors of
208 plateau attainment in the incremental test and; 2) verification of $\dot{V}O_{2max}$ through the use
209 of the supramaximal test. The variables tested in both models were age (years), sex, body
210 mass status (overweight/obese or non-overweight), APHV (pre-, circa-, post-PHV),
211 $\dot{V}O_{2max}$ expressed using the ratio standard and allometric methods, CRF status (low,
212 average or high) and incremental test TTE (s). The supramaximal test TTE (s),
213 supramaximal test intensity (% of peak power attained in the incremental test) and rest
214 period between the incremental and supramaximal test (s) were also included for
215 predicting $\dot{V}O_{2max}$ verification using the supramaximal test. Variables were entered using
216 the backward stepwise (likelihood ratio) method.

217 Primary and secondary objective criteria from the incremental test were assessed for their
218 sensitivity (ability to correctly identify attainment of "true" $\dot{V}O_{2max}$) and specificity
219 (ability to correctly identify non-attainment of "true" $\dot{V}O_{2max}$) to verify $\dot{V}O_{2peak}$ in the
220 incremental test as $\dot{V}O_{2max}$ when compared to the supramaximal test verification method.
221 Each of the criteria was also assessed for their positive and negative predictive value i.e.
222 the likelihood that a positive or negative result from the criteria for attainment of "true"
223 $\dot{V}O_{2max}$ in the incremental test is the correct result. The equations below (19) were used
224 to calculate sensitivity, specificity and the positive and negative predictive value for each
225 criteria. The ability of each criterion to confirm $\dot{V}O_{2max}$ was assessed, allowing for 5%
226 error at an individual level and compared to whether the supramaximal test was able to

227 verify $\dot{V}O_{2max}$. Receiver operator characteristic (ROC) curves (1-specificity vs sensitivity)
228 were also used to calculate the area under the curve (AUC).

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

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

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

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

233 **Results**

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

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

250 Incremental and supramaximal responses: whole group

251 Analysis of individual participant $\dot{V}O_2$ -intensity profiles revealed 27% ($n= 35$) of
252 participants demonstrated a $\dot{V}O_2$ plateau during the incremental test, 14% ($n=18$) had an
253 accelerated $\dot{V}O_2$ profile and 59% ($n=75$) had a linear $\dot{V}O_2$ profile. When comparing the
254 $\dot{V}O_{2peak}$ values obtained from the incremental and supramaximal test on an individual
255 basis, 88% ($n=112$) of children and adolescents had their $\dot{V}O_{2peak}$ in the incremental test
256 verified as their “true” $\dot{V}O_{2max}$. For the remaining 12% who did not have $\dot{V}O_{2max}$ verified,
257 the $\dot{V}O_{2peak}$ recorded was between 6 and 23% greater than that recorded in the incremental
258 test.

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

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

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

267 Incremental and supramaximal responses: influence of sex

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

271 $\dot{V}O_{2max}$ between males and females (89%; $n = 68$ vs. 85%; $n = 44$; $P=0.41$).

272 The mean absolute $\dot{V}O_{2peak}$ of the males recorded in the incremental test (2.48 ± 0.73
273 $L \cdot min^{-1}$) was greater than the supramaximal test ($2.36 \pm 0.72 L \cdot min^{-1}$; $P<0.001$; ES =
274 0.17). Likewise, the mean absolute $\dot{V}O_{2peak}$ recorded for the females in the incremental
275 test was higher than in the supramaximal test ($1.96 \pm 0.31 L \cdot min^{-1}$ and $1.89 \pm 0.34 L \cdot min^{-1}$,
276 respectively; $P=0.007$; ES = 0.22). Incremental and supramaximal test $\dot{V}O_{2peak}$ values
277 were correlated (males $r=0.95$ and females $r=0.85$; both $P<0.001$).

