A non-exercise prediction model for estimation of cardiorespiratory fitness in adults

Equações de predição da aptidão cardiorrespiratória de adultos sem teste de exercícios físicos

Juan Marcelo Simões Cáceres¹
Anderson Zampier Ulbrich¹
Tiago Facchini Panigas¹
Magnus Benetti¹

Abstract – The most accurate tool for assessment of cardiorespiratory fitness is cardiopulmonary exercise testing (CPET). However, CPET requires expensive equipment, trained technicians and time, which limits their use in population studies. In view of this issue, the present study aims to develop regression equations for predicting the cardiorespiratory fitness of adults using simple measurement variables. The study used data from 8,293 subjects, 5,291 male and 3,235 female (age range, 18 to 65 years). The sample was recruited in Florianópolis, Santa Catarina. To develop equations for prediction of peak oxygen uptake (VO₂peak), the data associated were: fitness, age, body mass, height, resting heart rate, hypertension, diabetes, dyslipidemia and smoking. After statistical analyses, two equations for men and two for women were developed. The complete equations showed an adjusted R² = 0.531 and a standard error of estimate (SEE) = 7.15 ml⁻¹·kg⁻¹·min for men and R² = 0.436 and SEE = 5.68 ml⁻¹·kg⁻¹·min for women. We conclude that the model developed for prediction of cardiorespiratory fitness is feasible and practical for prediction of VO₂peak in epidemiological studies or when CPET cannot be performed.

Key words: Cardiorespiratory fitness; Regression equation; VO₂peak.

Resumo – A ferramenta mais precisa para avaliação da aptidão cardiorrespiratória é o teste cardiopulmonar de esforço. Entretanto, para sua utilização, são necessários equipamentos de custo elevado, técnicos bem treinados e tempo, restringindo sua utilização em estudos populacionais. Desta forma, o objetivo do estudo foi desenvolver modelos de predição da aptidão cardiorrespiratória de adultos, por meio de variáveis de simples mensuração. Foram utilizados os dados de 8.293 sujeitos, sendo 5.291 homens e 3.235 mulheres. A amostra constituí-se de dados retrospectivos da cidade de Florianópolis – SC, abrangendo sujeitos entre 18 e 65 anos. Para estimar o consumo de oxigênio de pico (VO₂peak), mensurado de maneira direta, foram associados os dados de: idade, massa corporal, condicionamento, estatura, frequência cardíaca pré-esforço, dislipidemia, hipertensão arterial, tabagismo, diabetes. Após a realização dos procedimentos estatísticos por intermédio de regressão linear múltipla, foram desenvolvidas duas equações para o sexo masculino e duas para o sexo feminino. O modelo completo para o sexo masculino apresenta R² ajustado de 0,531 e erro padrão de estimativa (EPE) de 7,15 ml⁻¹·kg⁻¹·min, enquanto que o modelo completo para o sexo feminino apresenta R² ajustado de 0,436 e EPE de 5,68 ml⁻¹·kg⁻¹·min. Conclui-se que o modelo desenvolvido de predição da aptidão cardiorrespiratória é uma alternativa viável e prática para predição do VO₂peak em estudos epidemiológicos ou quando um teste cardiopulmonar de esforço não for possível ou acessível.

Palavras-chave: Aptidão cardiorrespiratória; Equação de regressão; VO₂peak.
INTRODUCTION

Cardiorespiratory fitness (CRF) is considered a physiological determinant of middle- and long-distance running performance. Its use, however, is not restricted to sports performance. It may also be used as a diagnostic measure of health and for prescribing physical exercise.

The determination of oxygen uptake (VO2) is the primary criterion for measuring cardiorespiratory fitness; the cardiopulmonary exercise test with direct gas analysis is the gold standard for measuring VO2. However, the use of this test is limited by equipment cost, space required for the equipment, and the time and qualified professionals required for the evaluations.

The limitations associated with this method stimulated the search for alternative methods for determination of VO2. The methods include prediction models without physical exercise, which are low-cost, time effective, and can be used in large population studies. Epidemiological studies have shown that low VO2 levels are associated with an increase in cardiovascular diseases, diabetes mellitus, and some types of cancer. Low VO2 levels are also considered an independent risk factor for death due to all causes. These associations justify the evaluation of VO2 with large population samples.

In Brazil, there is a need for studies that investigate alternative methods. The aim of the present study is to develop equations to predict CRF using variables that are easy to measure and do not require exercise testing.

