

OPEN

New Reference Values for Cardiopulmonary Exercise Testing in Children

JOEL BLANCHARD^{1,2}, SAMUEL BLAIS¹, PHILIPPE CHETAÏLLE³, MICHELE BISSON³, FRANÇOIS P. COUNIL¹, THELMA HUARD-GIRARD³, JADE BERBARI¹, PIERRE BOULAY², and FREDERIC DALLAIRE¹

¹Department of Pediatrics, Faculty of Medicine and Health Sciences, University of Sherbrooke, and Sherbrooke University Hospital Research Center, Sherbrooke, CANADA; ²Faculty of Physical Activity Sciences, University of Sherbrooke, Sherbrooke, CANADA; and ³Department of Pediatrics, Mother–Child Sun Center, CHU of Quebec City, and Research Center of the Quebec City Hospital Center, Laval University, Quebec City, CANADA

ABSTRACT

BLANCHARD, J., S. BLAIS, P. CHETAÏLLE, M. BISSON, F. P. COUNIL, T. HUARD-GIRARD, J. BERBARI, P. BOULAY, and F. DALLAIRE. New Reference Values for Cardiopulmonary Exercise Testing in Children. *Med. Sci. Sports Exerc.*, Vol. 50, No. 6, pp. 1125–1133, 2018. **Introduction:** Cardiopulmonary exercise testing is an essential tool to assess cardiorespiratory fitness (CRF) in children. There is a paucity of adequate pediatric reference values that are independent of body size and pubertal stage. The purpose of this study is to provide Z score equations for several maximal and submaximal CRF parameters derived from a prospectively recruited sample of healthy children. **Methods:** In this cross-sectional multicenter study, we prospectively recruited 228 healthy children 12 to 17 yr old in local schools. We performed a symptom-limited cardiopulmonary exercise testing progressive ramp protocol on an electronically braked cycle ergometer. Eighteen CRF parameters were analyzed. We tested several regression models to obtain prediction curves that minimized residual association with age, body size, and pubertal stage. Both the predicted mean and the predicted SD were modeled to account for heteroscedasticity. **Results:** We identified nonlinear association of CRF parameters with body size and significant heteroscedasticity. To normalize CRF parameters, the use of a single body size variable was not sufficient. We therefore used multivariable models with various combination of height, corrected body mass, and age. Final prediction models yielded adjusted CRF parameters that were independent of age, sex, body mass, height, body mass index, and Tanner stages. **Conclusions:** We present Z score equations for several CRF parameters derived from a healthy pediatric population. These reference values provide updated predicted means and range of normality that are independent of sex and body size. Further testing is needed to assess if these reference values increase sensitivity and specificity to identify abnormal cardiorespiratory response in children with chronic diseases. **Key Words:** EXERCISE TEST, NORMAL VALUES, ADOLESCENT, CHILD

Address for correspondence: Frederic Dallaire, M.D., Ph.D., Faculty of Medicine and Health Sciences, University of Sherbrooke, 3001, 12e Avenue Nord, Sherbrooke, Canada J1H 5N4; E-mail: frederic.a.dallaire@usherbrooke.ca.

Submitted for publication November 2017.

Accepted for publication January 2018.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/18/5006-1125/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE®

Copyright © 2018 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American College of Sports Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1249/MSS.0000000000001559

Cardiopulmonary exercise testing (CPET) assesses the ability of the respiratory and cardiovascular systems to respond to exercise. It is an important tool to evaluate the effect of chronic cardiac and pulmonary conditions on the integrity of the cardiorespiratory system (1–3). In children, aerobic capacity is strongly influenced by body size and pubertal stage (4). A typical healthy adolescent boy would see his peak $\dot{V}O_2$ go from 1000 to 3000 mL·min⁻¹ during puberty, and such variation is not completely accounted for even if $\dot{V}O_2$ is indexed for body mass (5). Removing the effect of body size on parameters measured during CPET is complex but essential to obtain body size-independent reference values in pediatrics (6).

Interpretation of cardiorespiratory fitness (CRF) requires knowledge of the normal response in healthy individuals so that abnormal response may be distinguished from physiological variations. Body size-independent reference values

are obtained by careful adjustment of measurements for body size. This process helps to remove the physiological effect of normal growth and thus improves the sensitivity and specificity of CRF parameters to detect the pathological effect of disease (7). We recently showed that high-quality pediatric contemporary CPET reference values with adequate adjustment for body size are still needed (5). We believe that a systematic approach to develop predictive equations would likely decrease the signal-to-noise ratio and therefore improve the predictive value of CPET. Moreover, assessments of slopes kinetics during exercise testing have shown great diagnostic and prognostic potential in adults, but reference values for these parameters are either scarce or inexistent for children (8–10).

The purpose of this study was to provide pediatric reference values for 12 submaximal and 6 maximal CRF parameters by using a standardized approach to normalization for body size and pubertal stage on a prospectively recruited sample of healthy children. We aimed to determine predictive formulas that would yield adjusted CRF parameters that are independent of sex, age, body size, puberty, and body mass index (BMI).

METHODS

Population and recruitment. This is a cross-sectional multicenter study. Targeted participants were healthy children between 12 and 17 yr of age. The recruitment took place in schools in the region of Sherbrooke and Quebec City (Canada). A representative of the study visited participating schools to contact potential participants. Parents of interested participants were contacted to provide informed consent. Inclusion criteria were as follows: healthy children and adolescents between 12 and 17 yr of age not involved in high-level competitive sports. Our definition of a competitive athlete was that of the 36th Bethesda Conference (11): “participation in an elite team or individual sport that requires regular competition, systematic and generally intense training, and a focus on excellence and achievement.” If potential participants met this definition, they were not eligible. We did not include competitive athletes because they are expected to have supranormal CRF, especially for maximal parameters, which could skew our predicting equations. Children were excluded for the following reasons: smoking; medical conditions forbidding intense exercise; history of exercise discomfort awaiting medical investigation; BMI-for-age Z score adjusted for sex of less than -2.0 or greater than 2.0 ; musculoskeletal, cardiovascular, or pulmonary condition limiting exercise performance (e.g., muscular dystrophy, congenital heart disease, asthma); and current medication influencing CRF, including inhaled corticosteroids and beta agonists. The CPET was performed either at the Research Center of Sherbrooke University Hospital or at the Research Center of Quebec City University Hospital. The institutional ethics boards approved and monitored this research project in both centers.

