Validity of Equations for Estimating $V_{O_2}$peak From the 20-m Shuttle Run Test in Adolescents Aged 11–13 Years

Mariana B. Batista, Edilson S. Cyrino, Miguel Arruda, Antonio C. Dourado, Manuel J. Coelho-E-Silva, David Ohara, Marcelo Romanzini, and Enio R.V. Ronque

1Group of Study and Research in Physical Activity and Exercise, Londrina State University, Londrina, Brazil; and 2Metabolism, Nutrition, and Exercise Laboratory, Londrina State University, Londrina, Brazil; 3Faculty of Physical Education, State University of Campinas, Campinas, Brazil; 4Department of Sport Science, Physical Education and Sport Center, Londrina State University, Londrina, Brazil; and 5Faculty of Sport Science and Physical Education, University of Coimbra, Coimbra, Portugal

Abstract

Batista, MB, Cyrino, ES, Arruda, M, Dourado, AC, Coelho-E-Silva, MJ, Ohara, D, Romanzini, M, and Ronque, ERV. Validity of equations for estimating $V_{O_2}$peak from the 20-m shuttle run test in adolescents aged 11–13 years. J Strength Cond Res 27(10): 2774–2781, 2013—The aim of this study was to analyze the validity of 4 regression equations to estimate the peak oxygen consumption ($V_{O_2}$peak) from the 20-m shuttle run test in adolescents aged 11–13 years. One hundred and fifteen adolescents, 61 boys (mean ± SD: age = 12.3 ± 0.9 years) and 54 girls (age = 12.1 ± 0.7 years) performed the 20-m shuttle run test and an incremental progressive maximal test for direct $V_{O_2}$peak analysis. Four linear regression equations were used to estimate the $V_{O_2}$peak: Barnett et al. (equation 1), Léger et al. (equation 2), Mahar et al. (equation 3), and Matsuzaka et al. (equation 4). For boys, only the $V_{O_2}$peak estimated by EQ3 did not differ from the value directly measured ($p > 0.05$). The EQ1, EQ2, and EQ4 underestimated the $V_{O_2}$peak, whereas the EQ3 overestimated, particularly in girls ($p < 0.05$). Large limits of agreement were found between the reference method and the 4 equations, with higher estimated values by EQ2 for boys ($8.36 ± 15.24$ mL·kg$^{-1}$·min$^{-1}$) and girls ($2.45 ± 12.63$ mL·kg$^{-1}$·min$^{-1}$). The highest correlation values were observed by EQ4 for boys ($r = 0.80$), EQ1 for girls ($r = 0.72$), and EQ3 for total sample ($r = 0.80$). The equations analyzed were not precise for individual $V_{O_2}$peak prediction; however, the EQ3 revealed better agreement, particularly for boys. Considering the data obtained in the boys and total sample, our results suggest that the EQ3 may provide the best predictive measure of $V_{O_2}$peak from the 20-m shuttle run test in adolescents aged 11–13 years.

Key Words cardiorespiratory fitness, field test, Bland-Altman, youth

Introduction

The assessment of the cardiorespiratory fitness (CRF) has been considered as an important factor in aspects related to public health because high levels of CRF can reduce the risk of morbidity and mortality from all causes and from cardiovascular diseases in the adult population (6,7,16). In children and adolescents, high levels of CRF tend to decrease the risk factors for cardiovascular diseases, such as obesity, high blood pressure, dyslipidemia, insulin resistance, among others (2,13,25).

In relation to the assessment of the CRF, peak oxygen consumption ($V_{O_2}$peak) has been used as one of the best indexes for measuring aerobic power in youth (3,4). Thus, different field tests have been developed to estimate $V_{O_2}$peak in children and adolescents from the walking, running, cycling, or swimming protocols, with predetermined time or distance, continuous or intermittent, and others with fixed distances and progressive increase in speed at every stage (9,10).

Among the field tests for the $V_{O_2}$peak prediction in youth, the 20-m shuttle run test (17) has been the most widely used and recommended to assess the health-related physical fitness in youth (11,12,31). The 20-m shuttle run test has been often used because it presents some advantages such as low cost, wide applicability, and can be administered in a relatively small space either indoors or outdoors for several subjects simultaneously (22,30). Moreover, the daily activities of children and adolescents have a strong intermittent feature; therefore, the 20-m shuttle run test may be a more specific test than continuous directional run protocols (18,22).