278 Figure 1 depicts the Bland-Altman plots for absolute (1B and 1C) and percentage (1E and
279 1F) difference in $\dot{V}O_{2peak}$ recorded between the incremental and supramaximal tests for
280 each sex. Mean absolute and relative bias for the males was $-0.12 L \cdot min^{-1}$ and -5.4% with
281 LoA as -0.58 to $0.33 L \cdot min^{-1}$ and -23 to 13% , respectively. The absolute and relative
282 mean bias for the females was $-0.06 L \cdot min^{-1}$ and -3.5% with LoA as -0.41 to $0.29 L \cdot min^{-1}$
283 and -21 to 14% , respectively.

284 Incremental and supramaximal responses: influence of body mass status

285 No difference was found between the proportion of $\dot{V}O_2$ plateau observations between
286 the non-overweight and overweight children and adolescents (26%; $n = 28$ vs. 37%; $n =$
287 7; $P=0.31$). Furthermore, no difference was found between the proportion of non-
288 overweight compared with overweight children and adolescents who had their $\dot{V}O_{2max}$
289 verified in the supramaximal test (89%; $n = 97$ vs. 79%; $n = 15$; $P=0.22$).

290 The mean absolute $\dot{V}O_{2peak}$ for non-overweight children and adolescents was greater in
291 the incremental test compared with the supramaximal test ($2.27 \pm 0.63 L \cdot min^{-1}$ vs. $2.13 \pm$
292 $0.54 L \cdot min^{-1}$; $P<0.001$; ES = 0.24). In contrast, the mean absolute $\dot{V}O_{2peak}$ was not

293 different for the overweight children and adolescents between the incremental and
294 supramaximal tests ($2.57 \pm 0.79 \text{ L} \cdot \text{min}^{-1}$ vs. $2.54 \pm 0.81 \text{ L} \cdot \text{min}^{-1}$; $P=0.65$; $ES = 0.04$). The
295 $\dot{V}O_{2\text{peak}}$ recorded in the incremental and supramaximal tests were correlated for each
296 group (non-overweight children and adolescents $r=0.94$, overweight children and
297 adolescents $r=0.95$; both $P<0.001$).

298 Figure 2 displays the absolute (2A & 2B) and percentage (2C & 2D) differences in the
299 $\dot{V}O_{2\text{peak}}$ from the incremental and supramaximal tests for non-overweight and overweight
300 children and adolescents. The absolute and relative mean bias for non-overweight
301 children and adolescents was $-0.11 \text{ L} \cdot \text{min}^{-1}$ and -5.2% , and LoA were -0.51 to $0.29 \text{ L} \cdot \text{min}^{-1}$
302 1 and -22 to 12% , respectively. Mean absolute and relative bias for overweight children
303 and adolescents was $-0.03 \text{ L} \cdot \text{min}^{-1}$ and -1.5% with LoA as -0.54 to $0.49 \text{ L} \cdot \text{min}^{-1}$ and -21
304 to 18% , respectively.

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

306 Incremental and supramaximal responses: influence of CRF status

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

313 Mean absolute $\dot{V}O_{2\text{peak}}$ was higher in the incremental compared with the supramaximal
314 test for the low ($2.08 \pm 0.62 \text{ L} \cdot \text{min}^{-1}$ vs. $2.02 \pm 0.64 \text{ L} \cdot \text{min}^{-1}$; $P=0.007$; $ES = 0.10$), average
315 ($2.38 \pm 0.55 \text{ L} \cdot \text{min}^{-1}$ vs. $2.26 \pm 0.50 \text{ L} \cdot \text{min}^{-1}$; $P=0.001$; $ES = 0.23$) and high (2.33 ± 0.87

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

319 Figure 3 displays the absolute (3A, 3B & 3C) and relative (3D, 3E & 3F) differences in
320 $\dot{V}O_{2\text{peak}}$ from the incremental and supramaximal tests for low (3A & 3D), average (3B &
321 3E) and high (3C & 3F) CRF groups. The absolute and relative mean bias for the low
322 CRF group was -0.06 L·min⁻¹ and -3.3% with LoA as -0.37 to 0.25 L·min⁻¹ and -18 and
323 12% . Average CRF absolute and relative mean bias was -0.12 L·min⁻¹ and -5.1% , and
324 LoA were -0.62 to 0.38 L·min⁻¹ and -25 to 14% . For the high CRF group, the absolute
325 and relative mean bias was -0.12 L·min⁻¹ and -6.0% with LoA as -0.49 to 0.24 L·min⁻¹
326 and -23 to 11% .