Anthropometrics and physical examination. Participants were examined and questioned by a pediatric cardiologist (F.D. or P.C.) and were submitted to a resting ECG and a forced spirometry to rule out contraindications to CPET. We measured body mass, height, waist circumference, and fat-free mass (FFM). Waist circumference was measured using an anthropometric tape at the minimum circumference between the iliac crest and the rib cage and was normalized for age and sex (12). FFM was estimated using a bioelectric impedance scale (TBF 543 (Tanita, Tokyo, Japan) or Inbody520 (Biospace, Urbandale, IA)). BMI was normalized for age and sex (13).

Questionnaire on habitual physical activities and puberty. A questionnaire administered by study personnel was used to determine the level of habitual physical activity. The questionnaire used was an adaptation of the “physical activities” and “sedentary activities” sections of the validated Canadian Community Health Survey (14). Briefly, the type, intensity, and duration of activities performed in the last 3 months were surveyed. The intensity (METs) and time spent doing each activity were used to estimate the mean daily metabolic equivalence ($\text{kcal}\cdot\text{kg}^{-1}\cdot\text{d}^{-1}$) as previously described (14). This mean daily metabolic equivalent allowed for comparison of our sample with the general Canadian population. Pubertal stage was self-reported by asking participants to confidentially identify their level of pubertal development using pictograms of Tanner criteria (15).

CPET. A symptom-limited CPET progressive ramp protocol was performed on two different electronically braked cycle ergometer, depending on the center: the VIAsprint™ 150P bicycle (CareFusion, Totowa, NJ) and the Corival bicycle (Lode, Groningen, the Netherlands). Participants were equipped either with a face mask (7450 Series Silicone V2™ Oro-Nasal; Hans Rudolph, Shawnee, KS) or with a rubber silicone mouthpiece with saliva trap (Hans Rudolph) connected to a Vmax Encore (Vmax Encore Metabolic Cart; SensorMedic, San Diego, CA) or a Ultima™ CardiO2 (MGC Diagnostics, St Paul, MN) metabolic Cart. The workload ramp was individualized to achieve maximum exertion within 8 to 12 min. This time target was chosen on the basis of recent recommendations as well as reports showing that peak $\dot{V}O_2$ is maximized at 10 min in ramp protocols (6,16,17). Selection of the workload ramp was based on predicted values for height (18). It was then tailored according to the participant’s level of habitual physical activity as described by Paridon et al. (10). Participants who did not achieve maximal exercise within 8 to 12 min were either excluded or asked to be retested at a later date. Participants were encouraged to keep a constant pedaling rate of 60 to 80 rpm. The test began with a 2-min rest phase (seated on the cycle ergometer without pedaling) followed by a 3-min warm-up of unloaded cycling. The exercise phase was followed by a 3-min recovery period of slow pedaling (40 rpm) at 10% of the predicted maximal workload. The recovery period was initiated if any of the following events occurred: plateauing of the $\dot{V}O_2$ for >30 s, the participant asked to stop the test (for any reason), or inability to maintain a constant pedaling of greater than 50 rpm.

Tests were interrupted if signs of cardiovascular distress were present (ventricular arrhythmias, falling blood pressure, dizziness, etc.). Participants were actively encouraged to exercise until voluntary exhaustion. Each test was supervised by a pediatric cardiologist and an experienced respiratory therapist (Sherbrooke) or kinesiologist (Quebec City).

We continuously monitored 12-lead ECG, oxygen saturation, and breath-to-breath gas exchanges (flow, O₂ uptake, and CO₂ production) (6). We selected CRF parameters that should be part of a routine CPET as per the latest recommendations (10). We also added CRF submaximal parameters that have shown potential value in assessing cardiovascular response in children with congenital heart diseases (19) (see Table S1, Supplemental Digital Content 1, CRF parameters that were evaluated and how they were calculated, <http://links.lww.com/MSS/B198>). The first ventilatory threshold (VT) was estimated using the ventilatory equivalence method and the V-slope method (6).

To standardize the performance of CPET between centers, we provided testing sites with a detailed testing protocol. The protocol specified all steps of the performance of CPET, including strategies to set workload rate, level of encouragement to give to participants, criteria to stop the test, and so on. In-person meetings took place before the study to review the testing protocol. Study team members performed CPET on themselves in both sites less than 2 wk apart to ensure uniformity in the protocols and measurements. Preliminary analyses were performed after 15 participants at the Quebec City Center to assess potential deviation from measurements already collected at the Sherbrooke site.

Modelization and Z scores. In preliminary analyses, we found an important interaction effect of sex on the association of body size and exercise capacity. Therefore, male and female participants were assessed separately and sex-specific models were developed. Prediction models and Z scores were computed using a systematic and standardized approach previously used by our group (20). The overall objective was to obtain normalized CRF parameters that are independent of body size and age (see the “Modelization and Z score” section in text document, Supplemental Digital Content 1, which presents the complete methodological approach to developing these prediction models, <http://links.lww.com/MSS/B198>).

Briefly, we first evaluated the relationship of the dependent variable (CRF parameter) with three independent variables: height, body mass, and age. We tested a first bivariate model with the independent variable that offered the best fit. For each model, we calculated the residuals of each observation (difference between the observed value and the value predicted by the model). To develop prediction models that are independent of body size, these residuals should no longer be associated with body size. When such associations were present, we tested if the inclusion of a second body size variable in the model diminished the residual association. During this process, we favored simpler models over more complex models to reduce the risk of overadjustment. These prediction models were then used to generate Z scores.

Assessment of the validity of Z scores. To assess if the newly computed Z scores were fully independent of body size, we plotted Z scores against age, body mass, height, BMI-for-age, waist circumference, pubertal stage, and FFM. When significant residual associations were found between body size variables and Z scores, we evaluated if modifications of the prediction models could reduce these associations. If so, the Z scores were recomputed and their validity was reassessed. Other tests were also performed to assess Z score validity (see the “Assessment of the validity of Z score” section in text document, Supplemental Digital Content 1, which explains in detail how the validity of Z scores was assessed, <http://links.lww.com/MSS/B198>).