Besides the increased use of the 20-m shuttle run test as an indicator of $V_{O_2}$peak in youth, there has been great interest...
from the scientific community in the development and validation of equations to predict \( V_{\text{O2peak}} \), including their regression model variables, such as sex, age, speed achieved in the last stage, and number of laps in the test, body mass, skinfold thickness, body mass index (BMI), among others. The results of these studies have indicated correlation values between \( V_{\text{O2peak}} \) obtained by the reference method and the estimate given by equation ranging from \( r = 0.65 \) to \( r = 0.87 \) and determination coefficient from \( R^2 = 0.68 \) to \( R^2 = 0.85 \) (5,17,19–21,29).

However, it is unclear which are the most appropriate equations for the \( V_{\text{O2peak}} \) prediction in children and adolescents. Melo et al. (22) investigated 7 equations for estimating \( V_{\text{O2peak}} \) in children aged 8–10 years and concluded that the equations proposed by Fernhall et al. (14) and Matsuzaka et al. (20) showed better agreement with reference values but with low performance for estimating \( V_{\text{O2peak}} \) individually. Additionally, Ruiz et al. (30) investigated 5 different equations for \( V_{\text{O2peak}} \) prediction from the 20-m shuttle run test in adolescents aged 13–19 years of both sexes and found that the equations proposed by Barnett et al. (5) and Ruiz et al. (29) seem to provide the best \( V_{\text{O2peak}} \) predictions.

Thus, the aim of this study was to analyze the validity of 4 regression equations available in the literature to estimate the \( V_{\text{O2peak}} \) from the 20-m shuttle run test in adolescents aged 11–13 years. A better comprehension of validity of these equations in adolescents can contribute to epidemiological studies and help health and fitness professionals in monitoring \( V_{\text{O2peak}} \) adaptations during interventions periods in both inside and outside the school environment. Our hypothesis is that at least one of the regression equations chosen for this study may provide a valid measure of \( V_{\text{O2peak}} \) in adolescents.

**METHODS**

**Experimental Approach to the Problem**

At baseline, after a previous contact with the school board and knowing that the institution had approximately 250 students enrolled in the fifth and sixth grades of elementary school, being within the age group from 11 to 13 years, all adolescents were invited to participate in the study. On the first visit to the school, the project proposal was presented and also assessments and procedures that would be involved. On the second visit to the school, 115 participants were submitted the anthropometric measurements in a room provided by the school. After this period, the 20-m shuttle run test was conducted in the laboratory for the determination of \( V_{\text{O2peak}} \) as reference value (dependent variable). All visits to perform the motor tests were performed at intervals of 48 hours, and the subjects were submitted to 2 protocols according to the order described. The subjects were instructed to take a light meal 2 hours before the experiments, avoid physical exercise for a minimum of 24 hours before the each testing session, and maintain similar sleeping hours and daily activities.

Before the collection of definitive data, a group of adolescents with the same characteristics of the study sample was randomly selected to make the pilot study. This group was submitted to 2 successive sets of measures to check the evaluators’ reproducibility, to check data collection operationalization, and to verify the equipment calibration. The group consisted of 63 adolescents, 31 boys (age = 12.2 years, body weight = 40.6 kg, height = 148.0 cm) and 32 girls (age = 11.7 years, body weight = 45.0 kg, height = 151.0 cm), who were submitted to repeated measurements of anthropometric variables and 20-m shuttle run test.

**Subjects**

The study included 115 adolescents, 61 boys (12.3 ± 0.9 years) and 54 girls (12.1 ± 0.7 years) enrolled in an educational institution in Londrina, Paraná, Brazil. Data collection was undertaken during the academic year 2009. A comprehensive verbal description of the nature and purpose of the study, and also the clinical implications of the investigation, was provided to the participants, their parents, and teachers. This information was also sent to parents or participants’ tutors. The following inclusion criteria were adopted to be within the group and be regularly enrolled in the public school selected and as exclusion criteria to have physical problem that temporarily or permanently prevents the individual from being submitted to the motor tests and not signing the informed consent term. This study was approved by the local ethical committees, and all participants gave verbal informed consent. Written informed consent was obtained from the adolescent’s parent or legal guardian. All procedures were in accordance with those outlined by the Declaration of Helsinki.

