The Steep Ramp Test in Healthy Children and Adolescents: Reliability and Validity

Bart C. Bongers, M.Sc.[1]

Sanne I. de Vries, M.Sc., Ph.D.[2]

Paul J.M. Helders, P.T., M.Sc., Ph.D.[1]

Tim Takken, M.Sc., Ph.D.[1]

[1]Child Development & Exercise Center, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, the Netherlands; [2]TNO Expertise Center Lifestyle, Leiden, the Netherlands.

Correspondence to:

Dr. Tim Takken, M.Sc., Ph.D.

Child Development & Exercise Center,

Wilhelmina Children's Hospital, University Medical Center Utrecht,

Postal address: KB.02.056.0, P.O. Box 85090, 3508 AB Utrecht, the Netherlands.

Email**: t.takken@umcutrecht.nl**, phone: +31 88 75 540 30, fax: +31 88 75 553 33.

Running title: Steep ramp test's reliability and validity

This study was funded by an unconditional research grant from Scientific Committee Physiotherapy of the Royal Dutch Society for Physiotherapy.

Conflict of interest: none declared for each author.

ABSTRACT

Purpose: To examine the reliability and validity of the steep ramp test (SRT), a feasible, maximal exercise test on a cycle ergometer which does not require the use of respiratory gas analysis, in healthy children and adolescents. Methods: Seventy-five children were randomly divided in a reliability group (n=37, 17 boys; 20 girls, mean±SD age: 13.86±3.22 years) that performed two SRTs within two weeks and a validity group (n=38, 17 boys; 21 girls, age: 13.85±3.20 years) that performed both a SRT and a regular cardiopulmonary exercise test (CPET) with respiratory gas analysis within two weeks. Peak work rate (WR_{peak}) was the main outcome of the SRT. Peak oxygen uptake ($\dot{V}O_{2peak}$) was the main outcome of the CPET. Reliability was examined with the intra class correlation coefficient (ICC) and a Bland and Altman plot, whereas validity was assessed using Pearson's correlation coefficients, and stepwise linear regression analysis. Results: Reliability statistics for the WR_{peak} values attained at the two SRTs showed an ICC of 0.986 ($P<0.001$). The average difference between the two SRTs was -6.4 W, with limits of agreement between +24.5 W and -37.5 W. A high correlation between WR_{peak} attained at the SRT and the $\dot{V}O_{2peak}$ achieved during the CPET was found (r=0.958; $P<0.001$). Stepwise linear regression analysis provided the following prediction equation: $\dot{V}O_{2peak}$ (mL·min^{-1}) = (8.262·WR_{peak} SRT) + 177.096 (R^2=0.917, SEE=237.4). Conclusion: The results suggest that the SRT is a reliable and valid exercise test in healthy children and adolescents, which can be used to predict $\dot{V}O_{2peak}$.

Key words: Exercise Testing; Physical Fitness; Reproducibility; Child

INTRODUCTION

Paragraph number 1 Aerobic capacity is an important determinant of overall health, in which a higher aerobic capacity has been related to a lower morbidity and mortality (5, 25). Direct measurement of aerobic capacity during a symptom-limited maximal cardiopulmonary exercise test (CPET) facilitates an accurate and objective assessment of the integrative response of the metabolic, cardiovascular, and pulmonary system to exercise. The results of a CPET represent the profiles and adequacy of the physiological responses to exercise, which provide clinically diagnostic and prognostic information (32). Measuring maximal oxygen uptake ($\dot{V}O_{2max}$) using respiratory gas analysis during incremental exercise is considered the gold standard for aerobic capacity by the World Health Organization (30) and others (1, 34). The physiological $\dot{V}O_{2max}$ requires the oxygen uptake ($\dot{V}O_2$) to attain a plateau despite a further increase in work rate (WR) (3). This plateau rarely occurs in pediatric populations (**4**, 28). Therefore, the highest $\dot{V}O_2$ measured during a symptom-limited maximal CPET ($\dot{V}O_{2peak}$) is often considered the best measurable indicator of aerobic capacity (9, 33). Nevertheless, direct measurement of $\dot{V}O_{2peak}$ in clinical settings is sometimes not feasible due to the expense, the need for special equipment for respiratory gas analysis, and the trained staff required (11, 12, 26).

Paragraph number 2 As exercise testing is sometimes underused in daily clinical practice (15, 31), there is a need for less demanding alternatives not requiring respiratory gas analysis. This might help to increase the utilization of clinical exercise testing. Maximal exercise testing with peak work rate (WR_{peak}) as primary outcome parameter is a much less demanding procedure (12). WR_{peak} has been indicated as an appropriate alternative measure of $\dot{V}O_{2peak}$ in healthy children (12) as well as in children and adolescents with juvenile idiopathic arthritis (11). The steep ramp test (SRT) is a feasible, short-time maximal exercise test with the achieved WR_{peak} as the main outcome, entitled maximum short-time exercise capacity. The SRT originates from determination and optimization of training WR in adult patients with chronic heart failure (20, 21, 22) and does not require the use of respiratory gas analysis. Hence, the SRT might contribute to an increase of the utilization of exercise testing in clinical settings. Despite its potential clinical applicability, the reliability and validity of the SRT in healthy children and adolescents are currently unknown. **Information concerning its reliability and validity is required for clinicians and researchers willing to use the SRT to evaluate (changes in) exercise capacity.** Therefore, the purpose of the current study was to investigate the reliability and validity of the SRT in healthy children and adolescents. Reliability **was** studied examining the test-retest reliability of the SRT, whereas validity **was** determined investigating the ability of the SRT to predict $\dot{V}O_{2peak}$ attained during a regular CPET.

