Citation for the original published paper (version of record):

Björkman, F., Eggers, A., Stenman, A., Bohman, T., Ekblom, B. et al. (2018)
Sex and maturity status affected the validity of a submaximal cycle test in adolescents..
Acta Paediatrica, 107(1): 126-133
https://doi.org/10.1111/apa.14080

Sex and maturity status affected the validity of a submaximal cycle test in adolescents

Frida Björkman MSc [a], Andrea Eggers MSc [a], Adam Stenman MD, PhD [a], Tony Bohman PhD [b], Björn Ekblom MD, PhD [a], Örjan Ekblom PhD [a]

Author affiliations: [a] Åstrand laboratory of work physiology, The Swedish School of Sports and Health Sciences, Stockholm, Sweden; [b] Department of Neurobiology, Care Sciences and Society, Division of Physiotherapy, Karolinska Institutet, Stockholm, Sweden.

Corresponding author: MS Frida Björkman, Åstrand laboratory of work physiology, The Swedish School of Sports and Health Sciences, Box 5626, SE-114 86 Stockholm, Sweden, Frida.Bjorkman@gih.se, +46 722 62 70 67.

Short Title: Validity of a submaximal cycle test

Funding: No specific funding.

Conflicts of Interests: No conflicts of interests.

ABSTRACT

Aim: This study assessed the validity and reliability of the Ekblom-Bak (EB) submaximal cycle test in adolescents and identified any sex or maturity-related factors for prediction errors.

Methods: We recruited 50 healthy subjects through a public announcement in Stockholm, Sweden, in 2016. The 27 boys and 23 girls were aged 10-15 years and in Tanner stages I-IV. They performed an EB test and incremental treadmill running test for direct measurement of maximal oxygen uptake (VO_2max).

Results: The estimation error of VO_2max was 0.09 L/min. The correlation (r) was 0.86 and the standard error of the estimate (SEE) was 0.29 L/min. The largest overestimation was seen in pre-pubertal boys (0.49 L/min). The best precision of the EB test was achieved when boys in Tanner stages I and II were re-calculated using the prediction equation developed for adult women. This yielded a mean difference of -0.05 L/min, r = 0.92 and SEE 0.23 L/min, in the entire sample. The prediction error was lowered in boys, but not girls, with increasing pubertal maturity.

Conclusion: The EB test was reasonably valid in adolescents, seemed to be related to sex and maturity status and our findings support its use.

Keywords: ergometry, exercise test, physical fitness, maximal oxygen uptake, sexual maturation

Key notes:

- We assessed the validity and reliability of the Ekblom-Bak (EB) submaximal cycle test in publically-recruited healthy adolescents and identified any sex or maturity-related factors for prediction errors.

- The EB test proved to be a reasonable and valid method for estimating maximal oxygen uptake in the 50 boys and girls who were 10 to 15 years of age.
- Its validity for estimating aerobic capacity depended on pubertal development, especially in boys.

INTRODUCTION

Physical fitness is an important factor for health and wellbeing in boys and girls of all ages. Maximal aerobic capacity, expressed as maximal oxygen uptake (VO_2max), is strongly correlated to health risks, such as cardiovascular diseases (1, 2), metabolic syndrome markers (3), insulin sensitivity (4), mental health and depression (5, 6). It is well known that exercise has positive outcomes for adolescents. Regular cardiovascular training increases both stroke volume and vital capacity in pre-pubertal adolescents (7). From a general health perspective, it is important to be able to assess VO_2max in a simple and precise manner, for example using submaximal exercise tests. The Ekblom-Bak test (EB test) is a submaximal cycle ergometer test that has been shown to be valid and reliable in adults (8, 9). However, it has not yet been tested in adolescents.

There are a number of ways in which children and adolescents differ from adults with regard to their circulatory response to acute exercise. The heart muscles of pre-pubertal children are smaller than adults and they have lower blood volume in relation to total body mass and a lower stroke volume at a given workload. With increased rate of work, the curve linear response in stroke volume displays more or less the same pattern in children and adolescents as it does in adults. It starts with a pronounced increase from rest to light or moderate intensity physical activity, followed by stabilisation with no further increase at higher intensities (10-12). However, it has also been shown that this plateau in stroke volume appears somewhat earlier in adolescents than in adults, at approximately 35-40% of VO_2max

in adolescents and 40-60% of VO_2max in adults (13). Adolescents have a lower cardiac output at a given oxygen uptake (VO_2) and displays a higher heart rate than adults, both for an absolute work rate and a given percentage of VO_2max (10). These circulatory differences between adolescents and adults have to be taken into account when applying adult-specific submaximal tests and prediction models for estimating VO_2max in adolescents. Furthermore, as boys mature, their shift in circulation is more pronounced than the parallel development in girls (14). As a consequence, pre-pubertal and mid-pubertal boys may display more circulatory and morphological similarities with adult women than men. We hypothesised that these divergences in maturity-related changes will result in differences in the test validity between boys and girls.