**Anthropometry**

Body mass was measured to the nearest 0.1 kg using a calibrated electronic scale (model PS 180A; Urano, Porto

<table>
<thead>
<tr>
<th>Variables</th>
<th>Boys (n = 61)</th>
<th>Girls (n = 54)</th>
<th>Total (n = 115)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>12.3 ± 0.9</td>
<td>12.1 ± 0.7</td>
<td>12.2 ± 0.8</td>
</tr>
<tr>
<td>Body mass (kg)</td>
<td>47.5 ± 13.6</td>
<td>45.1 ± 13.2</td>
<td>46.3 ± 13.4</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>150.9 ± 7.7</td>
<td>150.4 ± 7.3</td>
<td>150.7 ± 7.5</td>
</tr>
<tr>
<td>BMI (kg m(^{-2}))</td>
<td>20.7 ± 4.7</td>
<td>19.7 ± 4.3</td>
<td>20.2 ± 4.5</td>
</tr>
</tbody>
</table>

*Values are expressed as mean ± SD. No significant intergroups difference (p > 0.05).
Alegre, Brazil), with the subjects wearing light workout clothing and no shoes. Height was measured using a wooden stadiometer to the nearest 0.1 cm while subjects were standing without shoes. Body mass index (BMI) was calculated as the body mass in kilograms divided by the square of the height in meters. The anthropometric measurements were obtained during the afternoon by the same evaluator using standardized procedures described by Gordon et al. (15). The technical errors of measurement (TEMs) were less than 1% (body mass = 0.36 kg and stature = 0.57 cm), and intraclass correlation coefficients (ICC) were 0.99 for both measurements.

**Field Test and Estimation of \( V_{\text{peak}} \)**

The 20-m shuttle run test was conducted as described by Léger et al. (17) on an indoor court with slip resistant floor in a 20-m space bounded by 2 parallel lines. Participants were instructed to run in a straight line, to pivot, and turn on completing a shuttle and to pace themselves in accordance with the audio signals. The initial speed was set at 8.5 km·h\(^{-1}\) (2.4 m·s\(^{-1}\)), which was increased by 0.5 km·h\(^{-1}\) (0.1 m·s\(^{-1}\)) each minute (1 minute equals 1 stage). The test audio CD was calibrated for a duration of 1 minute. Subjects were advised on proper pacing strategy and motivated to give their best effort. During the test, the subjects were encouraged and verbally communicated at each change of stage. The test ended when the participant stopped because of fatigue or when they failed to reach the end lines concurrent with the audio signals on 2 consecutive occasions. The participants were familiar with the test in their physical education classes. Furthermore, the subjects were instructed to refrain from strenuous exercise in 48 hours before testing. All tests were conducted by the same investigators. The absolute and relative TEMs were 0.23 km·h\(^{-1}\) and 2.3%, respectively; the ICC was 0.93 and coefficient of variation (CV) was 2.9%.

\( V_{\text{peak}} \) was estimated in milliliters per kilometer per minute by 4 linear equations:

a) Barnett et al. (5): \( V_{\text{peak}} = 25.8 - 6.6 \times (G) - 0.2 \times (BM) + 3.2 \times (S) \), where \( G = \text{gender} (0 = \text{male and } 1 = \text{female}), \ BM = \text{body mass}, S = \text{speed in kilometers per hour at the end of the test}. \)

b) Léger et al. (17): \( V_{\text{peak}} = 31.025 + 3.238 \times (S) - 3.248 \times (A) + 0.1536 \times (A \times S) \), where \( S = \text{speed in kilometers per hour at the end of the test}, A = \text{age in years}. \)

c) Mahar et al. (19): \( V_{\text{peak}} = 47.438 + (\text{no.LAPS} \times 0.142) + (G \times 5.134) - (BM \times 0.197) \), where no.LAPS = numbers of laps in the SR-20m test, \( G = \text{gender} (0 = \text{male and } 1 = \text{female}), BM = \text{body mass.} \)

d) Matuzakaka et al. (20): \( V_{\text{peak}} = 25.9 - 2.21 \times (G) - 0.449 \times (A) - 0.831 \times (BM) + 4.12 \times (S) \), where \( G = \text{gender} (0 = \text{male and } 1 = \text{female}), A = \text{age in years}, BM = \text{body mass index}, S = \text{speed in kilometers per hour at the end of the test}. \)

**Direct Measurement of \( V_{\text{peak}} \)**

Direct analysis of \( V_{\text{peak}} \) was conducted in laboratory through open circuit spirometry, with the performance of a progressive and maximum test in treadmill ergometer. The test began with warm-up exercises for 3 minutes at a rate of 6 km·h\(^{-1}\) and 0% slope, and then the speed was increased by 1 km·h\(^{-1}\) every minute, maintaining the same slope up to completion of the test. The protocol adopted was tested in the previous pilot study in which adolescents with the same characteristics and age reached maximum effort in a time interval between 8 and 12 minutes, which has been recommended to obtain aerobic power indicators in youth (28).