METHODS

Participants

We carried out a cross-sectional, retrospective study with a descriptive correlation design. The study included 8,293 participants out of which 5,597 were physically active (1,969 women and 3,628 men) and 2,696 were sedentary (1,188 women and 1,508 men). The data were obtained from the CardioSport Clinic data base, in the city of Florianópolis, Santa Catarina, Brazil. The evaluations were carried out between January 2004 and December 2010. All participants had signed an informed consent form that allowed their data to be used in population studies. The research was approved by the Ethics Committee for Research with Human Beings of the Universidade do Estado de Santa Catarina, protocol number 97/2010.

Data collection

Body mass (BM) was measured using a Filizola scale with 100-g resolution; stature was measured using a SANNY stadiometer with 0.1-cm accuracy. The measures were collected following Anthropometric Standardization Reference Manual criteria.

To measure the pre-test heart rate (HR_{pre-t}) we obtained electrocardiograms after a five-minute rest with participants in a sitting position (three-derivation ECG Elite, Micromed, Brazil). This measurement preceded the
For the study, we used the lowest heart rate during the evaluation.

Sex, smoking habits (yes or no), age, and physical activity status (sedentary or active) were obtained from patient medical history. Participants were considered physically active based on the criteria reported by Blair et al.21: 30 min of exercise per day.

The data on arterial hypertension, diabetes, and dyslipidemia were obtained from the data base; they were obtained by specialists in the area following the recommendations of the Brazilian Society of Cardiology22.

Peak oxygen uptake (VO2 peak) was determined by uphill treadmill walking. The test did not have stages, only constant load increases (increase in treadmill slope and speed) gradually distributed during the test period. The ratio of load increase is defined individually, and it lasted from eight to 12 minutes. The participants started the test walking (speed between three and six km/h) with 0.0% treadmill slope. The test parameters evolved according to each participant’s physical condition and following criteria established by the Brazilian Cardiology Society (II Guidelines for Ergometric Tests)23.

We used a motorized treadmill (Imbrasport-ATL, Brazil 1999) and three-derivation ECG (ECG Digital; Micromed® - Brasília, DF - Brazil). The analysis of VO2 peak was carried out using a mixing chamber (Metalyzer II, Cortex™ Germany, 2003) and the Ergo PC Elite version 3.3.6.2 (Micromed™, Brazil) software. The VO2 values are reported relative to body mass (ml^-1·kg^-1·min).

The test was interrupted when two of the following three criteria were met: maximum exhaustion (fatigue or dyspnea); respiratory rate above 1.15; progressive angina that hindered continuation of the test; indication of significant change in ECG. The VO2 peak was determined based on visual inspection of the behavior of oxygen uptake curves and carbon dioxide output relative to body mass.

### Data Analysis

Data were analyzed using SPSS version 17.0 for Windows™. The normal distribution of the independent variables was assessed using the Kolmogorov-Smirnov test. Descriptive analysis was carried out to characterize the population sample; data are reported with averages, standard deviation, and minimum and maximum values. The contingency table chi-square test was used to measure the dispersion between categorical variables.

The dependent variable was the VO2 peak, the independent variables were sex, physical activity status, age, body mass index (BMI), HR pre-t, smoking habits, diabetes, hypertension, and dyslipidemia.

The equation was developed using multiple linear regression (stepwise regression) and 95.0% confidence interval. To validate the equation we calculated Person’s correlation coefficient, the coefficient of determination (adjusted R2), and the standard error of estimate (SEE). To estimate strength of prediction we calculated the beta weights (β weights) for the independent variables. The significance value adopted was p < 0.05.
RESULTS

Table 1 shows the description of the population sample (n = 8,293) variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Average</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women n=3,157</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>18.02</td>
<td>64.99</td>
<td>40.45</td>
<td>12.40</td>
</tr>
<tr>
<td>Stature (cm)</td>
<td>142.50</td>
<td>190.00</td>
<td>162.71</td>
<td>6.50</td>
</tr>
<tr>
<td>BM (kg)</td>
<td>39.00</td>
<td>111.00</td>
<td>62.64</td>
<td>10.50</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>15.99</td>
<td>39.84</td>
<td>23.68</td>
<td>3.90</td>
</tr>
<tr>
<td>HR pre-t (bpm)</td>
<td>41.00</td>
<td>130.00</td>
<td>83.90</td>
<td>13.80</td>
</tr>
<tr>
<td>VO₂ peak (ml⁻¹∙kg⁻¹∙min⁻¹)</td>
<td>16.00</td>
<td>62.84</td>
<td>31.38</td>
<td>7.60</td>
</tr>
</tbody>
</table>

| Men n=5,136       |         |         |         |       |
| Age (years)       | 18.00   | 64.99   | 39.58   | 12.50 |
| Stature (cm)      | 145.00  | 208.00  | 176.31  | 7.20  |
| BM (kg)           | 42.20   | 145.00  | 81.17   | 12.60 |
| BMI (kg/m²)       | 16.73   | 39.89   | 26.09   | 3.60  |
| HR pre-t (bpm)    | 41.00   | 130.00  | 77.39   | 13.30 |
| VO₂ peak (ml⁻¹∙kg⁻¹∙min⁻¹) | 16.05   | 75.57   | 40.65   | 10.40 |