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

328 Predicting plateau attainment in the incremental test:

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

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

332 TTE on the supramaximal test was the only predictor of whether the supramaximal test
333 can verify the $\dot{V}O_{2\text{peak}}$ attained in the incremental test as $\dot{V}O_{2\text{max}}$ ($P=0.040$; odds ratio =
334 0.978 ; 95% confidence limits = $0.958-0.999$). A longer supramaximal test TTE decreased
335 the likelihood that $\dot{V}O_{2\text{peak}}$ would be verified as $\dot{V}O_{2\text{max}}$ ($\beta=-0.022$; standard error = 0.011).

336 The regression equation below predicts supramaximal test verification.

337 Supramaximal test verification (Y) = $4.212 + (\text{Supramaximal Test TTE [s]} * -0.022)$

338 *Sensitivity and specificity of primary and secondary objective criteria:*

339 Table 2 displays the sensitivity and specificity analysis on the primary and secondary
340 objective criteria compared with the supramaximal test, as well as the positive and
341 negative predictive values. All criteria had low sensitivity (7.1-24.1%) but the majority
342 had high specificity (78.6-100%), apart from the plateau attainment (50%). Both primary
343 and secondary criteria had high positive predictive values (77.1-100%). By contrast,
344 negative predictive values were low for all criteria (8.6-14.7%) excluding the $HR_{max} >$
345 $195 \text{ beats}\cdot\text{min}^{-1}$ (92.9%). The AUC were low, ranging from 0.527 to 0.629.

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

347 **Discussion**

348 The main findings of the study were: 1) 88% of children and adolescents had their
349 absolute $\dot{V}O_{2max}$ verified in the supramaximal test which had a tendency to result in a ~
350 5% decrease in absolute $\dot{V}O_{2peak}$ in most, but not all of the sample; 2) the utility of the
351 supramaximal test to verify $\dot{V}O_{2max}$ was similar when stratified for sex, body mass or CRF
352 status; 3) TTE on the supramaximal test was the only significant predictor of $\dot{V}O_{2peak}$
353 being verified as $\dot{V}O_{2max}$ in the supramaximal test whereas there were no significant
354 predictors of plateau attainment in the ramp test and; 4) primary (plateau) and secondary
355 (HR and RER thresholds) objective criteria from the incremental test have insufficient
356 sensitivity and specificity to validate attainment of $\dot{V}O_{2max}$ in children.

357 The majority of our participants had their $\dot{V}O_{2max}$ verified in the supramaximal test (88%),
358 which is line with most (5, 31) but not all (7) of the literature. For example, recent findings
359 (7) showed that none of the obese and non-obese children had their $\dot{V}O_{2max}$ verified in the
360 supramaximal test. This difference in findings may be due to methodological differences

361 (e.g. their use of Douglas bags), or the smaller sample size (7). Alternatively, it could be
362 because the participants in Bhammar *et al.*'s study (7) had a supramaximal TTE of greater
363 than two minutes, which is uncommon in both the adult and pediatric literature (4, 5, 31)
364 and not in line with our findings (whole group = 98 ± 23 s, with 11 participants >120 s).
365 In the current study, supramaximal test TTE was a significant negative predictor for the
366 supramaximal test confirming attainment of $\dot{V}O_{2max}$ which could explain the lack of
367 verification of $\dot{V}O_{2max}$ observed in Bhammar *et al.*'s (7) study. The longer TTE may
368 reflect that the incremental test was sub-optimal and terminated early before $\dot{V}O_{2max}$ was
369 attained. However, it should be noted that although statistically significant, the finding
370 that the supramaximal test TTE was a significant predictor of $\dot{V}O_{2max}$ verification is
371 unlikely to be meaningful due to significance level in the logistic regression being
372 $P=0.040$ and CLs confidence limits = 0.958-0.999. Furthermore, we were not able to
373 identify a cut-off threshold for supramaximal test non-verification based on TTE.