To measure \( V_{\text{O}2} \) during the test, a portable gas analyzer model K4 b\(^2\) (Cosmed, Rome, Italy) was used. Before each test was conducted, the oxygen and carbon dioxide analyzers were calibrated according to the manufacturer’s instructions. During each test, a gel seal was used to help prevent air leaks from the face mask. Respiratory parameters were recorded breath-by-breath, which in turn were averaged over a 15-second period. The criteria adopted for the completion of the test were as follows: (a) subject’s voluntary exhaustion, with the request to finish the test; (b) reaching the maximum heart rate predicted for age (220 – age); (c) respiratory exchange ratio exceeding 1.1; (d) detection of a plateau in the \( V_{\text{O}2} \) curve, defined by an increase of less than 2 ml·kg\(^{-1}\)·min\(^{-1}\) in the \( V_{\text{O}2} \) with change of stage in the test. Therefore, when the subject showed ≥1 of these characteristics, the test was finished and then the highest \( V_{\text{O}2} \) value obtained was considered as the \( V_{\text{O}2\,\text{peak}} \), represented in relative form (ml·kg\(^{-1}\)·min\(^{-1}\)).

**Table 2.** \( V_{\text{O}2\,\text{peak}} \) values (ml·kg\(^{-1}\)·min\(^{-1}\)) measured from the reference method and estimated by 20-m shuttle run test from 4 linear equations (n = 115).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Boys (n = 61)</th>
<th>Girls (n = 54)</th>
<th>Total (n = 115)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference method</td>
<td>49.9 ± 9.5</td>
<td>42.2 ± 7.4</td>
<td>46.3 ± 9.4</td>
</tr>
<tr>
<td>Barnett et al. (5)</td>
<td>47.8 ± 4.5††§</td>
<td>40.3 ± 3.8††‡</td>
<td>44.3 ± 5.6††§</td>
</tr>
<tr>
<td>Léger et al. (17)</td>
<td>41.6 ± 4.2††§</td>
<td>39.7 ± 3.1††‡</td>
<td>40.7 ± 3.9††§</td>
</tr>
<tr>
<td>Mahar et al. (19)</td>
<td>50.2 ± 5.1††§</td>
<td>43.8 ± 4.9††‡</td>
<td>47.2 ± 5.6††§</td>
</tr>
<tr>
<td>Matuzakaka et al. (20)</td>
<td>43.7 ± 6.3††‡</td>
<td>40.7 ± 5.0††‡</td>
<td>42.3 ± 5.9††‡</td>
</tr>
</tbody>
</table>

\( p < 0.05 \) vs. reference method.  
\( \dagger p < 0.05 \) vs. Léger et al. (17).  
\( \ddagger p < 0.05 \) vs. Mahar et al. (19).  
\( § p < 0.05 \) vs. Matuzakaka et al. (20).  
\( || p < 0.05 \) vs. Barnett et al. (5).
**TABLE 3.** Performance criteria: intercept, slope, correlation coefficient ($r$), standard error of estimate (SEE), and coefficient of variation (CV) between $V_{O2}$ measured from the reference method and estimated by 20-m shuttle run test from 4 linear equations ($n = 115$).

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Slope</th>
<th>$R$ (95% CI)</th>
<th>SEE</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barnett et al. (5)</td>
<td>−14.14</td>
<td>1.40†</td>
<td>0.72 (0.50–0.85)</td>
<td>5.17</td>
<td>12.8</td>
</tr>
<tr>
<td>Léger et al. (17)</td>
<td>−3.57</td>
<td>1.15†</td>
<td>0.49 (0.24–0.68)</td>
<td>6.49</td>
<td>16.3</td>
</tr>
<tr>
<td>Mahar et al. (19)</td>
<td>−13.99</td>
<td>1.28†</td>
<td>0.71 (0.49–0.85)</td>
<td>5.22</td>
<td>11.9</td>
</tr>
<tr>
<td>Matsuzaka et al. (20)</td>
<td>0.58</td>
<td>1.02†</td>
<td>0.69 (0.47–0.83)</td>
<td>5.39</td>
<td>13.3</td>
</tr>
<tr>
<td>Total ($n = 115$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barnett et al. (5)</td>
<td>−12.29†</td>
<td>1.32†</td>
<td>0.79 (0.70–0.85)</td>
<td>5.81</td>
<td>13.1</td>
</tr>
<tr>
<td>Léger et al. (17)</td>
<td>−12.85</td>
<td>1.45†</td>
<td>0.60 (0.46–0.71)</td>
<td>7.59</td>
<td>18.6</td>
</tr>
<tr>
<td>Mahar et al. (19)</td>
<td>−16.34†</td>
<td>1.33†</td>
<td>0.80 (0.72–0.86)</td>
<td>5.69</td>
<td>12.1</td>
</tr>
<tr>
<td>Matsuzaka et al. (20)</td>
<td>−5.81</td>
<td>1.23†</td>
<td>0.77 (0.69–0.84)</td>
<td>5.97</td>
<td>14.1</td>
</tr>
</tbody>
</table>

*The CV refers to the effect of the respective SEE's on the mean $V_{O2}$ estimated for each equation, for boys and girls ((SEE/$V_{O2}$ estimated) × 100). SEE = standard error of estimate; CV = coefficient of variation.