Therefore, the aim of the present study was to assess criteria-related validity and reliability for the EB test in pre-pubertal and pubertal boys and girls. A secondary aim was to investigate whether any maturity related factors, or other individual factors and factors within the equation, were related to prediction errors.

METHODS

Subjects

The public recruitment drive to identify subjects in Stockholm, Sweden, during spring 2016 included posters in schools and word of mouth. Participants were offered two cinema tickets if they took part in the study and were enrolled as long as they were healthy, with no known diseases. None of the subjects were taking any medication that could have influenced the maximal work capacity and, or, the submaximal heart rate. The inclusion criteria were pre-pubertal and pubertal adolescents aged age 10 to 15 years with a Tanner status of I to IV (15), no serious illness and the physical capacity to perform maximal exercise tests. The complete study set-up consisted of a medical examination within the two weeks before or after the physical fitness tests. The participants also underwent a submaximal cycle test and maximal treadmill running test. All the tests and medical examinations were carried out at the Åstrand laboratory at the Swedish School of Sports and Health Sciences in Stockholm, Sweden. In total, 64 adolescents visited the laboratory at least once. Of those, 20 adolescents – 15 boys and five girls – agreed to pay a second visit within two weeks and they formed the test-retest group. The retest session was conducted in the same way as the first physical tests. We excluded 14 of the 64 volunteers from the analyses: nine participants did not want to be assessed for maturity status and three subjects did not meet the criteria for achieving a valid VO_2max, as described below. Furthermore, one participant was excluded because they failed to keep the pedalling rate, namely the unit revolutions per minute, in the submaximal cycle test and another one was excluded because they were in Tanner stage V. A total of 50 adolescents – 27 boys and 23 girls – completed the full medical examination, including the sexual maturity Tanner scale ratings and were included in the validation group. Two boys in the test-retest group were included even though they lacked maturity ratings, because the analysis of the test-retest reliability did not included any aspects of sexual maturity.

Before each test session, the participants were asked to refrain from vigorous physical activity for 24 hours and to not eat any heavy meals three hours before the tests. All subjects and their parents or guardians were fully informed about the details of the study and provided written, informed consent. The study was approved by the local research ethics committee in Stockholm, Sweden (ref. no. 2016/175-31/2).

Baseline measurements and sexual maturation

Body mass to the nearest 0.1kg and height to the nearest 0.1cm were measured with the participant wearing light clothes. Sexual maturation was visually assessed by a female nurse or a male doctor, using the indices developed by Tanner (15). The subjects were classified for pubic hair rating (boys and girls), genitalia (boys), and breasts (girls). As pubic hair rating was the only value that was assessed for both boys and girls, this value was used for the Tanner classification.

Test preparations

Before each test session at the laboratory, ambient conditions were measured with a portable humidity and temperature instrument, the HygroPalm 0 (Rotronic, Crawley, West Sussex, UK). The computerised metabolic system for measurements of ventilatory gas exchange, Oxycon Pro (Erich Jaeger GmbH, Hoechberg, Germany), was checked with automatic calibration before each test session. The inspiratory flowmeter was checked with the low and high air flow procedure. The gas analyse calibration was performed using a gas mixture of 15.000% oxygen and 5.999% carbon dioxide (Air Liquide, Paris, France). The calibrations for flow volume and gas sensitivity were repeated until variations between two calibrations were less than 1%.

The Ekblom-Bak test

The full description of the EB test procedure has previously been published (8). The EB test was performed on a cycle ergometer 828E (Monark Exercise AB, Vansbro, Sweden), that was manually calibrated before each test. The seat and handlebar were individually adjusted for each participant. Before the tests, all subjects were introduced to the Borg scale for ratings of perceived exertion (16) and practiced the cycling for 1-3 minutes with a breaking force of 0.5 kilopond and a pedalling rate of 60 revolutions per minute to get used to the cycling procedure in the EB test. Thereafter, the EB test was performed with continuous breath-by-breath measurements of VO_2 with the Oxycon Pro and heart rate was recorded in five-second epochs with the RS400 heart rate monitor (Polar Electro Oy, Kempele, Finland).

The EB test consists of continuous cycling for eight minutes with a pedal frequency of 60 revolutions per minute. The workload for the first four minutes was set to 0.5 kilopond for all individuals, corresponding to a work rate of 30 Watts. This was followed by a higher work rate, which was individually determined for each subject, in order to elicit an intensity corresponding to 14-15 on the Borg scale (16). The difference between work rates, or power output, was expressed as the delta output (Δpower output) and used in the prediction equation. The heart rate was recorded and averaged at 3.15, 3.30, 3.45 and 4.00 minutes at the standard and high work rate, respectively. Estimated VO_2max was calculated with the revised, sex-specific prediction equations for the EB test. Details about the prediction equations have previously been published (9).