Strategies to avoid bias related to fat mass. In preliminary analyses, we found a complex relationship between body mass and most CRF parameters, especially in children with increased fat mass. Similarly, we found that most CRF parameters had a strong correlation with FFM, but a weak correlation with fat mass. Because the estimation of FFM with the use of bioelectric impedance is not always available in centers performing CPET, we aimed to develop prediction equations that would not require the estimation of FFM (see the “Strategy to avoid bias caused by body mass” section in text document, Supplemental Digital Content 1, which presents the assessment of various models and correction strategies to account for the effect of body mass for individuals with increased fat mass, <http://links.lww.com/MSS/B198>). Prediction models were deemed adequate if there was no residual association with BMI-for-age after stratification for habitual physical activity.

Statistical analysis. We used SAS for Windows version 9.4 for all analyses (SAS Institute Inc, Cary, NC). Variables that could potentially influence CRF parameters were assessed by comparing the mean Z score between groups. These included subject characteristics and technical aspects such as the use of medication and the center where the CPET was performed. Mean Z scores between groups were compared using Student's *t*-test for two-group comparisons and ANOVA for comparisons of more than two groups. The Wilcoxon rank test was used for nonnormally distributed variables. The chi-square test was used to compare the percentage of physically inactive children in our study population versus the Canadian population. The *t*-statistic was used to estimate *P* values for residual association. Departure from a normal distribution was evaluated by using the Shapiro–Wilk test. A *P* value of <0.05 was considered statistically significant. The Hochberg correction was done for multiple comparisons.

RESULTS

A total of 411 potential participants showed an interest to participate in the study. Of them, 315 could be contacted and assessed for eligibility. Two hundred and forty-seven participants meeting the inclusion criteria were invited to participate. Of them, 4 were excluded for a suspicion of chronic pulmonary disease on spirometry and 11 with a BMI-for-age

Z score >2.0 . Thus, a total of 232 tests were performed. Four participants were subsequently excluded: one for non-sustained ventricular arrhythmia during exercise testing, one because of an obvious submaximal test, and two for not reaching peak exercise within 8–12 min. A total of 228 participants were included in the final analysis. Table 1 shows participants' characteristics. Psychostimulant medication was reported by 8.8% of the participants. There was no significant difference in all CRF Z scores between participants taking psychostimulant medication and those not taking medication (data not shown). Therefore, participants taking psychostimulant medication were not excluded. Similarly, there was no significant difference between the participants tested in Quebec City and those tested in Sherbrooke in all measured CRF parameters Z scores (data not shown).

Preliminary analysis and modelization. For all CRF parameters except the O_2 pulse/workload slope in boys, participants' height was the most strongly associated independent variable. Second degree polynomial or linear models with height were most effective to reduce residual associations with body size. However, residual association with body mass was often still present and the addition of body mass in the models allowed for further reduction of the residual association with body size, but also with pubertal stage. Some CRF parameters, such as $\dot{V}O_2$ /work slope and $\dot{V}_E/\dot{V}CO_2$ slope had little or no association with body size or age. In these cases, simple models such as linear models were sufficient to eliminate all residual associations with body size. For consistency, we also provide equations to calculate Z score even in cases of no association with body size.

Effect of fat mass. In our preliminary analyses, the exclusion of body mass from prediction models did not allow for adequate prediction for several CRF parameters. On the other hand, the inclusion of body mass introduced a bias and underestimated Z scores in overweight individuals for

some parameters (data not shown). Various strategies were used to correct this bias. These include the use of predicted lean body mass by Foster et al. (21), the introduction of BMI-for-age or adjusted waist circumference in prediction models, and the use of a cutoff for body mass in overweight individuals (see the "Strategy to avoid bias caused by body mass" section in text document, Supplemental Digital Content 1, which describes the various strategies tested to account for fat mass, <http://links.lww.com/MSS/B198>). Overall, these methods were equivalent to eliminate most of the residual association with BMI-for-age. We elected to use the latter in our final model owing to its simplicity. We thus used a "corrected body mass" in place of body mass in all prediction models. The corrected body mass is obtained from the participant's body mass from which the mass in excess of the 85th percentile limit of BMI-for-age is subtracted (see the "Strategy to avoid bias caused by body mass" section in text document, Supplemental Digital Content 1, which explains in more details the technique for calculating corrected body mass, <http://links.lww.com/MSS/B198>). This correction is only needed if the tested patient is overweight. For normal-weight individuals, the actual uncorrected body mass can be used.

Final models. Final estimates for prediction models are presented in Table 2 and Table 3 for male and female participants, respectively. These estimates are to be used in the following two equations to calculate the predicted values and the Z scores, respectively:

$$\text{predicted value} = (a \times \text{height}^2) + (b \times \text{height}) + (c \times \text{corrected body mass}) + (d \times \text{age}) + e$$

$$Z = \frac{\text{observed value} - [(a \times \text{height}^2) + (b \times \text{height}) + (c \times \text{corrected body mass}) + (d \times \text{age}) + e]}{(f \times \text{height}) + g}$$

In these equations, height is in centimeters, corrected body mass is in kilograms, and age is in years. To facilitate the interpretation and calculation of the predicted values and Z scores, we have developed an automated Excel sheet that calculates the Z score for each CRF parameter (see Excel file, Supplemental Digital Content 2, which is a Z score calculator for CPET in children, <http://links.lww.com/MSS/B199>). This excel spreadsheet automatically computes the corrected body mass when needed.

Evaluation of the validity of Z score equations. Overall, we were successful at removing most of the effect of body size for all CRF parameters. Z scores had adequate normal distribution and little or no residual association with body size, age, and pubertal stage. For example, Figure 1 demonstrates the absence of residual association for peak $\dot{V}O_2$ Z scores with BMI-for-age, pubertal stage, age, age-adjusted waist circumference, height, and body mass (see also the text document, Tables S2 and S3, and Figures S1 to S6, Supplemental Digital Content 3, which presents detailed results for residual associations and departure from the normal distribution for all CRF parameters, <http://links.lww.com/MSS/B200>).

TABLE 1. Participants' characteristics.