METHODS

Participants

Paragraph number 3 Healthy children and adolescents were recruited from primary and secondary schools in the Netherlands. Safety and possible risk of maximal exercise for an individual was assessed prior to inclusion using a modified Physical Activity Readiness Questionnaire (PAR-Q), leading to the exclusion of willing participants who answered yes to one or more questions. Three children were excluded due to musculoskeletal disease, one had cardiovascular disease, and two children reported chest pain in the month prior to exercise testing when performing physical activity. Eventually, the study population consisted of 75 healthy participants, who were randomly divided in a reliability (n=37) or a validity group (n=38), in which randomization was stratified by gender and age. Children between 8 and 19 years of age who were free from cardiovascular, pulmonary, neurological, or musculoskeletal disease were eligible. The study protocol was approved by institutional review board of the University Medical Center Utrecht, the Netherlands, and written informed consent was obtained from the legal guardians and/or from the children themselves if they were ≥12 years of age. Characteristics of both groups are presented in Table 1.

Study Design

Paragraph number 4 To assess the reliability of the SRT, the reliability group performed two SRTs within two weeks (mean between-visit time 8.03 ± 5.29 days). The WR_{peak} attained at the first SRT was compared with the WR_{peak} achieved at the second SRT. To assess the validity of the SRT, the validity group performed a SRT at the first visit and a symptom-limited maximal CPET including respiratory gas analysis at the second visit (mean between-visit time 8.26 ± 4.71 days). Both maximal exercise tests were performed at the same time of the day for a given participant. The reached WR_{peak} at the SRT was compared with the $\dot{V}O_{2peak}$ attained at the CPET.

Anthropometry

Paragraph number 5 Anthropometric measurements were conducted prior to exercise testing. Body mass was measured using an electronic scale (Seca 803, Seca, Hamburg, Germany) and body height was measured using a wall mounted stadiometer (Seca 206, Seca, Hamburg, Germany). Biological maturity was assessed by measuring sitting height in order to predict the age from peak height velocity (24). Body mass index (BMI) was calculated as the body mass divided by the square of the body height. Standard deviation (SD) scores were calculated for body height for age, body mass for age, and BMI for age, using Dutch normative values (16). Body surface area (BSA) was calculated using the equation of Haycock et al. (18), which has been validated in infants, children, and adults. Percentage body fat and subsequent fat free mass

(FFM) were determined by measuring subcutaneous fat of the biceps, triceps, subscapular and supra-iliacal regions with a Harpenden skin fold caliper (13). After estimating body density by means of the equations proposed by Deurenberg et al. (13), a modification of the Siri equation was used to estimate percentage body fat (35).

Exercise Testing

Paragraph number 6 Exercise tests were performed on an electronically braked cycle ergometer (Lode Corival, Lode BV, Groningen, the Netherlands). Seat height was adjusted to the participant's leg length. During the tests, heart rate (HR) was monitored by using an elastic belt with a HR sensor (Polar T31™ transmitter, Polar, Kempele, Finland). In order to examine validity, the participants in the validity group breathed through a facemask (Hans Rudolph, Kansas City, MO) during the SRT and the CPET, which was connected to a mobile respiratory gas analysis system (Cortex Metamax B^3, Cortex Medical GmbH, Leipzig, Germany). The metabolic test system was calibrated for respiratory gas analysis measurements (ambient air and a gas mixture of 17% oxygen and 5% carbon dioxide) and volume measurements (3 L syringe) twice a day: in the morning and at noontime. The metabolic test system consisted of the facemask and a transmitting unit with oxygen and carbon dioxide analyzers carried on the participant's chest (total weight 0.57 kg). The mobile respiratory gas analysis system had a wireless connection with a computer, so real-time physical strain of the children during the SRT and the CPET could be measured, as indicated by the minute ventilation (VE), VO_2, carbon dioxide production (VCO_2), and HR averaged at 10-second intervals. This metabolic test system was found to be a **reliable** and valid system for measuring ventilatory parameters during exercise (8, 19, 23). WR_{peak} was defined as the highest achieved WR, whereas peak VE (VE_{peak}), VO_{2peak}, and peak HR (HR_{peak}) were defined as the highest value achieved during the last 30 s before peak exercise. Prior to and directly after the exercise tests, participants completed a 10-point visual analog scale (VAS) indicating their level of fatigue. By doing this, the exhaustiveness of the SRT and the CPET (ΔVAS; post test VAS score minus pre test VAS score) was assessed.