Maximal running test

Before the maximal test, all subjects had a 5-10 minutes warm up to get them familiarised with running on the treadmill. The warm-up started at 5.0 km/h, with an increase in velocity to 7-9 km/h after 1-2 minutes. After a short jog, the subjects practiced grabbing the handle bars on the sides of the treadmill and exiting the belt securely, so that they felt comfortable with the stopping technique in the maximal tests. At the end of the warm-up session, the velocity was increased to 10.0-13.0 km/h for 20-30 seconds, and, after a short recovery time, the subjects also ran approximately one minute with an elevation of 3-6° and repeated the stopping procedure. The incremental test started with the treadmill at a 1° incline and the same speed as the adolescent felt comfortable with during the warm-up, typically between 7-8.5 km/h. The maximal test protocols were individually designed and the speed and, or, incline increased every minute until volitional exhaustion. All the subjects were given extensive verbal encouragement in order to achieve their VO_2max. Immediately after they finished the test, the subjects were asked to rate their perceived exertion for breathing, legs and whole-body exhaustiveness on the Borg scale (16). The criteria for achieving VO_2max were a maximal heart rate of more than 190, a respiratory exchange ratio of more than 0.95, a Borg scale rating of more than 17 in at least one variable, a running time that exceed five minutes and obvious signs of exhaustive and maximal effort. The test was only accepted if all of the first four variables were fulfilled. VO_2max was referred to as the highest 30 seconds. In adults, the criteria of leveling-off, a plateau of VO_2 versus work rate, is often used. However, this criteria has been questioned in children and adolescents (17).

Statistics

The descriptive statistics of the subject characteristics in Table 1 are presented as median and 25^{th} to 75^{th} percentiles, due to the measurements of VO_2max in L/min being skewed. The non-parametric Mann-Whitney U test was used to detect differences between groups. The difference between the measured and estimated VO_2max was found to be normally distributed and homoscedastic and for this reason it was evaluated as mean differences and limits of agreement and presented in Bland Altman plots. Significant predictors for misclassification in the estimate of VO_2max were identified using two separate multiple linear regressions with the backward method, using a probability of $F \geq 0.10$ for removal). The first analysis included the individual factors of age, sex, stature, heart rate, VO_2 and the respiratory exchange ratio at the standard work rate, maximal heart rate and Tanner stage. The other analyses included the variables in the EB test prediction equation, namely age, Δheart rate/Δpower output, Δpower output and heart rate at the standard work rate. The adjusted prediction coefficient (R^2) and standard errors of the estimate (SEE) were obtained by linear regression (enter mode), by entering measured VO_2max as the dependent variable and estimated VO_2max as the independent variable. Improvements in precision were tested with Levine's test. The Student´s t-test was applied to analyse mean differences between test-retest values. The distributions of the values for oxygen pulse were skewed and the differences between groups were evaluated with the non-parametric Mann-Whitney U test. The sample size was determined based on a minimum correlation of 0.60 and this showed that we needed 20 boys and 20 girls. Due to the number of comparisons when identifying possible predictors for measurement errors, only p values at or below 0.001 were deemed significant. For all other analyses, the statistical significance was set to $p < 0.05$. All statistical analyses were performed with SPSS statistical software version 24.0 (SPSS Inc, Chicago, Illinois, USA).

RESULTS

Subject characteristics for the validation and test-retest groups, respectively, are presented in Table 1. All included children were Caucasian. The group median (25^{th} to 75^{th} percentile) for estimated VO_2max in the validity group was 2.33 (1.95-2.69) L/min. Analysis of the differences between the estimated and measured VO_2max revealed an overall mean and standard deviation (SD) overestimation of 0.09 (0.35) L/min. The correlation coefficient between measured and estimated VO_2max was 0.86 (adjusted R^2 0.73) and the SEE was 0.29 L/min. The limits of agreement were -0.59 to 0.76 L/min.

For the boys, the median estimated VO_2max was 2.57 (2.27-3.20) L/min. The mean (SD) differences between the estimated and measured values showed that boys were overestimated by 0.23 (0.37) L/min. The correlation coefficient between the measured and estimated VO_2max was 0.85 (adjusted R^2 0.71) and the SEE was 0.35 L/min. The limits of agreement between the measured and estimated values were -0.50 to 0.96 for the boys. The corresponding values for the girls were a median estimated VO_2max of 1.97 (1.73-2.32) L/min and a mean underestimation of -0.08 (0.22) L/min. The correlation coefficient for the girls were $r = 0.76$ (adjusted R^2 0.56) and an SEE of 0.19 L/min. The limits of agreement were -0.51 to 0.34. A linear regression model was used to reveal potential predictors for the difference between estimated and measured VO_2. The results showed that the maximal heart rate ($p = 0.004$), sex ($p = 0.04$) and Tanner status ($p = 0.01$) were significant predictors for the difference between measured and estimated VO_2max. Further analysis of the mean differences between estimated and measured values showed that the greatest misclassifications were for boys in Tanner stage I and II, with mean differences of 0.49 (0.27) and 0.19 (0.42) L/min, respectively (Figure 1).