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

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

392 Both males and females had significantly lower mean absolute $\dot{V}O_{2peak}$ values obtained in
393 the supramaximal test than the incremental test, opposing the findings of Robben *et al.*
394 (31) who found no significant difference between tests for either sex. This could be
395 because of their much lower sample size, providing lower statistical power to detect small
396 differences. However, although the differences were significant, effect sizes showed this
397 difference was trivial for the males and small for the females. Our results also show a
398 significantly lower mean absolute $\dot{V}O_{2peak}$ recorded in the supramaximal test compared
399 with the incremental test for the non-overweight children and adolescents (5.2%) with a
400 small effect size for this difference but not for the overweight children and adolescents (-
401 1.5%) who had a trivial effect size. This contradicts the results of Bhammar *et al.*'s (7)
402 study who reported $\dot{V}O_{2peak}$ was significantly greater in the verification test in both obese
403 and non-obese children. In addition, this is the first study to assess the effect of CRF status
404 on the $\dot{V}O_2$ response between the incremental and supramaximal test in a pediatric
405 population. We found that there was a significant decrease in $\dot{V}O_{2peak}$ in the supramaximal
406 test than the incremental test for all CRF statuses, although effect sizes showed the
407 difference to be trivial for low and high CRF and small for average CRF. The magnitude
408 of the decrease within the groups was consistent with when we separated the group by

409 sex or body mass status. Astorino and DeRevere (4) demonstrated in adults that CRF may
410 be related to the ability of the supramaximal test to verify “true” $\dot{V}O_{2max}$ with less fit
411 individuals more likely to increase their $\dot{V}O_2$ in the supramaximal test than average or
412 high fit individuals (4 low fit participants vs. 1 moderate/high fit participant respectively).
413 In contrast, we found no significant differences between the different levels of CRF
414 statuses on the percentage of children and adolescents who had their $\dot{V}O_{2max}$ verified
415 through the supramaximal test. Astorino and DeRevere’s (4) finding in adults may result
416 from early termination of the initial ramp test in the low fit participants as they had a
417 significantly lower TTE on the incremental test compared with moderate and high fit
418 participants (9.1 ± 1.2 min vs. 10.4 ± 1.0 min and 10.8 ± 1.1 min, respectively). This
419 might be due to lack of motivation from the low-fit participants, possibly due to their
420 unfamiliarity with the demands of maximal intensity exercise (27). Despite these
421 differences, the original findings in the current study indicate that the validity of the
422 supramaximal test to verify $\dot{V}O_{2max}$ does not appear to be influenced by sex, body mass
423 or CRF status in children and adolescents.