Characteristics	Values
Participants, no. (%)	
Sherbrooke	184 (80.7)
Quebec City	44 (19.3)
Male gender, no. (%)	112 (49.1)
Age, mean \pm SD, yr	15.1 \pm 1.6
Height, mean \pm SD, cm	165.2 \pm 8.5
Body mass, mean \pm SD, kg	56.0 \pm 10.8
FFM, mean \pm SD, kg	45.7 \pm 8.7
BMI, mean \pm SD, $\text{kg}\cdot\text{m}^{-2}$	20.4 \pm 3.0
BMI Z score, mean \pm SD	0.1 \pm 0.9
Waist circumference, mean \pm SD, cm	73.6 \pm 8.9
Waist circumference Z score, mean \pm SD	0.8 \pm 1.3
Self-reported ethnicity, no. (%)	
Caucasian	200 (87.7)
Others	28 (12.0)
Self-reported pubertal stage, no. (%)	
Tanner stage II	9 (3.9)
Tanner stage III	38 (16.7)
Tanner stage IV	116 (50.9)
Tanner stage V	65 (28.5)
Level of habitual physical activity, mean \pm SD, $\text{kcal}\cdot\text{kg}^{-1}\cdot\text{d}^{-1}$	2.8 \pm 2.1
Medication, no. (%)	
Oral contraceptive	27 (11.8)
Psychostimulants	20 (8.8)
Other	12 (5.3)

TABLE 2. Models estimates for exercise parameters for male participants.

CRF Parameters	n	Predicted Mean					Predicted SD	
		a	b	c	d	e	f	g
Maximal parameters								
Peak $\dot{V}O_2$	109	-0.297	105.9	36.6	0	-8660	6.45	-717.1
Peak O_{2p}	107	-0.00131	0.459	0.214	0	-37.48	0.0277	-2.67
Peak workload	109	0.0182	-5.324	2.824	4.170	378.9	0.220	-7.62
Peak \dot{V}_e	109	0.00228	-0.419	0.981	3.168	2.704	0.405	-52.54
Peak HR	107	-0.000532	0.313	-0.259	0	169.5	0.0966	-7.47
Peak RER	109	0	0.00142	-0.000976	0.0155	0.786	-0.000161	0.0935
Submaximal parameters								
OUES	109	-0.171	57.8	39.1	0	-4247	8.61	-1043
OUES slope below VT	107	0.0923	-30.4	32.7	0	3181	7.27	-783.6
$\dot{V}_e/\dot{V}CO_2$ slope	107	0	-0.0407	0	0	35.1	-0.00559	4.48
$\dot{V}_e/\dot{V}CO_2$ slope below VT	107	-0.000918	0.319	-0.0466	-0.599	7.87	-0.0527	11.04
$\dot{V}_e/\dot{V}CO_2$ at VT	109	0.00128	-0.434	-0.0924	0	68.39	-0.0289	7.28
$\dot{V}O_2$ at VT	107	-0.146	56.3	18.0	-48.3	-3898	3.11	-90.9
$\dot{V}O_2$ /work slope	109	0	-0.00871	0	0	12.4	0.0121	-0.995
Workload at VT	107	0.00386	-0.939	1.27	0	104.4	0.251	-9.28
O_{2p} /work slope	109	0	0	0	0	0.0383	0	0.00748
O_{2p} increase (%)	107	0.0139	-4.03	-0.534	0	451.6	-0.599	137.7
HRR after 1 min of recovery	89	0.0168	-4.91	-0.439	0	536.4	0.000430	12.68
HRR after 2 min of recovery	86	0	0.627	-0.556	-0.679	80.17	0.0666	2.65

HR, heart rate; HRR, heart rate recovery; O_{2p} , O_2 pulse; OUES, oxygen uptake efficiency slope; RER, respiratory exchange ratio.

Because it has been suggested that FFM should be used to adjust for body size in overweight individual (22–25), we compared our models with models using FFM. Regardless of body mass (normal or overweight), there was no significant difference in the mean Z score for peak $\dot{V}O_2$ when corrections were done using corrected body mass or FFM (see Figure S7, Supplemental Digital Content 3, which compares mean peak $\dot{V}O_2$ Z scores calculated from models using corrected body mass or FFM, stratified for normal-weight or overweight participants, <http://links.lww.com/MSS/B200>). Similar results were found for other exercise parameters (not shown). Blant–Altman analysis of our models and models with FFM showed that they were highly correlated with small mean differences and little bias (see Figure S8, Supplemental Digital Content 3, which illustrates the relationship and agreement between the Z score of the peak $\dot{V}O_2$, modeled with the FFM and modeled with height and corrected weight,

<http://links.lww.com/MSS/B200>). After adjustment for the level of habitual physical activity, mean peak $\dot{V}O_2$ Z scores produced by our model in overweight subjects were similar to those in normal-weight subjects (see Figure S9, Supplemental Digital Content 3, which illustrates peak $\dot{V}O_2$ Z scores stratified by level of habitual physical activity for normal weight and overweight participants, <http://links.lww.com/MSS/B200>).

Effect of habitual physical activity. We observed a statistically significant association between the level of habitual physical activity and maximal CRF parameters, but not with submaximal CRF parameters. Figure 2 shows box plots of Z scores for maximal (peak O_2 pulse, peak $\dot{V}O_2$, and peak workload) and submaximal ($\dot{V}O_2$ /workload, $\dot{V}_e/\dot{V}CO_2$ slope, and $\dot{V}_e/\dot{V}CO_2$ slope below VT) CRF parameters, according to the level of habitual physical activity.

To evaluate the possibility of a selection bias (more active individuals being more interested in participating in our

TABLE 3. Models estimates for exercise parameters for female participants.