Under the assumption that pre-pubertal and mid-pubertal boys share more circulatory and morphological similarities with adult women than men, the data on the boys in Tanner stage I and II were re-calculated using the equation developed for women. For boys, the new modified results gave a correlation coefficient of 0.92 (adjusted R^2 0.84) and an SEE of 0.25 L/min and limits of agreement of -0.59 to 0.54. This modification of the calculations resulted in an enhanced precision in the entire sample, with a mean difference of -0.05 (0.26) L/min, correlation coefficient of 0.92 (adjusted R^2 0.83) and an SEE of 0.23 L/min. Following the above mentioned adjustment of the test results, the narrower limits of agreement for the entire sample were -0.55 to 0.45 L/min with a p value of > 0.05, according to Levine's test (Figure 2).

Further explanatory variables for the misclassification of the boys in Tanner I and II can be found in the prediction equation itself. In a backward multiple regression analysis, Δheart rate/Δpower output was the only variable in the prediction equation related to the difference between estimated and measured VO_2max, with Δpower output being a proxy for ΔVO_2. The ratio heart rate/power output was related to the oxygen pulse. The oxygen pulses for standard work rate and maximal work, respectively, were relatively similar for boys and girls in Tanner I (medians 6.3 and 10.6 mL/beat versus 5.3 and 8.6 mL/beat, respectively) and Tanner II (medians 7.0 and 11.5 mL/beat versus 6.2 and 9.8 mL/beat, respectively). However, the boys in Tanner III-IV showed significantly higher values (8.1 and 16.2 mL/beat, respectively) compared to Tanner III- IV girls (6.6 and 10.5 mL/beat, respectively).

Reliability of the EB test in adolescents

Analysis of the test-retest reliability for the EB test showed a non-significant mean difference of -0.08 (0.28) L/min, limits of agreement of -0.63 to 0.47 L/min and a correlation coefficient of 0.87 (adjusted R^2 0.74) between the two test occasions. The SEE was 0.28 L/min. The coefficient of variation for the test-retest was low, at 3.1%. In comparison, the coefficient of variation for the measured VO_2max was 2.7% in the 14 participants who performed the repeated maximal tests.

DISCUSSION

The EB test was a valid and reliable test in our study cohort, which indicates that VO_2max can be accurately estimated in adolescence with this submaximal cycle ergometer test. The present study found that the VO_2max of pre-pubertal and pubertal boys and girls aged 10-15 years was estimated within ± 0.68 L/min, compared to ± 0.65 L/min for men and ± 0.53 L/min for women in a mixed adult population in another study (9). The validity in adolescents seemed to depend on maturity and sex and the boys in Tanner stage I and II were the most overestimated. The variable Δheart rate/Δpower output, an analogue to oxygen pulse, was a highly contributing factor to the prediction error. However, the relatively low number of observations in some of the Tanner stages hindered detailed analyses and conclusions.

Previous studies have shown that submaximal cycle ergometer tests for adults can be used for children and adolescents, but this is preferably done after some adjustments or modifications (18). For example, the Åstrand test has been shown to underestimateVO_2max in 10 to 12-

year-old boys by 12% underestimation, with a correlation coefficient of $r = 0.60$ ($p < 0.05$) when expressed in L/min, and $r = 0.55$ ($p < 0.05$) when expressed in mL/kg/min. In that study, the test falsely indicated an enhanced VO_2max after a training period, although no actual change in VO_2max took place (19). Other studies have developed child-specific regression models for estimating VO_2max from the Åstrand test. Woynarowska and et al found a correlation of $r = 0.82$ in girls, but only 0.52 in boys (20). Binyildiz et al used multiple regression analysis to predict VO_2max in 11 to 13-year-old boys, resulting in a significant correlation of $r = 0.70$ between the steady state heart rate at the submaximal work rate and measured VO_2max. When the authors included height and age in the analysis, r was improved to 0.80. and the standard error of the predicted VO_2max calculated from the presented formula was ±18% of the mean (21). In a comparative study, Ekblom evaluated the Åstrand-Ryhming nomogram and the two above mentioned regression models in 62 boys and girls who were aged 11 to 12 years (18). VO_2max was measured during a treadmill running test and estimated via the three methods mentioned above. Low mean misclassification was found for the Woynarowska regression model and Åstrand-Ryhming nomogram with the age correction factor for 12 year olds, at 14 mL/min and 23 mL/min, respectively. The tests had moderately to high correlation ($r = 0.81$ and 0.73, respectively) to measured VO_2max and a high SEE of 398 mL/min and 340 mL/min, respectively. The Binyildiz method demonstrated a large mean underestimation, but a high correlation with measured VO_2max ($r = 0.87$) and relatively low SEE (298 mL/min). Furthermore, all the methods underestimated VO_2max in well-trained subjects (18). In the present study, no such general underestimation was seen for well-trained adolescents. The physical working capacity test is another commonly used cycle test. The different versions of the test have been shown to predict VO_2max with varying accuracy, with correlation coefficients from 0.46 to 0.81 (22-24).