Steep Ramp Test

Paragraph number 7 In order to make the test suitable for pediatric populations, the original SRT protocol (WR increments of 25 W·10 s^{-1} (20)) was modified. After a three-minute warming-up at 25 W, the test started by applying resistance to the ergometer with increments of 10, 15, or 20 W·10 s^{-1}, depending on the participant's body height (<120 cm, between 120 and 150 cm, and >150 cm respectively). The participant was instructed to maintain a pedaling frequency between 60 and 80 revolutions·min^{-1} and the protocol continued until **there was a sustained drop in the** participant's pedaling frequency **from** 60 revolutions·min^{-1} despite strong verbal encouragement. Peak exercise was defined as the point at which the participant's pedaling frequency definitely dropped below 60 revolutions·min^{-1}. Efforts were considered to be maximal when participants showed subjective signs of intense effort (e.g. unsteady biking, sweating, facial flushing, and clear unwillingness to continue despite encouragement).

Cardiopulmonary Exercise Test

Paragraph number 8 During the CPET, participants started with a three-minute warming-up at 25 W where after the WR was increased by 10, 15, or 20 W·min^{-1} depending on the participant's body height (<120 cm, between 120 and 150 cm, and >150 cm respectively) (17). Participants had to maintain a pedaling frequency between 60 and 80 revolutions·min^{-1}. Peak exercise was defined as the point at which **there was a sustained drop in the** participant's pedaling frequency **from** 60 revolutions·min^{-1} despite strong verbal encouragement. A test was considered to be at or near the maximal level if at least one of the following criteria was met: a HR_{peak} >180 beats·min^{-1} or a respiratory exchange ratio (RER) at peak exercise (RER_{peak}) >1.0 (2).

Statistical Analysis

Paragraph number 9 Data analysis was performed using the Statistical Package for the Social Sciences **(SPSS** version 15.0; SPSS Inc., Chicago, IL). All data were expressed as mean ± SD and [range] and were verified for normality with Shapiro-Wilk tests. Since all variables were normally distributed, Paired Samples T-tests were completed in order to determine whether there were significant differences for test duration, exercise variables, and exhaustiveness between the two SRTs performed by the reliability group and between the SRT and the regular CPET executed by the validity group. The two-way mixed intraclass correlation coefficient (ICC) for WR_{peak} and WR_{peak} normalized for body mass were computed to assess reliability of the SRT. ICC values higher than 0.75 were considered acceptable (27). To analyze agreement, limits of agreement were calculated for WR_{peak} according to the procedure described by Bland and Altman (5) using the two WR_{peak} values attained at the two SRTs. To examine the validity of the SRT, the Pearson's correlation coefficient was calculated between the attained WR_{peak} at the SRT and the $\dot{V}O_{2peak}$ achieved during the CPET. Stepwise linear regression analysis was used in order to develop an equation to predict $\dot{V}O_{2peak}$ reached at the regular CPET with the SRT performance (WR_{peak}). First, univariate regression analyses were completed to determine which demographic and anthropometric variables were the best candidate predictors of $\dot{V}O_{2peak}$ achieved at the CPET. Based on their goodness of fit variables were selected to be included into the stepwise linear regression analysis. Statistically significant differences were inferred from P-values <0.05.

RESULTS

Paragraph number 10 The SRTs were well-tolerated by all participants of the reliability group and they all performed the two SRTs at a maximal effort without any complications or adverse effects. They all met signs of the subjective criteria of maximal effort during the two SRTs and the majority of the participants also showed objective signs of maximal effort at the SRT, as indicated by a HR_{peak} >180 beats·min^{-1} (53%). The participants of the validity group met the subjective criteria of maximal effort at the SRT and the CPET as well, and they all attained a HR_{peak} >180 beats·min^{-1} and/or a RER_{peak} >1.0 during the CPET. A plateau in $\dot{V}O_2$ during maximal exercise (29) was observed in 13 children (34%).

Reliability

Paragraph number 11 The results of the two SRTs performed by the reliability group are shown in Table 2. Although the differences in test duration (3.24 s), WR_{peak} (6.41 W), and WR_{peak} normalized for body mass (0.11 W·kg^{-1}) between the two SRTs were small and therefore not clinically relevant, significantly higher values were observed during the second SRT. HR_{peak} and exhaustiveness (ΔVAS) were not significantly different between the two SRTs.

Paragraph number 12 Reliability statistics for the SRT showed an ICC of 0.986 (95% confidence interval (CI): 0.973 to 0.993; P<0.001) for WR_{peak} and an ICC of 0.935 (95% CI: 0.878 to 0.966; P<0.001) for WR_{peak} normalized for body mass. The ICC for the attained HR_{peak} at the SRT was 0.676 (95% CI: 0.451 to 0.821; P<0.001). To analyze agreement between the two SRTs, a Bland and Altman plot is depicted in Figure 1. The average bias ± 1.96 SD between the two SRTs was -6.4 ± 30.9 W. Hence, the limits of agreement for WR_{peak} were +24.5 W and -37.3 W.