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

431 The low plateau attainment in the incremental test in this study (27%) is consistent with
432 the pediatric literature (5, 33) and was consistently found between sex, body mass and
433 CRF statuses, highlighting the need for the supramaximal verification test as the
434 alternative method to identify $\dot{V}O_{2\max}$ in children and adolescents. Similarly, Wood *et al.*
435 (40) found overweight adults were no less likely to show a plateau than non-overweight
436 adults. However, early treadmill work by Myers *et al.* (23) led to the suggestion that the
437 occurrence of a $\dot{V}O_2$ plateau might be a random occurrence because, although all
438 participants demonstrated a plateau in the initial incremental test, three of these did not
439 plateau in the subsequent incremental test. The more recent findings provide some
440 evidence to refute the suggestion by Myers *et al.* (23) since if it were a random occurrence,
441 studies would be reporting different attainment levels of a plateau with some reporting
442 lower plateau attainment and others reporting much higher plateau attainment. Our
443 investigation into potential predictors of plateau attainment during the incremental test
444 (e.g. age, sex, CRF status) did not find any significant predictors. Thus, we are unable to
445 offer further explanation as to why plateau attainment is low in children and adolescent
446 during the incremental test. Previous research has suggested that maturation may
447 influence attainment of a $\dot{V}O_2$ plateau because almost double the number of adult males
448 (23.8%) achieved a plateau compared with the boys at Tanner stages 1 or 2 (12.5%) (9),
449 but we did not find maturity (somatic) status to be a significant predictor for attaining a
450 plateau, nor for supramaximal test verification. However, this could be due to differences
451 in sample size as Brown *et al.* (9) only studied 16 young boys, protocol differences since
452 the study was conducted on a treadmill rather than a cycle ergometer or due to Brown *et*
453 *al.* having a wider range of maturation statuses (comparing children Tanner stages 1-2 to
454 adults Tanner stage 5) than the present study.

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

479 intensity exercise (27) and less experienced research teams may have lower validation
480 rates with the supramaximal test. Consequently, our data support previous proposals for
481 pediatric and adults groups (5, 27) that the use of primary and secondary objective criteria
482 from the incremental test should be discontinued in favour of the use of the supramaximal
483 test (5) when determining $\dot{V}O_{2max}$.

484 The major strength of this study is that, for the first time, the sample has been stratified
485 based on sex, body mass and CRF statuses, made possible by our large sample size of
486 128 ostensibly healthy children and adolescents. Within the large sample, there was a
487 broad range of CRF statuses (22.6-72.1 mL·kg⁻¹·min⁻¹) and maturation status'. However,
488 although the overall sample size was large, it was lacking participants who were classed
489 as overweight, especially for the girls – likely due to a self-selection bias for involvement
490 in exercise studies. A further limitation of this study is that CRF status was determined
491 using the ratio standard scaling for body mass, which may have resulted in
492 misclassification for some participants. However, we are not aware of normative CRF
493 data to classify CRF status using allometric scaling for body mass. Emerging data show
494 that the supramaximal test is equally useful in clinical groups (34) and the variables
495 investigated in this paper should be assessed in clinical and adult populations in case they
496 are significant predictors of a $\dot{V}O_2$ plateau or supramaximal test verification. Additionally,
497 the literature needs to address the issue of the remaining 12% of children who did not
498 have their $\dot{V}O_{2max}$ verified in the supramaximal test, whether that is by investigating the
499 utility of conducting a secondary supramaximal test on the same day or on a separate day,
500 or whether a supplementary incremental and supramaximal test is required.

501 In conclusion, although only 27% had a plateau after the incremental test, the
502 supramaximal test verified $\dot{V}O_{2max}$ in 88% of children and adolescents and was equally