There were a number of reasons for the prediction errors in the adolescent in our study cohort. The largest mean misclassification was seen when pre-pubertal boys were estimated with the prediction equation for men. Previous research has shown that boys and girls display more circulatory and physiological similarities before puberty and that boys and men demonstrate greater maturity-related differences than girls and women. This led to our assumption that we would achieve more accurate estimated values for boys in Tanner stage I and II if we used the prediction equation for women and this was confirmed by the sensitivity analysis. The grouping of boys in Tanner I and II resulted in the most precise values, compared to when only boys in Tanner I were calculated as women and when all subjects were calculated as women (data not shown).

Furthermore, the regression analysis revealed that the variable Δheart rate/Δpower output in the prediction equation was a significant predictor for the misclassification, mostly in boys. The ratio heart rate/power output is related to oxygen pulse. It was not possible to conduct direct measurements of stroke volume and arteriovenous oxygen difference in the present study. Instead, the oxygen pulse served as a surrogate measure (12, 25). The error explained by the variable Δheart rate/Δpower output in the prediction equation, when applied in adolescents, is probably due to the divergent heart rate response to increase workload in pre-pubertal adolescents compared to adults. Because of the fact that the prediction equations are sex-specific – and developed for mature women and men – the differences in Δheart rate/Δpower output response between the sexes is being corrected for in the model. The variable is -0.90742*Δheart rate/Δpower output in the male model and -0.62578*Δheart rate/Δpower output in the female model. This meant that the pre-pubertal boys, who have more morphological and circulatory similarities with women than men, were the most misclassified group when we applied the adult models for the adolescents.

When we analysed the data for the boys and girls separately, the pattern of maturity-related changes was greater in boys than in girls. From the pre-pubertal Tanner I and II stages to the mid and late-pubertal Tanner III and IV stages, boys showed larger circulatory changes than girls. The maturity induced changes in oxygen pulse or respiratory exchange ratio at maximal workload or heart rate at the standard work rate were not as great in the girls as in the boys (data not shown).

The test-retest data revealed that no familiarisation test seemed to be needed, since no significant differences between the repeated EB tests were found (10). This was in line with previous observations in adults (8). When comparing the reliability of the cycle test – with a coefficient of variation of 3.1% in the present study – to measured VO_2max, it is important to keep in mind that there is also a small variability in directly measured VO_2max. In the present study, 14 subjects performed a test-retest for measured VO_2max and the coefficient of variation was 2.7%. Others have reported a variability of 3-6% for maximal tests in adolescents (26, 27). When adults performed a graded exercise test using the same procedures as the subjects in the present study, the coefficient of variation was 2.7% (8). We tested reliability in the form of stability. Another important issue is the ability to detect changes in VO_2max, another aspect of reliability, which was not tested in the present study.

Strengths and limitations

To our knowledge, this was the first study where the validity and reliability for a submaximal cycle test in adolescents was examined together with the influence of biological and sexual maturation. This study focused on maturity-related changes in morphology and circulatory regulation to explain the validity of the EB test. A parallel process during childhood and adolescent growth is the increase in body size. Our only measure of growth was stature, which is a very crude measure. However, stature was not a significant predictor for the measurement error (data not shown), indicating that maturity and not growth seemed to be important for validity. Likewise, age had no significant influence on the misclassification of the estimated VO_2max. This might be explained by the fact that age was already a variable in the prediction equation itself, but it is even more likely that the correlation between chronological age and sexual maturation was very weak. Age cannot be a substitute for the Tanner stage assessment in settings where a proper maturity examination is impossible to perform.

The ability to use the test with the adult female prediction equation for Tanner I and II boys may be relevant in the clinical setting, when a proper maturity rating is simple to perform using the Tanner scale. In the present study, the Tanner assessment was a sufficient substitute for direct and invasive measurements of, for example, hormonal levels. However, in most practical settings, it may be challenging to assess sexual maturity in the same manner as it was carried out in the present study. To categorise adolescents by maturity, without using the Tanner scale, is a drawback that is particularly pronounced for boys. A possible solution is self-assessment via the Tanner scale, but the reliability of that procedure in this context may be an area for future research. Also, the median aerobic capacity of the participants in the present study was higher than in an average paediatric population. Thus, the validity and reliability were not tested in those with poor fitness, which may hinder the ability to generalise the results to untrained or overweight adolescents. Another aspect with regard to the generalisability of data is that the validity was tested in Caucasian adolescents, which may limit the applicability to subjects of other ethnicities.

The sample size was *a priori* based on a minimum correlation of 0.60 between the estimated and measured VO_2max and the correlations were found to be higher. Also, when we compared the deviations in Figure 2a and 2b the number of observations provided sufficient power. However, when we compared the sexes across the Tanner stages, the number of observations were low and the results should be interpreted with caution due to the risk of a type II error.

Another issue that needs to be addressed with regard to the applicability of our findings in clinical versus practical settings is the commonly existing limitations in the standard equipment that is used for cycle tests. Standard cycle ergometers are designed for adults and the Monark 828E ergometer used in this study was only suitable for subjects who were at least 130cm high. Furthermore, the EB test duration of eight minutes is manageable for most adolescents, while other tests with prolonged duration must be modified to take into consideration the shorter attention span and motivational level of adolescents in general (28). Also, the crucial aspect of keeping the pedalling rate throughout the whole test can be a problem for younger participants. Based on the findings in the present study, the EB test is a clinically applicable test, to the same extent as submaximal testing in adults.