Validity

Paragraph number 13 Table 3 presents the results of the SRT and the CPET completed by the validity group. Although significantly higher values were found for the WR_{peak} attained at the SRT compared to the achieved WR_{peak} at the CPET, significantly lower values at the SRT compared to the CPET were observed for test duration, HR_{peak} and $\dot{V}E_{peak}$. All participants of the validity group indicated that they favored the SRT over the CPET when they were asked about their pReference:rential maximal exercise test. This is confirmed by the fact that the CPET received significantly higher values for exhaustiveness (ΔVAS) than the SRT. Figure 2 shows the strong linear relationship between the WR_{peak} attained at the SRT and the $\dot{V}O_{2peak}$ achieved during the CPET. Both variables correlated highly with each other (r=0.958; P<0.001). Based on univariate regression analysis FFM and BSA were also included in the stepwise linear regression analysis. The results however indicated that WR_{peak} attained at the SRT (P<0.001) remained the only significant predictor of $\dot{V}O_{2peak}$, whereas FFM (P=0.377) and BSA (P=0.391) were removed from the model. The following equation was developed to predict $\dot{V}O_{2peak}$ achieved during a CPET from the attained WR_{peak} at the SRT: $\dot{V}O_{2peak}$ ($mL \cdot min^{-1}$) = ($8.262 \cdot WR_{peak}$ SRT) + 177.096 (R^2=0.917, standard error of estimate (SEE)=237.4).

DISCUSSION

Paragraph number 14 The aim of this study was to investigate the reliability and validity of the SRT in healthy children and adolescents. The main results indicate that the SRT comprises good test-retest reliability and is a valid maximal exercise test that can predict $\dot{V}O_{2peak}$ as reached during a regular symptom-limited CPET. In addition, the SRT seems to put a smaller burden on the cardiopulmonary system compared with a regular CPET, as shown by the significantly lower values for HR_{peak} and $\dot{V}E_{peak}$ attained during the SRT. The latter is caused by the short duration of the SRT as a result of the fast increase in WR compared to the regular CPET. Hence, peripheral muscle strength predominates in limiting SRT performance, with consequential higher WR_{peak} values and lower HR_{peak} and $\dot{V}E_{peak}$ values during the SRT.

Paragraph number 15 Especially in pediatric clinical populations it is important that an exercise test can be easily performed by the participant. The SRT is a simple, short-time maximal exercise test, which was well tolerated by all participants. The current study in healthy children and adolescents demonstrates that the SRT seems to be appropriate for pediatric clinical populations because of the fact that it does not require respiratory gas analysis, it has a short duration (approximately between two to three minutes, excluding warming-up and cooling-down), its good reliability, and the valid equation to predict an individuals' $\dot{V}O_{2peak}$.

Paragraph number 16 Regarding its reliability, the average difference between the absolute WR_{peak} values attained at the two SRTs was -6.4 W, indicating that the reliability group on average attained slightly higher WR_{peak} values at the second SRT. Since the differences are scattered symmetrically around the zero bias line up to 400 W, there is no evidence for a significant learning effect. Very high ICCs (>0.9) (27) were found for both WR_{peak} and WR_{peak} normalized for body mass attained at the SRT. This indicates that the SRT is appropriate to utilize for discriminative purposes in cross-sectional samples. For clinicians however, agreement of the measurements is more of interest, as they intend to determine meaningful improvements in a single individual (14). Concerning agreement, or individual variation between the test and re-test, the average absolute WR_{peak} achieved at the two SRTs showed acceptable limits of agreement (24.5 W to -37.3 W), which means that the agreement as indicated by the smallest detectable change (SDC) at the SRT equals 30.9 W. Expressed as a percentage, the limits of agreement were

9% to -13% (SDC: 11%) and appropriate to employ in support of evaluative purposes subsequent to exercise testing of individual subjects.

Paragraph number 17 It is difficult to compare the current study outcomes with existing literature since this is, to our knowledge, the first reliability study of the SRT in pediatric participants. De Backer et al. (10) investigated the test-retest reliability of the WR_{peak} in adult oncology patients who performed an SRT during cancer rehabilitation and reported an ICC of 0.996 (95% IC: 0.989 to 0.998). This is comparable to the ICCs observed in the current study in healthy children and adolescents. Overall, it seems that the SRT performance can be reproducibly performed by healthy children and adolescents.

Paragraph number 18 The WR_{peak} attained at the SRT was highly associated with the $\dot{V}O_{2peak}$ achieved during the CPET, showing its validity as a measure of aerobic capacity. Our results are comparable to those of De Backer et al. (10) in adult oncology patients who also observed a significant correlation between the SRT's WR_{peak} and the CPET's $\dot{V}O_{2peak}$ ($r=0.82$; $P<0.01$). With the attained WR_{peak} at the SRT it was therefore possible to predict a child's aerobic capacity. Several other studies predicted aerobic capacity in pediatric populations during exercise testing, including the regular CPET (11, 12) and a sub-maximal treadmill test (26). $\dot{V}O_{2peak}$ ($mL \cdot min^{-1}$) could be estimated from the WR_{peak} accomplished at a CPET in healthy children ($R^2=0.83$, SEE=114) (12) as well as in children with juvenile idiopathic arthritis ($R^2=0.91$, SEE=180) (11). By means of a sub-maximal treadmill test it was found that $\dot{V}O_{2peak}$ ($mL \cdot min^{-1}$) could be predicted (based on HR and walking speed amongst others) in overweight children ($R^2=0.75$, SEE=271) (26). De Backer et al. (10) developed a prediction equation to predict $\dot{V}O_{2peak}$ ($mL \cdot min^{-1}$) from WR_{peak} attained at the SRT in adult oncology patients, and reported a SEE of 308 ($R^2=0.67$). The current study observed a SEE of 237 when predicting $\dot{V}O_{2peak}$ ($mL \cdot min^{-1}$) which is comparable to those reported above. One can argue that this SEE is larger than those observed by Dencker et al.