CRF Parameters	n	Predicted Mean					Predicted SD	
		a	b	c	d	e	f	g
Maximal parameters								
Peak $\dot{V}O_2$	119	-0.244	86.80	14.70	0	-6424	2.12	-45.9
Peak O_{2p}	116	-0.00019	0.075	0.1007	0	-1.83	-0.00320	2.17
Peak workload	119	-0.06025	20.57	0.741	0	-1622	0.284	-24.41
Peak \dot{V}_e	119	-0.00697	2.56	0.528	1.14	-202.86	0.0681	3.72
Peak HR	116	-0.0213	7.198	-0.193	-0.809	-391.1	-0.121	28.41
Peak RER	119	0	0.00122	-0.00195	0.0143	0.906	-0.00109	0.251
Submaximal parameters								
OUES	119	-0.251	91.4	13.8	0	-6768	4.48	-406.1
OUES slope below VT	118	-0.0333	12.8	15.9	0	35.6	5.13	-476.0
$\dot{V}_e/\dot{V}CO_2$ slope	118	0.000191	-0.112	0.0697	0	37.9	-0.0103	5.37
$\dot{V}_e/\dot{V}CO_2$ slope below VT	118	-0.00558	1.83	0.0191	-2.901	-120.7	0.00481	1.98
$\dot{V}_e/\dot{V}CO_2$ at VT	118	-0.00548	1.81	-0.0232	0	-119.1	0.0181	-0.00423
$\dot{V}O_2$ at VT	118	-0.00407	-2.14	15.9	-26.7	1282	0.454	215.3
$\dot{V}O_2$ /work slope	119	0.00145	-0.500	0.0152	0	52.17	-0.00247	1.58
Workload at VT	118	-0.0122	4.09	0.601	0	-276.7	0.0369	17.77
O_{2p} /work slope	119	0	-0.000232	0.0000108	0	0.0515	-0.00000512	0.0154
O_{2p} increase (%)	116	-0.114	38.2	-0.290	-3.54	-3002	0.509	-47.4
HRR after 1 min of recovery	100	0	0.225	-0.0470	-0.669	144.8	-0.289	60.02
HRR after 2 min of recovery	98	0	0.397	-0.239	-0.624	110.8	-0.118	32.48

HR, heart rate; HRR, heart rate recovery; O_{2p} , O_2 pulse; OUES, oxygen uptake efficiency slope; RER, respiratory exchange ratio.

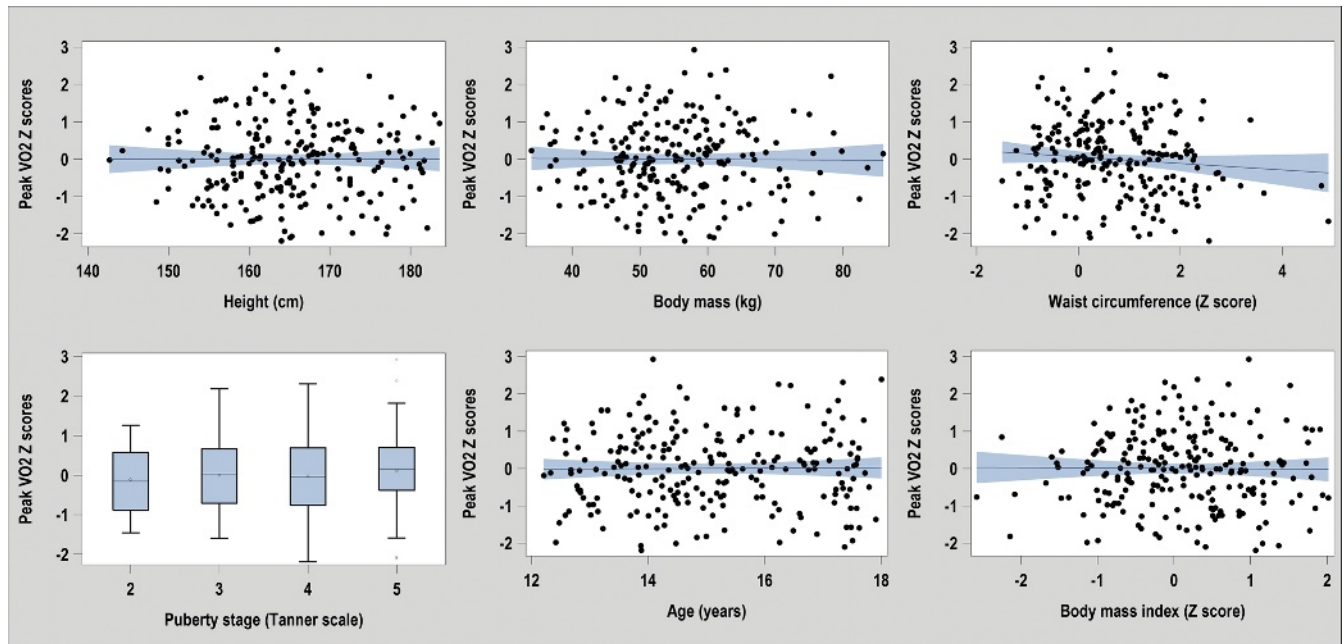


FIGURE 1—Absence of residual associations with height, body mass, pubertal stage, age, waist circumference, and BMI with the peak $\dot{V}O_2$ Z score.

study), we compared the level of habitual physical activity of our sample with those published by Statistics Canada in 2014 (similar questionnaire administered to 3,112,914 Canadian children) (14). The percentages of physically inactive girls ($<1.5 \text{ kcal}\cdot\text{kg}^{-1}\cdot\text{d}^{-1}$) were 32.8% and 34.9% for our study sample and for that of Statistics Canada, respectively ($P = 0.62$). The percentage of physically inactive boys were 26.4% and 24.5% for our sample and for that of Statistics Canada, respectively ($P = 0.61$).

Comparison with other reference values. Figure 3 shows unadjusted and adjusted peak $\dot{V}O_2$ according to body mass. When peak $\dot{V}O_2$ was expressed as an index of body mass (panel A) or as a percentage of predicted values (panel B), we observed an important residual association with body mass (nonzero slope). This indicated an overestimation of predicted peak $\dot{V}O_2$ in heavier participants. This residual association was removed by the current predicting equations (panel C). Similar results were observed for other CRF parameters (e.g., peak O_2 pulse) and other body size variables (e.g., height; data not shown). Residual associations could lead to misclassification. For example, in our sample of healthy adolescents, $>12\%$ of participants taller than 170 cm fell below the limit of the 80% of predicted according to Cooper's equation (26). This potential overestimation of predicted values in taller individuals was corrected by our predicted values ($<2\%$ were classified below a Z score of -2 , which is in line with second percentile limit).