Estimating aerobic fitness from submaximal data is an attempt to identify the true VO_2max in an individual, namely the maximal oxygen uptake during maximal whole body work. However, recording oxygen consumption during an incremental treadmill test until voluntary exhaustion is no guarantee that the actual VO_2max has been obtained, especially not in children and adolescents (17). A more suitable nomenclature may be VO_2peak, namely the highest recorded oxygen uptake reached during maximal voluntary effort engaging a large proportion of skeletal muscles. This discrepancy may explain a fraction of the differences between the measured and estimated values. We do, however, believe that this fraction was of limited importance because of the use of a whole body type of exercise, the design of the test protocols and the high maximal values obtained, for heart rate, respiratory exchange ratio and VO_2.

CONCLUSION

The EB test proved to be a reliable and valid test for estimating VO_2max in pre-pubertal and pubertal adolescents. This study supports the use of submaximal testing in adolescents outside the laboratory environment, which may be beneficial for sports, health and clinical settings. It also demonstrated that it was possible to improve the validity for pre-pubertal boys by using the equation developed for adult women.

ABBREVIATIONS

EB test, Ekblom-Bak test; SD, standard deviation; SEE, standard error of estimate; VO_2, oxygen uptake, VO_2max, maximal oxygen uptake.

CONFLICT OF INTERESTS

The authors have no conflicts of interest to declare.

FUNDING

This study did not receive any specific funding.

ACKNOWLEDGEMENTS

Sincere thanks to nurse Catherine Fahlen for her helpful advice and assistance in the Tanner assessments. Also, thanks to Olga Tarassova at the Swedish School for Sports and Health Sciences, for help with the design and formatting of the figures.

REFERENCES

1. Andersen LB, Sardinha LB, Froberg K, Riddoch CJ, Page AS, Anderssen SA. Fitness, fatness and clustering of cardiovascular risk factors in children from Denmark, Estonia and Portugal: the European Youth Heart Study. *Int J Pediatr Obes* 2008; 3 Suppl 1: 58-66

2. Hurtig-Wennlof A, Ruiz JR, Harro M, Sjostrom M. Cardiorespiratory fitness relates more strongly than physical activity to cardiovascular disease risk factors in healthy children and adolescents: the European Youth Heart Study. *Eur J Cardiovasc Prev Rehabil* 2007; 14: 575-81

3. Rizzo NS, Ruiz JR, Hurtig-Wennlof A, Ortega FB, Sjostrom M. Relationship of physical activity, fitness, and fatness with clustered metabolic risk in children and adolescents: the European youth heart study. *J Pediatr* 2007; 150: 388-94

4. Henderson M, Gray-Donald K, Mathieu ME, Barnett TA, Hanley JA, O'Loughlin J, et al. How are physical activity, fitness, and sedentary behavior associated with insulin sensitivity in children? *Diabetes care* 2012; 35: 1272-8

5. Esmaeilzadeh S. The association between depressive symptoms and physical status including physical activity, aerobic and muscular fitness tests in children. *Environ Health Prev Med* 2015; 20: 434-40

6. Brown HE, Pearson N, Braithwaite RE, Brown WJ, Biddle SJ. Physical activity interventions and depression in children and adolescents : a systematic review and meta-analysis. Sports Med 2013; 43: 195-206

7. Ekblom B. Physical training in normal boys in adolescence. *Acta Paediatr Scand Suppl* 1971; 217: 60-2

8. Ekblom-Bak E, Bjorkman F, Hellenius ML, Ekblom B. A new submaximal cycle ergometer test for prediction of VO2max. *Scand J Med Sci Sports* 2014; 24: 319-26

9. Bjorkman F, Ekblom-Bak E, Ekblom O, Ekblom B. Validity of the revised Ekblom Bak cycle ergometer test in adults. *Eur J Appl Physiol* 2016; 116: 1627-38

10. Katsuura T. Influences of age and sex on cardiac output during submaximal exercise. *Ann Physiol Anthropol* 1986; 5: 39-57

11. de Prado DM, Dias RG, Trombetta IC. Cardiovascular, ventilatory, and metabolic parameters during exercise: differences between children and adults. *Arq Bras Cardiol* 2006; 87: e149-55

12. Unnithan V, Rowland TW. Use of Oxygen Pulse in Predicting Doppler-Derived Maximal Stroke Volume in Adolescents. Pediatr Exerc Sci 2015; 27: 412-8

13. Eriksson BO, Grimby G, Saltin B. Cardiac output and arterial blood gases during exercise in pubertal boys. *J Appl Physiol* 1971; 31: 348-52

14. Turley KR, Wilmore JH. Cardiovascular responses to treadmill and cycle ergometer exercise in children and adults. J Appl Physiol (1985) 1997; 83: 948-57