†Significantly different from 0, $p < 0.05$.

**Statistical Analyses**

The data were stored and analyzed using the Statistical Package for the Social Sciences (SPSS for Windows Version 17.0) and MedCalc (Version 10.1.2.0). Normality was checked by Kolmogorov-Smirnov test. In the variables body mass and body mass index, the data were transformed logarithmically. Thus, data are presented in terms of means and SD. Student’s $t$-test for independent samples was used for comparisons between sexes. Analysis of variance (ANOVA) for repeated measures was used for comparisons between $V_{O2}$peak measured by the reference method and $V_{O2}$peak estimated by 4 different equations for the 20-m shuttle run test according to sex. In variables that sphericity was violated as indicated by Mauchly’s test, the analyses were adjusted using a Greenhouse-Geisser correction. A post hoc Bonferroni’s test was used to identify the differences. The relationship between $V_{O2}$peak measured by the reference method (dependent variable) and $V_{O2}$peak estimated by each equation for the 20-m shuttle run test (independent variable) was measured by the Pearson’s linear correlation coefficient and also by simple linear regression analysis, indicating the following values: intercept, slope, coefficient of correlation ($r$), and standard error of estimate (SEE). In addition, the CV was determined by dividing SEE by the average $V_{O2}$peak estimated by field tests, multiplied by 100. To verify the agreement between

**TABLE 4.** Agreement values (bias, limits, and trend) between $V_{O2}$ measured from the reference method and estimated by 20-m shuttle run test from 4 linear equations ($n = 115$).

<table>
<thead>
<tr>
<th></th>
<th>Bias</th>
<th>Limits</th>
<th>Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys ($n = 61$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barnett et al. (5)</td>
<td>2.19*</td>
<td>15.3; −10.9</td>
<td>0.79†</td>
</tr>
<tr>
<td>Léger et al. (17)</td>
<td>8.36*</td>
<td>23.6; −6.9</td>
<td>0.75†</td>
</tr>
<tr>
<td>Mahar et al. (19)</td>
<td>−0.25</td>
<td>12.4; −12.9</td>
<td>0.72†</td>
</tr>
<tr>
<td>Matsuzaka et al. (20)</td>
<td>6.22*</td>
<td>17.8; −5.4</td>
<td>0.57†</td>
</tr>
<tr>
<td>Girls ($n = 54$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barnett et al. (5)</td>
<td>1.89*</td>
<td>12.4; −8.6</td>
<td>0.71†</td>
</tr>
<tr>
<td>Léger et al. (17)</td>
<td>2.45*</td>
<td>15.1; −10.2</td>
<td>0.73†</td>
</tr>
<tr>
<td>Mahar et al. (19)</td>
<td>−1.60*</td>
<td>8.8; −12.0</td>
<td>0.66†</td>
</tr>
<tr>
<td>Matsuzaka et al. (20)</td>
<td>1.49*</td>
<td>12.0; −9.0</td>
<td>0.48†</td>
</tr>
</tbody>
</table>

*Significantly different from 0, $p < 0.05$.

†$p < 0.001$. 
measured \( \text{VO}_2 \) peak and that estimated the 4 equations in the 20-m shuttle run test, the analysis proposed by Bland and Altman (8) was used, which allowed assessing the bias values and upper and lower limits. The trend was analyzed by measuring the correlation between the difference of methods and the average between them. Significant values were considered when \( p < 0.05 \). The sample size was estimated according to the calculation for comparing dependent means (27), with a power of 90% and an alpha error of 5%. Based on the literature (26) and on values obtained in the pilot study, an expected difference of 3.0 ml·kg\(^{-1} \)·min\(^{-1} \) and a SD of the difference of 4.0 ml·kg\(^{-1} \)·min\(^{-1} \) were considered.

**RESULTS**

The descriptive characteristics of the sample stratified by sex are shown in Table 1. No significant difference was identified between boys and girls for age, body mass, stature, and BMI (\( p > 0.05 \)).