Legend: n = number of participants; BM = body mass; SD = standard deviation; HR pre-t = pre-test heart rate; BMI = body mass index; VO₂ peak = peak oxygen uptake

The highest percentage of dyslipidemic and diabetic participants was observed among men. The other comorbidities did not show statistically significant differences between the sexes.

Table 2. Percentage of comorbidities and difference between men and women

<table>
<thead>
<tr>
<th></th>
<th>Men (n=5,136)</th>
<th>Women (n=3,157)</th>
<th>χ²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>394</td>
<td>7.7</td>
<td>4,628</td>
<td>92.3</td>
</tr>
<tr>
<td>Hypertension</td>
<td>568</td>
<td>11.1</td>
<td>4,454</td>
<td>88.9</td>
</tr>
<tr>
<td>Smoking</td>
<td>222</td>
<td>4.3</td>
<td>4,800</td>
<td>95.7</td>
</tr>
<tr>
<td>Diabetes</td>
<td>114</td>
<td>2.2</td>
<td>4,908</td>
<td>97.8</td>
</tr>
</tbody>
</table>

Legend: n = number of participants; *p<0.05; χ²: chi-square.

Box 1 shows the Female Gender-Specific Equation 1 (F1), Female Gender-Specific Equation 2 (F2), Male Gender-Specific Equation 1 (M1), and Male Gender-Specific Equation 2 (M2). M1 includes nine variables and for F1 the diabetes variable was removed (it was not statistically significant and was thus removed from the model). For equations M1 and F1 the adjusted R² were 0.551 and 0.715 and the SEE 7.309 and 5.687 ml⁻¹∙kg⁻¹∙min⁻¹, respectively. For M2 and F2 we removed the variables HR pre-t, dyslipidemia, diabetes, and arterial hypertension to make the equation simpler and the calculation faster. M2 and F2 had an adjusted R² of 0.510 and 0.425 and SEE of 7.309 and 5.743 ml⁻¹∙kg⁻¹∙min⁻¹, respectively. The independent variables were statistically significant predictors (p<0.001) for cardiorespiratory fitness in all models.
Box 1. Gender-specific equations for predicting cardiorespiratory fitness

<table>
<thead>
<tr>
<th>Equation</th>
<th>R</th>
<th>Adjusted R²</th>
<th>SEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VO₂peak = 47.189 – 0.394 (age) – 0.282 (BM) – 4.289 (physical activity) + 0.231 (stature) – 0.090 (HRpre-t) – 2.901 (dyslipidemia) – 1.925 (hypertension) – 2.95 (diabetes)</td>
<td>0.729</td>
<td>0.531</td>
<td>7.150 ml⁻¹·kg⁻¹·min</td>
</tr>
<tr>
<td>M2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VO₂peak = 39.390 – 0.409 (age) – 0.307 (BM) – 4.437 (physical activity) + 0.254 (stature) – 3.081 (smoking)</td>
<td>0.714</td>
<td>0.510</td>
<td>7.309 ml⁻¹·kg⁻¹·min</td>
</tr>
<tr>
<td>F1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VO₂peak = 37.844 – 0.250 (age) – 0.208 (BM) – 3.428 (physical activity) + 0.139 (stature) – 0.053 (HRpre-t) – 1.327 (dyslipidemia) – 1.508 (smoking) – 1.009 (hypertension)</td>
<td>0.661</td>
<td>0.436</td>
<td>5.687 ml⁻¹·kg⁻¹·min</td>
</tr>
<tr>
<td>F2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VO₂peak = 31.733 – 0.244 (age) – 0.219 (BM) – 3.598 (physical activity) + 0.151 (stature) – 1.486 (smoking)</td>
<td>0.652</td>
<td>0.425</td>
<td>5.743 ml⁻¹·kg⁻¹·min</td>
</tr>
</tbody>
</table>

Legend: M1: Male equation 1; M2: Male equation 2; F1: Female equation 1; F2: Female equation 2; r: correlation; R² adjusted: adjusted coefficient of determination; SEE: standard error of estimate; VO₂ peak: peak oxygen uptake (ml⁻¹·kg⁻¹·min); Values applied in the formula: age in years; stature in centimeters (cm); body mass in kg; physical activity: sedentary (0) and active (1); dyslipidemia: no (0) and yes (1); smoking: no (0) and yes (1); hypertension: no (0) and yes (1); diabetes: no (0) and yes (1) and FCpre-t: value in beats per minute (bpm).