15. Tanner JM. Growth at Adolescence, 2nd Edition. Oxford: *Blackwell scientific publications.* 1962.

16. Borg G. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med* 1970; 2: 92-8

17. Armstrong N, Kirby BJ, McManus AM, Welsman JR. Aerobic fitness of prepubescent children. *Ann Hum Biol* 1995; 22: 427-41

18. Ekblom O. Testing the validity of three submaximal ergometer tests for estimating maximal aerobic capacity in children. *Acta Paediatr* 2014; 103: 559-63

19. Stewart KJ, Gutin B. The prediction of maximal oxygen uptake before and after physical training in children. *J Hum Ergol (Tokyo)* 1975; 4: 153-62

20. Woynarowska B. The validity of indirect estimations of maximal oxygen uptake in children 11-12 years of age. *Eur J Appl Physiol Occup Physiol* 1980; 43: 19-23

21. Binyildiz PO. Prediction of maximal oxygen uptake in boys 11--13 years of age. *Eur J Appl Physiol Occup Physiol* 1980; 43: 213-9

22. McMurray RG, Guion WK, Ainsworth BE, Harrell JS. Predicting aerobic power in children. A comparison of two methods. *J Sports Med Phys Fitness* 1998; 38: 227-33

23. Rowland TW, Rambusch JM, Staab JS, Unnithan VB, Siconolfi SF. Accuracy of physical working capacity (PWC170) in estimating aerobic fitness in children. *J Sports Med Phys Fitness* 1993; 33: 184-8

24. Wells CL, Scrutton EW, Archibald LD, Cooke WP, De la Mothe JW. Physical working capacity and maximal oxygen uptake of teenaged athletes. Med Sci Sports 1973; 5: 232-41

25. Wait J. Cardiopulmonary stress testing. A review of noninvasive approaches. *Chest* 1986; 90: 504-10

26. Paterson DH, McLellan TM, Stella RS, Cunningham DA. Longitudinal study of ventilation threshold and maximal O2 uptake in athletic boys. *J Appl Physiol (1985)* 1987; 62: 2051-7

27. Kenneth R. Turley DMR, Kevin M. Harper, Kathleen I. Kujawa, and Jack H. Wilmore. Maximal Treadmill Versus Cycle Ergometry Testing in Children: Differences, Reliability, and Variability of Responses. *Pediatr Exerc Sci* 1995: 49-60

28. Sady SP. Cardiorespiratory exercise training in children. *Clin Sports Med* 1986; 5: 493-514

Figure legends

Figure 1.

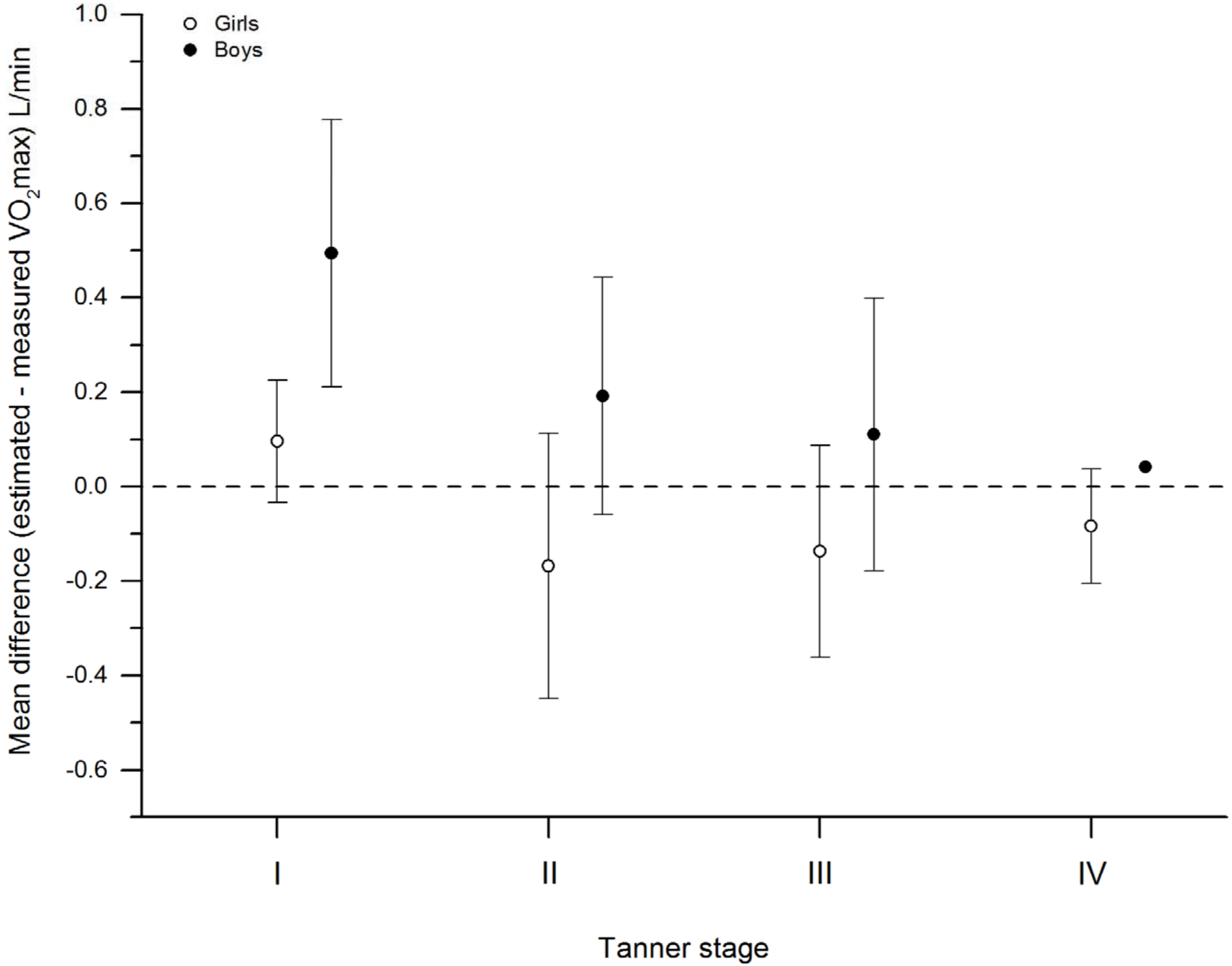
Mean differences between estimated and measured values of VO_2max for the different Tanner stages (I-IV). Error bars denotes 95% confidence interval.