DISCUSSION

In this study, we present Z score equations for several submaximal and maximal CRF parameters, which are commonly used for the interpretation of a CPET. We identified nonlinear

associations with body size and significant heteroscedasticity. We used a standardized approach to optimize regression modeling and produce Z scores with adequate data distribution and little residual association with body size. In most cases, the use of a single body size variable was not sufficient to normalize CRF parameters. We thus used multivariable models with various combinations of height, corrected body mass, and age, and we were able to produce adequate Z score equations that were independent of body size and pubertal stage.

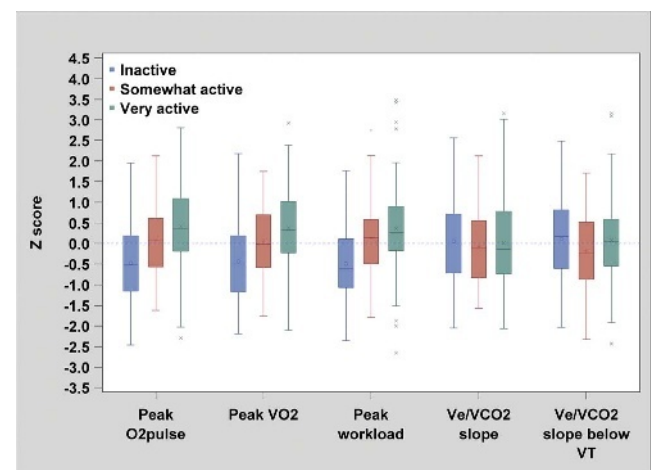


FIGURE 2—Mean Z score of maximal and submaximal CRF parameters stratified according to level of habitual physical activity. We found significant differences ($P < 0.0001$) between the Z score of maximal CRF parameters (peak O_2 pulse, peak $\dot{V}O_2$, and peak workload) and the level of habitual physical activity. We found no significant differences between the Z score of submaximal CRF parameters ($\dot{V}O_2$ /workload, $\dot{V}_e/\dot{V}CO_2$ slope, and $\dot{V}_e/\dot{V}CO_2$ slope below VT) and the level of habitual physical activity ($P = 0.13$, $P = 0.89$, and $P = 0.16$, respectively).

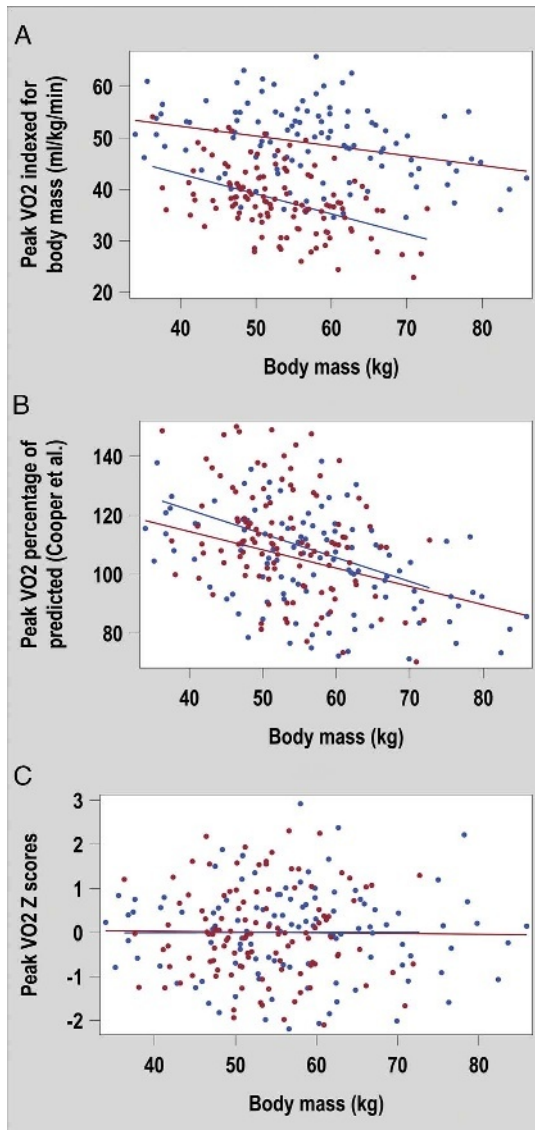


FIGURE 3—Peak $\dot{V}O_2$ for boys (*blue*) and girls (*red*) according to body mass. The peak $\dot{V}O_2$ is expressed as an uncorrected value indexed for body mass in panel A, as a percentage of predicted values according to Cooper et al. (4) in panel B, and as a Z score according to the equations proposed in this study in panel C.

In a recent systematic review, we showed that the number of CPET pediatric reference values was limited, and when available, the methods used for normalization were heterogeneous with inconsistent—and often incomplete—adjustment for body size (5). We pointed out that the assessment of potential biases (residual association, residual heteroscedasticity, and departure from the normal distribution) had been inconsistent and insufficiently described. Furthermore, there were few studies reporting reference values for some submaximal and maximal CRF parameters that may have interesting prognostic significance, such as O_2 pulse, $\dot{V}O_2$ /work slopes, and the oxygen uptake efficiency slope. This paucity of appropriate recent CPET reference values in children has for effect that some laboratories and metabolic cart software still rely on reference values proposed >30 yr ago (4,18,27).

The relation between body size and CRF is complex and it is unlikely that a single body size measurement, like the body mass, can account for variation in CRF parameters (8,28,29). It has been advocated that lean body mass or total muscle mass was a better predictor of CRF than height, weight, or body surface area alone. We have previously shown that using multivariable models including both weight and height was superior to predict the diameter of the aorta compared with models including only one body size measurement (30). Others have also advocated for the use of multivariable models to better predict CRF (31). In our previous review, we found that some authors did assess and compare multiple independent variables (32–35), but none proposed the use of a multivariable model incorporating more than one independent variable. A precise evaluation of body composition would likely be preferable to accurately predict CRF. However, because such an evaluation is not practical in routine clinical settings, we believe that the multivariable models presented in this study strike the right balance between accurately predicting CRF while only using body size measurements that are easily acquired.