Comparisons between \( \text{VO}_2 \) peak measured from the reference method and estimated by the 20-m shuttle run test using 4 different equations are present in Table 2. For boys, \( \text{VO}_2 \) peak estimated only by the equation proposed by Mahar et al. (19) was not considered different (\( p > 0.05 \)) from the measured value, a fact also observed in the analysis of total sample. Moreover, all estimates by the 20-m shuttle run test
were considered different from each other \((p < 0.05)\). In addition, for girls, the \(\text{V}_{O2\text{peak}}\) measured by the reference method was considered different from all estimates provided by the 20-m shuttle run test. Among the estimates provided by different equations, only the 1 proposed by Mahar et al. (19) evaluated \(\text{V}_{O2\text{peak}}\) differently from the others \((p < 0.05)\).

In addition to the comparisons, a simple linear regression analysis was performed to verify the performance of the 20-m shuttle run test and its 4 different equations (independent variables) as predictor variables of \(\text{V}_{O2\text{peak}}\) measured by the reference method (dependent variable). Table 3 shows the regression results between \(\text{V}_{O2\text{peak}}\) measured and estimated by 4 linear equations from the 20-m shuttle run test. Considering the total sample and stratified by sex, the intercept values were not different from the zero value of the identity line \((p > 0.05)\) for most analyses, except for \(\text{V}_{O2\text{peak}}\) estimated by the equation of Barnett et al. (5) and Mahar et al. (19) for boys and total sample. Moreover, the inclination data were statistically different from the one \((p < 0.05)\) in all analyses. The correlation coefficient values ranged from 0.49 to 0.80, with higher values found for the equation of Matsuzaka et al. (20) for boys \((r = 0.80)\), Barnett et al. (5) for girls \((r = 0.72)\) and Mahar et al. (19) for the total analysis \((r = 0.80)\). Corroborating these results, the lowest SEE values for boys, girls, and total sample were obtained with the same equations \((5.81 \text{ml-kg}^{-1}\text{-min}^{-1}, 5.17 \text{ml-kg}^{-1}\text{-min}^{-1}, \text{and } 5.69 \text{ml-kg}^{-1}\text{-min}^{-1}, \text{respectively})\). It was also observed that the individual \(\text{V}_{O2\text{peak}}\) values estimated by the 4 equations in the 20-m shuttle run test varied approximately 12–18% of the group mean (CV).

Table 4 shows the values obtained in the analysis of agreement between estimates provided by 4 different equations for the 20-m shuttle run test and measurement using the reference method. The correlation analysis showed that the mean differences (bias) for boys were not considered different from zero only for the estimation of \(\text{V}_{O2\text{peak}}\) by the equation of Mahar et al. (19), which behavior was also observed when the total sample was considered. For all other analyses, there was a significant bias, i.e., the mean of differences between methods was considered different from zero \((p < 0.05)\). Moreover, the agreement limits were considered large, with higher values for estimate given by the equation of Le\(^{e}\)ger et al. (17) for boys \((8.36 \pm 15.24 \text{ml-kg}^{-1}\text{-min}^{-1})\) and girls \((2.45 \pm 12.63 \text{ml-kg}^{-1}\text{-min}^{-1})\). Measurement trends were observed in all analyses for the 4 equations \((p < 0.001)\), i.e., estimates by the 20-m shuttle run test tend to overestimate the \(\text{V}_{O2\text{peak}}\) of adolescents less physically fit and underestimate \(\text{V}_{O2\text{peak}}\) of those considered more physically fit. It is noteworthy that the lowest values and correlation between mean and the difference of methods was found for the equation of Matsuzaka et al. (20) for boys \((r = 0.57)\) and girls \((r = 0.48)\).

The agreement data of total sample can be seen in Figure 1. The plotting proposed by Bland and Altman (8) confirms the low performance of equations for estimating \(\text{V}_{O2\text{peak}}\) individually, with wide agreement limits and measurement trend for the 4 equations.

**Discussion**

The main finding of this study was that the estimates provided by 4 regression equations analyzed are imprecise for \(\text{V}_{O2\text{peak}}\) prediction. Although some reasonable predictions in the group analysis have been found, as in the equation proposed by Mahar et al. (19) for boys and total sample, the accuracy of individual estimates was very low for all equations. Our finding was similar to a previous study (22) that showed a systematic bias of regression equations investigated to estimate the \(\text{V}_{O2\text{peak}}\) from the 20-m shuttle run test in 8- to 10-year-old children, resulting in the overestimation for youth considered less physically fit and underestimation for those more physically fit.