(12) and De Backer et al. (11), however, in these studies WR_{peak} and $\dot{V}O_{2peak}$ were obtained during the same test. In the current study, the SRT and the CPET were performed approximately eight days apart, which includes also some day-to-day variance in performance (see reliability section). The same test approach was used by De Backer et al. (10) and comparison of the results revealed that our SEE and R^2 values were more favorable than observed in their study. A Bland-Altman plot for the predicted versus the measured $\dot{V}O_{2peak}$ in the current study showed a mean difference between the predicted and the measured $\dot{V}O_{2peak}$ of 0.3 $mL \cdot min^{-1}$, with all values scattered symmetrically around the zero bias line. The limits of agreement were +459.4 and -458.9 $mL \cdot min^{-1}$. Nevertheless, the conversion to $\dot{V}O_{2peak}$ might be unnecessary, since gender- and age-related reference values for the SRT performance (WR_{peak}) have recently been developed in healthy children and adolescents (7), which facilitates interpretation of SRT results for clinicians and researchers.

Paragraph 19 Compared with a regular CPET, the significantly lower values for HR_{peak} and $\dot{V}E_{peak}$ indicate that the SRT puts a smaller burden on the cardiopulmonary system as has previously been described in heart failure patients (21). In relation with this finding, all participants in the validity group indicated that they preferred performing a SRT over a CPET. Since exercise testing strongly depends on motivational factors, a more positive affective response during exercise will result in better adherence to the exercise protocol. Hence, the results of the exercise test will be more reliable and valid.

Study Limitations

Paragraph number 20 One of the limitations of this study is that only healthy participants were tested. In future studies the reliability and validity of the SRT in clinical populations should be investigated. Although the participants' anthropometry differed not significantly from the general Dutch population norms, the currently developed regression equation for the prediction of aerobic capacity by SRT performance should be cross-validated in a healthy population, as well as in clinical populations. **The lack of habitual physical activity data of the participants as well as the lack of a randomized testing order within the validity group are additional limitations of the current study.**

CONCLUSION

Paragraph number 21 The SRT seems to be a reliable and valid exercise test, which can predict VO_{2peak} in healthy children and adolescents. As the SRT seems to be cardiopulmonary less demanding than a regular CPET, it might be of interest for use in clinical populations as well as in less motivated participants.

ACKNOWLEDGEMENTS

Paragraph number 22 This study was funded by an unconditional research grant from the Scientific Committee Physiotherapy of the Royal Dutch Society for Physiotherapy and the results of the current study do not constitute endorsement by the American College of Sports Medicine (ACSM). We are very grateful to Lode BV, Groningen, the Netherlands and ProCare BV, Groningen, the Netherlands for the technical support during this study. We also would like to thank IVECO Schouten, Utrecht, the Netherlands, for their logistical support. We are very thankful to Anouk Schouten and Mark Mulder for their assistance during the data collection stage. We thank the participating schools, Basisschool Lucas Galecop, Nieuwegein, the Netherlands, Cals College, Nieuwegein, the Netherlands, Graaf Huyn College, Geleen, the Netherlands, Wellantcollege, Gorinchem, the Netherlands, and Zuyd University of Applied Sciences (school of Biometrics and school of Physiotherapy), Heerlen, the Netherlands. Last but not least we are especially grateful to all our participants.

CONFLICT OF INTEREST

Paragraph number 23 The authors declare no conflict of interest.

REFERENCES

1. American Thoracic Society/American College Chest Physicians. ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med*. 2003;167:211-77.
2. Armstrong N, Welsman JR. Aerobic fitness. In: Armstrong N, van Mechelen W, editors. *Paediatric Exercise Science and Medicine*. 2nd ed. Oxford: Oxford University Press; 2008. p. 97-108.
3. Balady GJ, Arena R, Sietsema K, Myers J, Coke L, Fletcher GF, Forman D, Franklin B, Guazzi M, Gulati M, Keteyian SJ, Lavie CJ, Macko R, Mancini D, Milani RV. Clinician's Guide to cardiopulmonary exercise testing in adults: a scientific statement from the American Heart Association. *Circulation*. 2010;122:191-225.
4. **Bar-Or O, Rowland TW. Procedures for exercise testing in children. In: Bar-Or O, Rowland TW. *Pediatric exercise medicine. From physiologic principles to health care application*. Champaign: Human Kinetics; 2004. p. 343-65.**
5. Blair SN, Kohl HW, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA*. 1989;262:2395-401.
6. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-10.
7. Bongers BC, De Vries SI, Obeid J, Van Buuren S, Helders PJM, Takken T. The steep ramp test in children and adolescents: reference values in relation to gender and age (abstract). *Biennial Conference of the North American Society for Pediatric Exercise Medicine;* 2012 Aug 15-18: Philadelphia (Pennsylvania, USA).

8. Brehm MA, Harlaar J, Groepenhof H. Validation of the portable VmaxST system for oxygen-uptake measurement. *Gait Posture*. 2004;20:67-73.