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

510

511 **References**

- 512 1. Armstrong N, Welsman JR. Assessment and interpretation of aerobic fitness in children
513 and adolescents. *Exercise & Sport Sciences Reviews*. 1994;22:435-76.
- 514 2. Armstrong N, Kirby BJ, McManus AM, Welsman JR. Aerobic fitness of prepubertal
515 children. *Annals of Human Biology*. 1995;22(5):427-41.
- 516 3. Armstrong N, Welsman J, Winsley R. Is peak VO₂ a maximal index of children's
517 aerobic fitness? *International Journal of Sports Medicine*. 1996;17(5):356-9.
- 518 4. Astorino TA, DeRevere J. Efficacy of constant load verification testing to confirm VO₂
519 max attainment. *Clinical Physiology & Functional Imaging*. 2017;38(4):703-9.
- 520 5. Barker AR, Williams CA, Jones AM, Armstrong N. Establishing maximal oxygen
521 uptake in young people during a ramp cycle test to exhaustion. *British Journal of Sports*
522 *Medicine*. 2011;45(6):498-503.
- 523 6. Barker AR, Williams CA, Tolfrey K, Fawkner S, Sandercock G. The BASES expert
524 statement on measurement and interpretation of aerobic fitness in young people. *The Sport &*
525 *Exercise Scientist*. 2013;37.
- 526 7. Bhammar DM, Stickford JL, Bernhardt V, Babb TG. Verification of maximal oxygen
527 uptake in obese and nonobese children. *Medicine & Science in Sports & Exercise*.
528 2017;49(4):702-10.
- 529 8. Bongers BC, van Brussel M, Hulzebos EHJ, Takken T. Pediatric norms for
530 cardiopulmonary exercise testing: In relation to sex and age. Second ed. Den Bosch: BoxPress;
531 2014.
- 532 9. Brown JD, Mahon AD, Plank DM. Attainment of maximal exercise criteria in boys and
533 men. *Journal of Sports Medicine & Physical Fitness*. 2002;42(2):135-40.
- 534 10. Caspersen CJ, Powell KE, Christenson GM. Physical-activity, exercise, and physical-
535 fitness - Definitions and distinctions for health-related research. *Public Health Reports*.
536 1985;100(2):126-31.
- 537 11. Cohen J. A power primer. *Psychological Bulletin*. 1992;112(1):155-9.
- 538 12. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for
539 child overweight and obesity worldwide: international survey. *British Medical Journal*.
540 2000;320(7244):1240-3.

541 13. Day JR, Rossiter HB, Coats EM, Skasick A, Whipp BJ. The maximally attainable VO₂
542 during exercise in humans: the peak vs. maximum issue. *Journal of Applied Physiology*.
543 2003;95(5):1901-7.

544 14. Dencker M, Thorsson O, Karlsson MK, Linden C, Eiberg S, Wollmer P, et al. Gender
545 differences and determinants of aerobic fitness in children aged 8-11 years. *European Journal of*
546 *Applied Physiology*. 2007;99(1):19-26.

547 15. Duff DK, De Souza AM, Human DG, Potts JE, Harris KC. A novel treadmill protocol
548 for exercise testing in children: the British Columbia Children's Hospital protocol. *BMJ Open*
549 *Sport & Exercise medicine*. 2017;3(1):e000197-e.

550 16. Hill DW, Poole DC, Smith JC. The relationship between power and the time to achieve
551 VO₂max. *Medicine & Science in Sports & Exercise*. 2002;34(4):709-14.

552 17. Hogstrom G, Nordstrom A, Nordstrom P. High aerobic fitness in late adolescence is
553 associated with a reduced risk of myocardial infarction later in life: a nationwide cohort study in
554 men. *Eur Heart J*. 2014;35(44):3133-40.

555 18. Högström G, Nordström A, Nordström P. Aerobic fitness in late adolescence and the
556 risk of early death: a prospective cohort study of 1.3 million Swedish men. *International Journal of*
557 *Epidemiology*. 2016;45(4):1159-68.

558 19. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. *Continuing*
559 *Education in Anaesthesia Critical Care & Pain*. 2008;8(6):1743-816.

560 20. Martin Bland J, Altman D. Statistical methods for assessing agreement between two
561 methods of clinical measurement. *The Lancet*. 1986;327(8476):307-10.

562 21. Moore SA, McKay HA, Macdonald H, Nettlefold L, Baxter-Jones ADG, Cameron N, et
563 al. Enhancing a somatic maturity prediction model. *Medicine & Science in Sports & Exercise*.
564 2015;47(8):1755-64.

565 22. Murias JM, Pogliaghi S, Paterson DH. Measurement of a true VO₂max during a ramp
566 incremental test is not confirmed by a verification phase. *Frontiers in Physiology*. 2018;9(143).

567 23. Myers J, Walsh D, Buchanan N, Froelicher VF. Can maximal cardiopulmonary capacity
568 be recognised by a plateau in oxygen-uptake. *Chest*. 1989;96(6):1312-6.