Although many studies of cross-validation have been performed with several equations to estimate the \(\text{V}_{O2\text{peak}}\) from the 20-m shuttle run test \((1,5,18,19,22,24,29,30,33)\), the most investigations have only used Pearson’s inter-class correlation coefficient, paired \(t\)-test or ANOVA for repeated measures for the analysis of the results. Considering that additional procedures are required to verify the data consistency, it remains unknown whether these equations are able to estimate with precision the \(\text{V}_{O2\text{peak}}\). In this study were also analyzed intercept, slope, SEE, CV, systematic bias and random error, and homoscedasticity, important criteria to evaluate the performance of equations investigated.

The original equation for the 20-m shuttle run test proposed by Léger et al. (17) showed in our study a poor performance \((r = 0.60, \text{SEE} = 7.59 \text{ml-kg}^{-1}\text{-min}^{-1}, \text{CV} = 18.69\%)\) and low agreement values \((\text{bias} = 5.58 \text{ml-kg}^{-1}\text{-min}^{-1}, \text{limits: } 20.76/-9.59 \text{ml-kg}^{-1}\text{-min}^{-1}, \text{trend: } r = 0.78, p < 0.001)\) when compared with the 3 equations applied for estimating \(\text{V}_{O2\text{peak}}\). Analogous studies \((5,19,29,30)\) have also confirmed the potential systematic underestimation of \(\text{V}_{O2\text{peak}}\) values estimated by the equation of Léger et al. (17). However, when analyses were stratified by sex, the equation of Léger et al. (17) yields better estimates of \(\text{V}_{O2\text{peak}}\) for girls, information similar to that previously reported by Melo et al. (22).

The discrepancy between \(\text{V}_{O2\text{peak}}\) values estimated by the equation of Léger et al. (17) and other equations for the 20-m shuttle run test may be, at least in part, explained by aspects such as testing protocol and the form of direct \(\text{V}_{O2\text{peak}}\) analysis. The original equation used the speed achieved in the last stage (1 minute) as test parameter, so if an individual interrupts the test a few seconds from finishing a stage, only the speed of the previous stage will be taken into account. This can produce errors in estimating the \(\text{V}_{O2\text{peak}}\) because, in fact, the maximum speed corresponding to the moment in which the individual interrupted the test was not considered. Thus, some studies have considered the count of at least half of the stage (30 seconds) in which the individual completed
the test, so that an increased accuracy in estimating \( V_{\text{O}_2\text{peak}} \) is obtained (29,32).

The low performance of equation proposed by Léger et al. (17) for estimating \( V_{\text{O}_2\text{peak}} \) in youth, according to some studies in the literature (5,19,22,29,30), has led to the development of new equations using other variables, not only age and maximum speed. The equations analyzed in our study included variables such as sex (5,19,20), body mass (5,19), BMI (20), and also the number of laps obtained during the test (19). It has been established that in addition to chronological age, biological maturation directly influences \( V_{\text{O}_2\text{peak}} \) and the performance in motor tests (3).

Another interesting aspect to be considered is how the reference measurement of \( V_{\text{O}_2\text{peak}} \) is obtained because the motor task involved in the performance of 20-m shuttle run test seems to be greater complexity than that required by treadmill running protocols used for the direct evaluation of \( V_{\text{O}_2\text{peak}} \) in the laboratory. Adolescents started the 20-m shuttle run test at speed from moderate to high (8.5 km·h\(^{-1}\)) and need to accelerate and decelerate every time to meet the coming and going path (20 m) determined in the test. This was investigated by Metsios et al. (23), who obtained lower \( V_{\text{O}_2\text{peak}} \) values in the modified Bruce protocol in treadmill (48.1 ± 2.39 ml·kg\(^{-1}\)·min\(^{-1}\), \( p < 0.001 \)) than those directly measured from the 20-m shuttle run test (52.3 ± 2.28 ml·kg\(^{-1}\)·min\(^{-1}\)). A possible alternative to solve the problem of differences found between motor tasks required from the treadmill and 20-m shuttle run test has been using of portable gas analyzers that allow measuring the \( V_{\text{O}_2\text{peak}} \) directly during the performance in field test (29,30). However, it was not possible to adopt this procedure in our study.

Additionally, other possible factors may contribute to the comprehension of results produced by the different equations for estimating \( V_{\text{O}_2\text{peak}} \) from the 20-m shuttle run test in this work. The mechanical efficiency or economy of movement, together with \( V_{\text{O}_2\text{peak}} \) and aerobic capacity, determines CRF and cannot be measured only by controlling chronological age (variable present in some equations) (28). Motivating youth to perform the 20-m shuttle run test also seems to be quite important because the results depend on the subjects’ subjective perception to achieve the maximum effort required (22). Thus, one should be careful and aware about the standardized instructions provided by the evaluators before and during the test. Therefore, the importance of checking the quality of measurements in repeated situations also should be highlighted.