Figure 2.

a) Mean VO_2max (x-axis) and difference between estimated and measured VO_2max (y-axis) for boys and girls, analysed as men and women, respectively.

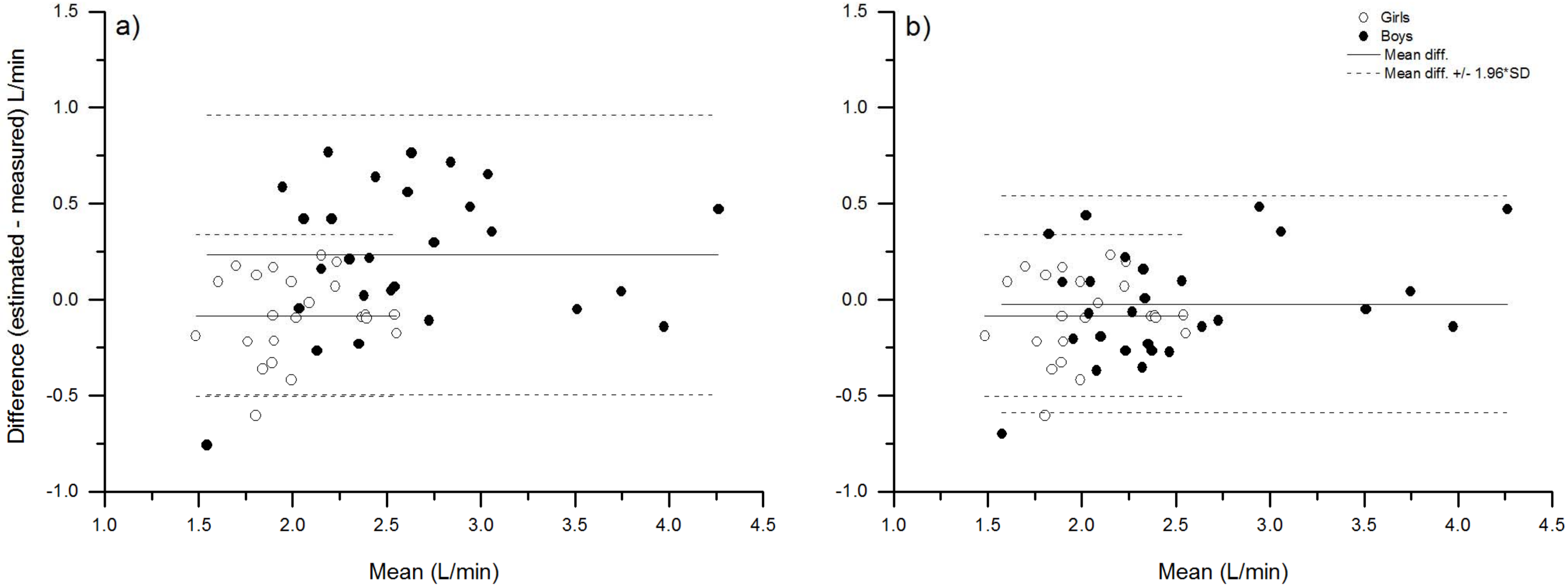
b) Mean VO_2max (x-axis) and difference between estimated and measured VO_2max (y-axis) with boys in Tanner stage I and II, and girls, analysed as women. Boys in Tanner stage III and IV analysed as men.

Figure 1



Girls
Boys
Mean difference (estimated - measured VO2max) L/min
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
I
II
III
IV
Tanner stage

Figure 2



a)
b)
Difference (estimated - measured) L/min
Mean (L/min)
Mean (L/min)
1.5
1.0
0.5
0.0
-0.5
-1.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
Girls
Boys
Mean diff.
Mean diff. +/- 1.96*SD