The data in Box 2 show that in terms of the strength of prediction of the independent variables (β weights), age, BM, and physical activity status contributed the most in the prediction.

Box 2. β weights for the three values with the highest strength of prediction

<table>
<thead>
<tr>
<th>Variable</th>
<th>Equation</th>
<th>F1</th>
<th>F2</th>
<th>M1</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td>-0.408*</td>
<td>-0.398*</td>
<td>-0.474*</td>
<td>-0.491*</td>
</tr>
<tr>
<td>BM</td>
<td></td>
<td>-0.290*</td>
<td>-0.304*</td>
<td>-0.339*</td>
<td>-0.369*</td>
</tr>
<tr>
<td>Physical activity</td>
<td></td>
<td>-0.219*</td>
<td>-0.230*</td>
<td>-0.187*</td>
<td>-0.198*</td>
</tr>
</tbody>
</table>

Legend: F1: Female equation 1; F2: Female equation 2; M1: Male equation 1; M2: male equation 2; BM: body mass; BMI: body mass index; * p < 0.001

DISCUSSION

CRF is considered an indicator of risk for development of cardiovascular diseases and other chronic degenerative diseases. However, its evaluation and use, either by direct gas analysis or equations based on physical exercise, is limited by equipment cost, space required for the equipment, and the time and qualified professionals required for the evaluations15. The evaluation of CRF using regression analyses and variables that do not include physical exercise is a cost-effective and practical alternative for the other methods; it may also be effective for epidemiological studies3,16,17.

Matthews et al.24 concluded that VO₂ may be predicted without physical exercise in epidemiological studies. The participants evaluated were divided into quintiles according to their CRF (measured directly). The model showed that 83.0% of all subjects were either classified correctly or within one quintile of measured CRF. Misclassification of a low fit individual as high fit was only observed in 0.13% of cases24.

Whaley et al.25 agreed that models for predicting CRF are valid because they satisfy statistical criteria. However, the authors argued that the models
are insufficiently accurate for predicting CRF in epidemiological studies aimed at evaluating the risk for development of chronic degenerative diseases; an opinion contrary to other studies\textsuperscript{14,16,17}.

Our results show that VO\textsubscript{2} prediction models that do not require physical exercise can produce valid estimates for the VO\textsubscript{2}\textsubscript{peak} for both men and women who are apparently healthy, active or sedentary, and aged 18 to 65 years. Our results are similar to those reported by Blair et al.\textsuperscript{26} (adjusted $R^2 = 0.59$; SEE not reported); Nes et al.\textsuperscript{27} ($R^2 = 0.61$ SEE = 5.70 ml\textsuperscript{1}\textcdot kg\textsuperscript{-1}\textcdot min for men and $R^2 = 0.56$ and SEE 5.15 ml\textsuperscript{1}\textcdot kg\textsuperscript{-1}\textcdot min for women); and the Jackson et al.\textsuperscript{12} equation (adjusted $R^2 = 0.61$ and SEE = 5.70 ml\textsuperscript{1}\textcdot kg\textsuperscript{-1}\textcdot min), which used BMI as an alternative for body fat percentage (adjusted $R^2 = 0.66$ and SEE = 5.35) in the main equation. The reason for using BMI is that it is easy to measure.

We also produced alternative equations (M2 and F2). The alternatives did not compromise accuracy to a great extent, and the variables used are simple to measure and can be used in situations where it is no longer possible to obtain data on HR\textsubscript{pre-ex}, arterial hypertension, diabetes, and dyslipidemia. It is a viable, simple and fast alternative.

Wier et al.\textsuperscript{28} investigated the use of waist circumference as a replacement for body indexes in equations for predicting CRF without exercise. Three models were developed which differed in terms of waist circumference ($R^2 = 0.65$ SEE= 4.80), body fat percentage ($R^2 = 0.67$ SEE = 4.72) and BMI ($R^2 = 0.64$ SEE = 4.90). However, the model systematically overestimated 3 ml\textsuperscript{1}\textcdot kg\textsuperscript{-1}\textcdot min among individuals of low CRF, and 7 ml\textsuperscript{1}\textcdot kg\textsuperscript{-1}\textcdot min among individuals with high CRF. The equation was more accurate for individuals older than 50 years, less physically active, and with VO\textsubscript{2}\textsubscript{max} between 30 and 50 ml\textsuperscript{1}\textcdot kg\textsuperscript{-1}\textcdot min\textsuperscript{27}. Nes et al.\textsuperscript{27} also obtained a high accuracy for values between 35 and 50 ml\textsuperscript{1}\textcdot kg\textsuperscript{-1}\textcdot min for men, and 30 and 40 ml\textsuperscript{1}\textcdot kg\textsuperscript{-1}\textcdot min for women.