The results of this study indicated that among the 4 equations selected, the equation proposed by Mahar et al. (19) in a sample of American adolescents aged 12–14 years using the number of laps obtained in the test was the only one that overestimated the \( V_{\text{O}_2\text{peak}} \) values. However, no difference was found between means when the value measured by the reference method for boys (bias = –0.25 ml·kg\(^{-1}\)·min\(^{-1}\), \( p > 0.05 \)) and total sample (bias = –0.89 ml·kg\(^{-1}\)·min\(^{-1}\), \( p > 0.05 \)) was contrasted. Moreover, this equation showed lowest agreement limits for boys (–0.25 ± 12.68 ml·kg\(^{-1}\)·min\(^{-1}\)), girls (–1.60 ± 10.39 ml·kg\(^{-1}\)·min\(^{-1}\)), and total sample (–0.89 ± 11.69 ml·kg\(^{-1}\)·min\(^{-1}\)).

In the study by Mahar et al. (19), the results found were considered adequate for the sample analyzed (\( r = 0.65 \), \( \text{SEE} = 6.38 \text{ ml·kg}^{-1} \cdot \text{min}^{-1} \)). These values were less accuracy for estimating \( V_{\text{O}_2\text{peak}} \) than those found in our study (\( r = 0.80 \), \( \text{SEE} = 5.69 \text{ mL·kg}^{-1} \cdot \text{min}^{-1} \)). However, the original study (19) did not use the agreement analysis to assess the performance of equation for estimating \( V_{\text{O}_2\text{peak}} \) individually.

It is also noteworthy that all equations showed measurement trend in our study, but despite significant (\( p < 0.001 \)), the lowest correlation values between mean and the difference of methods were those observed for estimates provided by the equation of Matsuzaka et al. (20) for boys (\( r = 0.57 \)), girls (\( r = 0.48 \)), and total sample (\( r = 0.60 \)). Similarly, Ruiz et al. (30) also investigated the validity of equations for estimating \( V_{\text{O}_2\text{peak}} \) from the 20-m shuttle run test in youth aged 13–19 years. The authors found correlation coefficient (\( r \)), \( \text{SEE} \), and mean differences between measured and estimated \( V_{\text{O}_2\text{peak}} \) values very similar to those found in this study. However, the work of Ruiz et al. (30) identified no measurement trend, i.e., the estimation errors were independent from the physical fitness. This may be related to methodological issues such as direct \( V_{\text{O}_2\text{peak}} \) measurement that was carried out by portable gas analyzer during the performance of the 20-m shuttle run test protocol in field situation.

In conclusion, the data shown in this work indicate wide variation in the accuracy and agreement between equations used in estimating \( V_{\text{O}_2\text{peak}} \) from the 20-m shuttle run test. Thus, our results suggest that the equations of Barnett et al. (5), Léger et al. (17), Mahar et al. (19), and Matsuzaka et al. (20) are inadequate to assess the CRF from the 20-m shuttle run test in adolescents aged 11–13 years. Therefore, one should give special attention when comparing results from studies using different models for estimating \( V_{\text{O}_2\text{peak}} \) from the 20-m shuttle run test.

**Practical Applications**

The field tests used to evaluate CRF area viable alternative to check this component in various populations because it can be administered in a relatively small space either indoors or outdoors and in places without extensive facilities. However, it is unclear which regression equation is most accurate for estimating \( V_{\text{O}_2\text{peak}} \) from the 20-m shuttle run test in adolescents. In this study, all equations analyzed showed low performance and agreement in estimating \( V_{\text{O}_2\text{peak}} \) individually. Thus, although the accuracy of individual \( V_{\text{O}_2\text{peak}} \) predictions is unacceptable for all equations analyzed, when the assessment is addressed to the school environment, our results suggest the use of equation of Mahar et al. (19) because it seems to show the best estimates considering total group (males and females) or only male adolescents.
ACKNOWLEDGMENTS

We would like to express thanks to all the participants for their engagement in this study and also the Foundation for the Support of Scientific and Technological Development of Paraná (FAADCT/Brazil) and National Council of Technological and Scientific Development (CNPq/Brazil) for the grants conceded to E.R.V.R and E.S.C., respectively. The authors recognize no conflicts of interest. The results of this study do not constitute endorsement of the National Strength and Conditioning Association.

REFERENCES