9. Day JR, Rossiter HB, Coats EM, Skasick A, Whipp BJ. The maximally attainable VO2 during exercise in humans: the peak vs. maximum issue. *J Appl Physiol*. 2003;95:1901-7.

10. De Backer IC, Schep G, Hoogeveen A, Vreugdenhil G, Kester AD, van Breda E. Exercise testing and training in a cancer rehabilitation program: the advantage of the steep ramp test. *Arch Phys Med Rehabil*. 2007;88:610-6.

11. De Backer IC, Singh-Grewal D, Helders PJ, Takken T. Can peak work rate predict peak oxygen uptake in children with juvenile idiopathic arthritis? *Arthritis Care Res*. 2010;62:960-4.

12. Dencker M, Thorsson O, Karlsson MK, Lindén C, Wollmer P, Andersen LB. Maximal oxygen uptake versus maximal power output in children. *J Sports Sci*. 2008;26:1397-402.

13. Deurenberg P, van der Kooy K, Hautvast JG. The assessment of the body composition in the elderly by densitometry, anthropometry and bioelectrical impedance. *Basic Life Sci*. 1990;55:391-3.

14. De Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006;59:1033-9.

15. Forman DE, Myers J, Lavie CJ, Guazzi M, Celli B, Arena R. Cardiopulmonary exercise testing: relevant but underused. *Postgrad Med*. 2010;122:68-86.

16. Fredriks AM, van Buuren S, Wit, JM, Verloove-Vanhorick SP. Body index measurements in 1996-7 compared with 1980. *Arch Dis Child*. 2000;82:107-12.

17. Godfrey S. *Exercise testing in children: applications in health and disease*. London: W.B. Saunders Company, Ltd; 1974. p. 1-168.

18. Haycock GB, Schwartz GJ, Wisotsky DH. Geometric method for measuring body surface area: a height-weight formula validated in infants, children, and adults. *J Pediatr.* 1978;93:62-6.

19. Medbo JL, Mamen A, Welde B, et al. Examination of the Metamax I and II oxygen analysers during exercise studies in the laboratory. *Scand J Clin Lab Invest*. 2002;62:585-98.

20. Meyer K, Samek L, Schwaibold M, Westbrook S, Hajric R, Lehmann M, Essfeld D, Roskamm H. Physical responses to different modes of interval exercise in patients with chronic heart failure - application to exercise training. *Eur Heart J*. 1996;17:1040-7.

21. Meyer K, Samek L, Schwaibold M, Westbrook S, Hajric R, Beneke R, Lehmann M, Roskamm H. Interval training in patients with severe chronic heart failure: analysis and recommendations for exercise procedures. *Med Sci Sports Exerc*. 1997;29:306-12.

22. Meyer K. Exercise training in heart failure: recommendations based on current research. *Med Sci Sports Exerc*. 2001;33:525-31.

23. Meyer T, Georg T, Becker C, Kindermann W. Reliability of gas exchange measurements from two different spiroergometry systems. *Int J Sports Med*. 2001;22:593-7.

24. Mirwald RL, Baxter-Jones ADG, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. *Med Sci Sports Exerc*. 2002;34:689–694.

25. Myers J, Prakash M, Froelicher V, Do D, Partington S, Atwood JE. Exercise capacity and mortality among men referred for exercise testing. *N Engl J Med*. 2002;346:793-801.

26. Nemeth BA, Carrel AL, Eickhoff J, Clark RR, Peterson SE, Allen DB. Submaximal treadmill test predicts VO2max in overweight children. *J Pediatr.* 2009;154:677-81.

27. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 3rd ed. Upper Saddle River: Pearson Education, Inc; 2009. p. 585-618.

28. Rowland TW, Cunningham LN. Oxygen-uptake plateau during maximal treadmill exercise in children. *Chest*. 1992;101:485-9.

29. Rowland TW. Does peak VO2 reflect VO2max in children?: evidence from supramaximal testing. *Med Sci Sports Exerc*. 1993;25:689-93.

30. Shephard RJ, Allen C, Benade AJ, Davies CT, Di Prampero PE, Hedman R, Merriman JE, Myhre K, Simmons R. The maximum oxygen intake. An international reference standard of cardiorespiratory fitness. *Bull World Health Organ*. 1968;38:757-64.

31. Stevens D, Oades PJ, Armstrong N, Williams CA. A survey of exercise testing and training in UK cystic fibrosis clinics. *J Cyst Fibros*. 2010;9:302-6.

32. Vanhees L, Lefevre J, Philippaerts R, Martens M, Huygens W, Troosters T, Beunen G. How to assess physical activity? How to assess physical fitness? *Eur J Cardiovasc Prev Rehabil*. 2005;12:102-14.

33. Washington RL, Bricker JT, Alpert BS, Daniels SR, Deckelbaum RJ, Fisher EA, Gidding SS, Isabel-Jones J, Kavey RE, Marx GR. Guidelines for exercise testing in the pediatric age group. *Circulation*. 1994;90:2166-79.

34. Weisman IM, Zeballos RJ. Clinical exercise testing. *Clin Chest Med*. 2001;22:679-701.

35. Weststrate, JA, Deurenberg P. Body composition in children: proposal for a method for calculating body fat percentage from total body density or skinfold-thickness measurements. *Am J Clin Nutr.* 1989;50:1104-15.