In order to be useful as a tool for stratifying health risks, the model must be able to identify low CRF individuals\textsuperscript{27}. The models\textsuperscript{27,28} show that the values in the lower tiers of CRF may misclassify individuals at risk as healthy; therefore, the models should be used with caution. One of the limitations of the present study was that we did not investigate the loss of accuracy at the extremes of VO\textsubscript{2}\textsubscript{peak}.

The beta weights (Picture 2) show that the most important variable for predicting CRF was age, with values between -0.398 and -0.491. These values are similar to those reported by Mailey et al.\textsuperscript{17}: -0.450, and Nest et al.\textsuperscript{27} who obtained scores of -0.435 and -0.436 for men and women, respectively. The results are higher than those obtained by Heil et al.\textsuperscript{13}, Whaley et al.\textsuperscript{25} and Bradshaw et al.\textsuperscript{29}. The literature indicates that VO2 reduces between eight and 10.0% per decade of life after 25 years of age\textsuperscript{30,31}. Ravagnani et al.\textsuperscript{31} observed that an individual aged between 60-69 years has approximately 60.0% of VO\textsubscript{2}\textsubscript{max} than between ages 20-29 years. Therefore, the use of this variable is explained by the broad age range the equations aim to include.

Body mass had the second highest beta weight among the variables. The weights varied between -0.290 and -0.369. These values are lower than the 0.434 score reported by Whaley et al.\textsuperscript{25}. 

\textsuperscript{292}
The independent variables age, body mass, physical activity status, stature, HR_{pre}, and smoking have been widely investigated in studies that predict CRF without physical exercise; the importance of these variables is shown by the statistical results. However, to our knowledge, no other study evaluated the effect of metabolic (dyslipidemia and diabetes) and hemodynamic (arterial hypertension) disorders for the prediction of VO2. Poorly-explored variables may help improve the strength of prediction of models. These variables were statistically significant for the prediction; however, when removed, their impact is relatively small in the reduction of adjusted R^2 and SEE. This is shown by the comparison between equations M1 and M2 (Figure 1). The simpler equation shows a reduction of 0.021 in adjusted R^2 and an increase of 0.159 ml^{-1}kg^{-1}min in SEE; a similar effect can be verified in the F1 and F2 equations (Figure 1), for which the reduction in R^2 is of 0.011 and the increase in SEE is 0.056 ml^{-1}kg^{-1}min. Therefore, the F2 and M2 models may be used without a significant loss in accuracy when it is either not possible or convenient to obtain information on these disorders.

Few studies developed equations for predicting CRF without using tests that involve physical exercise in the Brazilian population. These studies had a reduced population sample in comparison to the present study, and they analyzed specific populations. The importance of the present study is in its population size (8,293 individuals, 3,157 women and 5,136 men, broad age range (19 to 65 years), and CRF range (16 to 75.57 ml^{-1}kg^{-1}min). The combination of these factors may allow for useful generalization of the data. However, if compared to the Barbosa et al. study, the present study showed a reduced accuracy. The authors reported an adjusted R^2 of 0.90 and an SEE of 3.44 ml^{-1}kg^{-1}min. The difference is possibly associated with the reduced range of individuals evaluated in the Barbosa et al. study, and with the classification of CRF into four categories. One of the limitations of the present study is the retrospective characteristic of the variables used.

CONCLUSION

The equations developed in the present study showed satisfactory predictive accuracy based on the statistical requirements for predicting CRF in adults. The equations applied easily-obtained variables and aimed at the evaluation of large population sizes. The equations must be used with caution if the goal is to more accurately predict VO2_{peak} in patients or athletes.

REFERENCES

18. Lima DE, Abatti PJ. Formulação de equação preditiva do VO2máx. baseada em dados que independem de exercícios físicos. Semina 2006;27:139-49.


Address for Correspondence
Juan Marcelo Simões Cáceres
Rod. Amaro Antonio Vieira 2489/107
Bairro Itacorubi.
CEP 88034-102 - Florianópolis, SC.
Brazil.
E-mail: juansimoes@yahoo.com.br