569 24. Ortega FB, Ruiz JR, Castillo MJ, Sjostrom M. Physical fitness in childhood and
570 adolescence: A powerful marker of health. *International Journal of Obesity*. 2008;32(1):1-11.

571 25. Pivarnik JM, Dwyer MC, Lauderdale MA. The reliability of aerobic capacity VO₂max
572 testing in adolescent girls. *Res Q Exerc Sport*. 1996;67(3):345-8.

573 26. Poole DC, Wilkerson DP, Jones AM. Validity of criteria for establishing maximal VO₂
574 uptake during ramp exercise tests. *European Journal of Applied Physiology*. 2008;102(4):403-
575 10.

576 27. Poole DC, Jones AM. Measurement of the maximum oxygen uptake VO₂max :
577 VO₂peak is no longer acceptable. *Journal of Applied Physiology*. 2017;122(4):997-1002.

578 28. Raine L, Drollette E, Kao S-C, Westfall D, Chaddock-Heyman L, Kramer AF, et al.
579 The Associations between Adiposity, Cognitive Function, and Achievement in Children.
580 *Medicine & Science in Sports & Exercise*. 2018;50(9):1868-74.

581 29. Rivera-Brown A, Rivera MA, Frontera WR. Reliability of VO₂max in Adolescent
582 Runners: A Comparison Between Plateau Achievers and Nonachievers. *Pediatric Exercise*
583 *Science*. 1995;7:203-10.

584 30. Rivera-Brown AM, Rivera MA, Frontera WR. Applicability of Criteria for VO₂max in
585 Active Adolescents. *Pediatric Exercise Science*. 1992;4(4):331-9.

586 31. Robben KE, Poole DC, Harms CA. Maximal oxygen uptake validation in children with
587 expiratory flow limitation. *Pediatric Exercise Science*. 2013;25(1):84-100.

588 32. Rowland TW, Cunningham LN. Oxygen uptake plateau during maximal treadmill
589 exercise in children. *Chest*. 1992;101(2):485-9.

590 33. Rowland TW. Does peak VO₂ reflect VO₂max in children - Evidence from
591 supramaximal testing. *Medicine & Science in Sports & Exercise*. 1993;25(6):689-93.

- 592 34. Saynor ZL, Barker AR, Oades PJ, Williams CA. A protocol to determine valid
593 VO₂max in young cystic fibrosis patients. *Journal of Science & Medicine in Sport*.
594 2013;16(6):539-44.
- 595 35. Schaun GZ. The maximal oxygen uptake verification phase: A light at the end of the
596 tunnel? *Sports Medicine - Open*. 2017;3(1):44-.
- 597 36. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of
598 cardio-respiratory performance. *Journal of Applied Physiology*. 1955;8(1):73-80.
- 599 37. Welsman J, Bywater K, Farr C, Welford D, Armstrong N. Reliability of peak VO₂ and
600 maximal cardiac output assessed using thoracic bioimpedance in children. *European Journal of*
601 *Applied Physiology*. 2005;94(3):228-34.
- 602 38. Welsman JR, Armstrong N, Nevill AM, Winter EM, Kirby BJ. Scaling peak VO₂ for
603 differences in body size. *Medicine and Science in Sports and Exercise*. 1996;28(2):259-65.
- 604 39. Welsman JR, Armstrong N. Statistical techniques for interpreting body size-related
605 exercise performance during growth. *Pediatric Exercise Science*. 2000;12(2):112-27.
- 606 40. Wood RE, Hills AP, Hunter GR, King NA, Byrne NM. VO₂max in overweight and
607 obese adults: Do they meet the threshold criteria? *Medicine & Science in Sports & Exercise*.
608 2010;42(3):470-7.

609

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

	Males			Females		
	Overall	Non-overweight	Overweight	Overall	Non-overweight	Overweight
	(n =76)	(n =65)	(n =11)	(n =52)	(n=44)	(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·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·kg ⁻¹ ·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·kg ^{-0.66} ·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·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

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

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

	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.

620

621

622

623

624

625