FIGURE CAPTIONS

Figure 1: Bland and Altman plot of the WR_{peak} as attained at the first SRT versus the second SRT.

Figure 2: The linear relationship between the $\dot{V}O_{2peak}$ attained at the CPET and the WR_{peak} attained at the SRT.

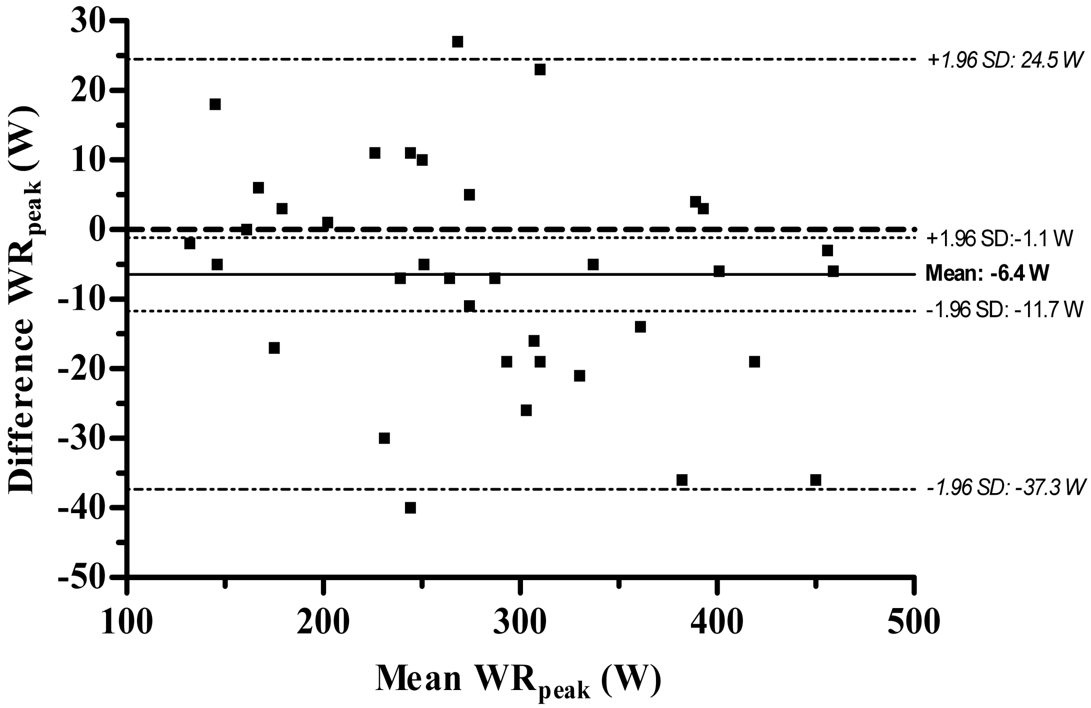
TABLE LEGENDS

Table 1: Group characteristics.

Table 2: Steep ramp test results of the reliability group.

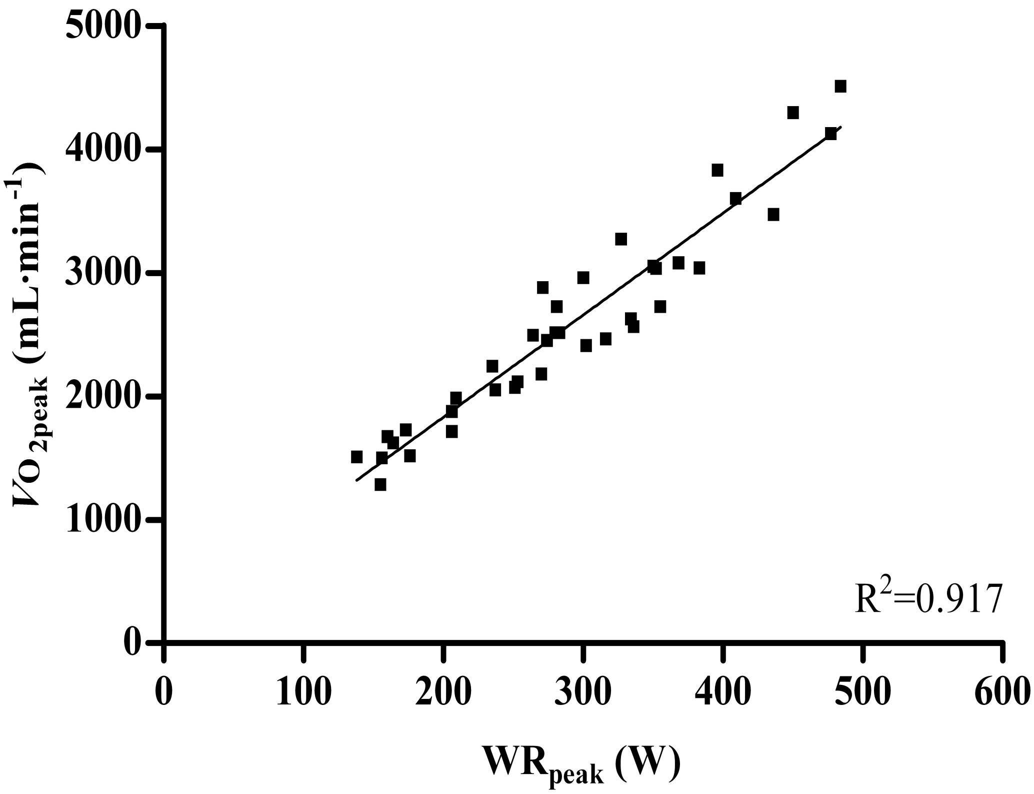
Table 3: Steep ramp test and cardiopulmonary exercise test results of the validity group.