Historically, investigators and clinicians have used 80% of the predicted values as the limit of normality. This lower limit seemed to have been borrowed from the literature in forced spirometry (36,37), but its validity in categorizing normal CRF is lacking. We observed significant heteroscedasticity for some measured CRF parameters, meaning that the variance around the mean was not constant as children grew older and taller. We found few previous studies testing for the presence of heteroscedasticity in which the SD was modeled to account for its variation with body size (5). If the nonconstant variance is unaccounted for, one could expect a Z score that is not representative of the true distribution of the adjusted parameter. In our sample, the lower limit of normal (Z score -2.0) corresponded to a percentage of predicted values that varied from 66% to 77%. Using our predicted mean, as much as 11% of our sample fell below the arbitrary level of 80% of predicted value. Hence, we believe that the routine practice of considering 80% of the predicted value as being an abnormally low CRF result is biased and should be abandoned. One should rather use Z score equations with adequate modeling of the SD and consider results above the 98th percentile or below the 2nd percentile (roughly $Z > 2.0$ and $Z < -2.0$) as being abnormal.

Body mass includes fat mass and FFM, which have markedly different influence on CRF (38). Inclusion of body mass in prediction equations may thus introduce bias in overweight individuals, and it has been suggested that FFM has a better correlation with CRF (31). In this study, we excluded obese children with a BMI-for-age Z score of >2 , but overweight children were included. When we included body mass in prediction models, we observed a residual negative association with BMI-for-age and waist circumference for some parameters, indicating a possible underestimation of predicted CRF with increasing BMI-for-age. This was also observed after controlling for habitual physical

activity. We were successful in reducing these residual associations by correcting body mass for overweight subjects. This approach was simple and intuitive yet yielded results similar to other approaches, such as using predicted lean body mass, using measured FFM by bioelectrical impedance, or introducing other parameters in the model.

Maximal CRF parameters were associated with the level of habitual physical activity. We expected such results because physical activity is known to increase CRF (39). For submaximal parameters, the level of habitual physical activity had very little effect with similar results in active and nonactive children. Further studies are needed to determine if abnormal submaximal CRF parameters will be helpful in differentiating between healthy but sedentary children and children with impaired cardiovascular or respiratory response caused by chronic diseases.

This study has limitations. Our sample size is relatively small. This increases the possibility of bias due to sampling and selection. In our recent review, all prospective studies on cycle ergometer included <300 pediatric healthy participants. For this study, we chose a ramp protocol on a cycle ergometer. We know that CRF parameters may be different according to the exercise protocol (10). For example, peak $\dot{V}O_2$ has been shown to be 9% to 11% higher on treadmill compared with cycle ergometer (40). Consequently, the reference values proposed here may not apply to other exercise protocols. This was a multicenter study in which two metabolic carts were used by two separate teams, with a possibility of variation in the level of encouragement to participants. The test protocols were, however, strictly standardized and equipment was calibrated before each test. When CRF data were compared according to the center, we found no significant differences (not shown). Our population was mostly Caucasian. Considering

the possible differences in cardiovascular responses to exercise according to ethnic background, our results should be validated in non-Caucasian populations.

CONCLUSIONS

In this study, we present *Z* score equations for several CRF parameters derived from a healthy pediatric population. We believe that our careful adjustment for body size yielded *Z* scores that are independent of body size, pubertal, age and sex. Using weighted models to account for heteroscedasticity has also allowed us to define more precisely the nonconstant range of normality with body size. These new and updated reference values should provide a more accurate lower normal limit. This may improve their value for the diagnosis, prognosis, and risk stratification of children with chronic diseases. We are currently recruiting children with tetralogy of Fallot and with heart failure to confirm this hypothesis.

The authors would like to show their appreciation to the participants and the teachers involved in this study. We are grateful to Patrice Cotoir for facilitating this study in his school. We thank the teachers and students of De Rochebelle, De la Montée, and Mitchell-Montcalm secondary schools for their enthusiast participation. We are indebted to Mario Geoffroy for his expertise in exercise testing and respiratory physiology. We wish to acknowledge the Fondation des Étoiles for supporting this study. F. D. and the Centre de Recherche du Centre Hospitalier Universitaire de Sherbrooke are supported in part by the Fonds de Recherche du Québec—Santé. J. B. is supported by scholarships from the Faculty of Physical Activity Sciences and from the Centre d'Excellence de l'Université de Sherbrooke en recherche Mère-Enfant.

The authors have no conflict of interest to declare. The results of the study are presented clearly, honestly, without fabrication, falsification, or inappropriate data manipulation, and do not constitute endorsement by the American College of Sports Medicine.

REFERENCES

- Ohuchi H, Negishi J, Noritake K, et al. Prognostic value of exercise variables in 335 patients after the Fontan operation: a 23-year single-center experience of cardiopulmonary exercise testing. *Congenit Heart Dis*. 2015;10(2):105–16.
- Myers J. Applications of cardiopulmonary exercise testing in the management of cardiovascular and pulmonary disease. *Int J Sports Med*. 2005;26(1 Suppl):S49–55.
- Guazzi M, Bandera F, Ozemek C, Systrom D, Arena R. Cardiopulmonary exercise testing: what is its value? *J Am Coll Cardiol*. 2017;70(13):1618–36.
- Cooper D, Weiler-Ravell D, Whipp BJ, Wasserman K. Aerobic parameters of exercise as a function of body size during growth in children. *J Appl Physiol Respir Environ Exerc Physiol*. 1984;56(3):628–34.
- Blais S, Berbari J, Counil FP, Dallaire F. A systematic review of reference values in pediatric cardiopulmonary exercise testing. *Pediatr Cardiol*. 2015;36(8):1553–64.
- Wasserman K, Hansen JE, Sue DY, Stringer WW, Whipp BJ. Principles of exercise testing and interpretation: including pathophysiology and clinical applications. *Med Sci Sports Exerc*. 2005;37(7):52–131.
- Colan SD. The why and how of *Z* scores. *J Am Soc Echocardiogr*. 2013;26(1):38–40.
- Hansen JE, Sue DY, Wasserman K. Predicted values for clinical exercise testing 1–3. *Am Rev Respir Dis*. 1984;129(2P2):S49–55.
- Neder JA, Nery LE, Castelo A, et al. Prediction of metabolic and cardiopulmonary responses to maximum cycle ergometry: a randomised study. *Eur Respir J*. 1999;14(6):1304–13.
- Paridon SM, Alpert BS, Boas SR, et al. Clinical stress testing in the pediatric age group: a statement from the American Heart Association Council on Cardiovascular Disease in the Young, Committee on Atherosclerosis, Hypertension, and Obesity in Youth. *Circulation*. 2006;113(15):1905–20.
- Maron BJ, Zipes DP. Introduction: eligibility recommendations for competitive athletes with cardiovascular abnormalities—general considerations. *J Am Coll Cardiol*. 2005;45(8):1318–21.
- Katzmarzyk PT. Waist circumference percentiles for Canadian youth 11–18 y of age. *Eur J Clin Nutr*. 2004;58(7):1011–5.
- de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martinez J. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. *Food Nutr Bull*. 2004;25(1 Suppl 1):S15–26.
- Health Canada. *Canadian Community Health Survey Cycle 2.2, Nutrition (2004): Income-Related Household Food Security in Canada*. Ottawa (Ontario): Health Canada; 2007.
- Tanner JM. *Growth at Adolescence: With a General Consideration of the Effects of Hereditary and Environmental Factors Upon Growth and Maturation From Birth to Maturity*. 2nd ed. Oxford (England): Scientific B; 1962.