Figure 1:



30
20
10
0
-10
-20
-30
-40
-50
100
200
300
400
500
Difference WRpeak (W)
Mean WRpeak (W)
+1.96 SD: 24.5 W
+1.96 SD:-1.1 W
Mean: -6.4 W
-1.96 SD: -11.7 W
-1.96 SD: -37.3 W

Figure 2:



5000
4000
3000
2000
1000
0
0
100
200
300
400
500
600
VO_{2peak} (mL·min^{-1})
WR_{peak} (W)
R^2=0.917

Table 1:

	Reliability group (n=37)	Validity group (n=38)	*P*-value
Gender boys/girls	17/20	17/21	
Age (years)	13.9 ± 3.2[8.1-18.9]	13.9 ± 3.2[8.1-18.9]	0.99
Body mass (kg)	52.8 ± 15.0[30.0-97.8]	51.1 ± 15.3[23.6-94.2]	0.630
Body height (m)	1.62 ± 0.16[1.29-1.87]	1.61 ± 0.14[1.26-1.85]	0.809
Age from peak height velocity (years)	0.8 ± 2.5[-4.0-4. 0]	0.8 ± 2.4[-4.0-4.0]	0.978
BMI ($kg \cdot m^{-2}$)	19.9 ± 3.2[15.3-28.8]	19.3 ± 3.3[13.2-31.5]	0.463
BSA (m^2)	1.53 ± 0.28[1.07-2.27]	1.50 ± 0.29[0.90-2.16]	0.630
Body fat (%)	21.0 ± 6.1[10.7-35.5]	19.7 ± 4.7[10.3-30.0]	0.288
FFM (kg)	41.5 ± 11.0 [23.7-63.1]	40.8 ± 11.3 [21.2-68.5]	0.790

Values are presented as means ± SD, [range]. Abbreviations: BMI=body mass index; BSA=body surface area; FFM=fat free mass.

Table 2:

	First SRT	Second SRT	*P*-value
Duration (s) [a]	131 ± 42 [63-220]	135 ± 44[70-223]	0.020
WR_{peak} (W)	277 ± 93[131-456]	284 ± 97[133-468]	0.018
WR_{peak}/kg ($W \cdot kg^{-1}$)	5.2 ± 0.8[3.6-6.5]	5.3 ± 0.9[3.7-6.7]	0.038
HR_{peak} ($beats \cdot min^{-1}$)	182 ± 10[163-203]	183 ± 10 [b][166-201]	0.659
ΔVAS	5.5 ± 1.9[0.7-9.3]	6.1 ± 1.8[1.6-9.6]	0.053

Values are presented as means ± SD, [range]. Abbreviations: HR_{peak}=peak heart rate; ΔVAS=visual analog scale difference addressing the participants' level of fatigue (post SRT minus pre SRT); WR_{peak}=peak work rate (maximal short-time exercise capacity). [a]: duration of the load phase, excluding warming-up and cooling-down. [b]: HR_{peak} was not determinable in 1 boy so in this case n=16 for boys.

Table 3:

	SRT	CPET	***P*-value**
Duration (s) [a]	139 ± 41[73-232]	558 ± 183[278-949]	<0.001
WR_{peak} (W)	290 ± 94[138-484]	203 ± 69[94-348]	<0.001
WR_{peak}/kg ($W \cdot kg^{-1}$)	5.7 ± 0.7[4.5-7.9]	4.0 ± 0.6[2.7-5.8]	<0.001
HR_{peak} ($beats \cdot min^{-1}$)	181 ± 10 [b][157-201]	193 ± 9 [b][170-209]	<0.001
$\dot{V}E_{peak}$ ($L \cdot min^{-1}$)	80.7 ± 30.2[27.4-170.3]	93.3 ± 30.7[44.8-166.0]	<0.001
$\dot{V}O_{2peak}$ ($mL \cdot kg^{-1} \cdot min^{-1}$)	*NA NA*	50.7 ± 7.8[36.9-71.2]	*NA*
ΔVAS	5.9 ± 1.7 [2.2-9.1]	7.2 ± 1.8 [2.3-9.9]	<0.001

Values are presented as means ± SD, [range]. Abbreviations: HR_{peak}=peak heart rate; NA=not available; RER_{peak}=peak respiratory exchange ratio; ΔVAS=visual analog scale difference addressing the participants' level of fatigue (post SRT minus pre SRT); $\dot{V}E_{peak}$=peak minute ventilation; $\dot{V}O_{2peak}$=peak oxygen uptake; WR_{peak}=peak work rate (maximal short-time exercise capacity). [a]: duration of the load phase, excluding warming-up and cooling-down. [b]: HR_{peak} was not determinable in 1 girl during both exercise tests so in this case n=20 for girls.