16. Takken T, Bongers BC, van Brussel M, Haapala E, Hulzebos EH. Cardiopulmonary exercise testing in pediatrics. *Ann Am Thorac Soc*. 2017;14(1 Suppl):S123–6.
17. Buchfuhrer MJ, Hansen JE, Robinson TE, Sue DY, Wasserman K, Whipp BJ. Optimizing the exercise protocol for cardiopulmonary assessment. *J Appl Physiol Respir Environ Exerc Physiol*. 1983;55(5):1558–64.
18. Godfrey S, Davies C, Wozniak E, Barnes CA. Cardio-respiratory response to exercise in normal children. *Clin Sci*. 1971;40(5):419–31.
19. Dallaire F, Wald RM, Marelli A. The role of cardiopulmonary exercise testing for decision making in patients with repaired tetralogy of Fallot. *Pediatr Cardiol*. 2017;1–9.
20. Dallaire F, Slorach C, Hui W, et al. Reference values for pulse wave Doppler and tissue doppler imaging in pediatric echocardiography. *Circ Cardiovasc Imaging*. 2015;8(2):e002167.
21. Foster BJ, Platt RW, Zemel BS. Development and validation of a predictive equation for lean body mass in children and adolescents. *Ann Hum Biol*. 2012;39(3):171–82.
22. Janz KF, Dawson JD, Mahoney LT. Tracking physical fitness and physical activity from childhood to adolescence: the Muscatine study. *Med Sci Sports Exerc*. 2000;32(7):1250–7.
23. Janz KF, Burns TL, Witt JD, Mahoney LT. Longitudinal analysis of scaling VO_2 for differences in body size during puberty: the Muscatine study. *Med Sci Sports Exerc*. 1998;30(9):1436–44.
24. Ortega FB, Ruiz JR, Mesa JL, Gutiérrez Á, Sjörström M. Cardiovascular fitness in adolescents: the influence of sexual maturation status—the AVENA and EYHS studies. *Am J Hum Biol*. 2007;19(6):801–8.
25. Maffei C, Schena F, Zaffanello M, Zocante L, Schutz Y, Pinelli L. Maximal aerobic power during running and cycling in obese and non-obese children. *Acta Paediatr*. 1994;83(1):113–6.
26. Cooper DM, Weiler-Ravell D. Gas exchange response to exercise in children 1, 2. *Am Rev Respir Dis*. 1984;129(2P2):S47–8.
27. Cooper DM, Weiler-Ravell D, Whipp BJ, Wasserman K. Growth-related changes in oxygen uptake and heart rate during progressive exercise in children. *Pediatr Res*. 1984;18(9):845–51.
28. Jones N, Summers E, Killian K. Influence of age and stature on exercise capacity during incremental cycle ergometry in men and women. *Am Rev Respir Dis*. 1989;140(5):1373–80.
29. Welsman JR, Armstrong N, Nevill AM, Winter EM, Kirby BJ. Scaling peak VO_2 for differences in body size. *Med Sci Sports Exerc*. 1996;28(2):259–65.
30. Dallaire F, Bigras JL, Prsa M, Dahdah N. Bias related to body mass index in pediatric echocardiographic Z scores. *Pediatr Cardiol*. 2015;36(3):667–76.
31. Loftin M, Sothorn M, Abe T, Bonis M. Expression of $\text{VO}_{2\text{peak}}$ in children and youth, with special reference to allometric scaling. *Sports Med*. 2016;46(10):1451–60.
32. Cooper DM, Kaplan MR, Baumgarten L, Weiler-Ravell D, Whipp BJ, Wasserman K. Coupling of ventilation and CO_2 production during exercise in children. *Pediatr Res*. 1987;21(6):568–72.
33. Gulmans V, De Meer K, Binkhorst R, Helders P, Saris W. Reference values for maximum work capacity in relation to body composition in healthy Dutch children. *Eur Respir J*. 1997;10(1):94–7.
34. Gursel Y, Sonel B, Gok H, Yalcin P. The peak oxygen uptake of healthy Turkish children with reference to age and sex: a pilot study. *Turk J Pediatr*. 2004;46(1):38–43.
35. Mandadzhieva S, Marinov B, Kostianev S, Turnovska T. Anthropometric and cardiopulmonary parameters in Bulgarian and Romany children: cross-sectional study. *Croat Med J*. 2005;46(2):294–301.
36. Miller MR, Quanjer PH, Swanney MP, Ruppel G, Enright PL. Interpreting lung function data using 80% predicted and fixed thresholds misclassifies more than 20% of patients. *Chest*. 2011;139(1):52–9.
37. Hansen JE. Lower limit of normal is better than 70% or 80%. *Chest*. 2011;139(1):6–8.
38. Goran M, Fields D, Hunter G, Herd S, Weinsier R. Total body fat does not influence maximal aerobic capacity. *Int J Obes (Lond)*. 2000;24(7):841.
39. Payne VG, Morrow JR Jr. Exercise and VO_2 max in children: a meta-analysis. *Res Q Exerc Sport*. 1993;64(3):305–13.
40. Turley KR, Rogers DM, Harper KM, Kujawa KI, Wilmore JH. Maximal treadmill versus cycle ergometry testing in children: differences, reliability, and variability of responses. *Pediatr Exerc Sci*. 1995;7(1):49–60.