Validity and reliability of two submaximal tests for estimating maximal oxygen consumption in adults

Validez y confiabilidad de dos pruebas submáximas para estimar el consumo máximo de oxígeno en adultos

Authors

Thiago Barbosa Lima [1]
Willemax dos Santos Gomes [1]
Vinícius de Oliveira Damasceno [2]
Tony Meireles Santos [1]
Eduardo Zapaterra Campos [1]

[1] Federal University of Pernambuco (Brazil)
[2] Air Force University (Brazil)

Corresponding author:
Eduardo Zapaterra Campos
eduardo.zapaterracampos@ufpe.br

How to cite in APA

Damasceno, V., Barbosa Lima, T., Meireles Santos, T., dos Santos Gomes, W., & Zapaterra Campos, E. (2025). Validity and reliability of two submaximal tests for estimating maximal oxygen consumption in adults. *Retos*, *69*, 1330–1338. https://doi.org/10.47197/retos.v69.108833

Abstract

Objective: Verify the reliability and validity of two submaximal tests – walking-running test (WRT) and self-selected intensity test (SSIT) – to predict $\dot{V}O_{2MAX}$ in healthy adults.
Methodology: A total of 20 adults of both sexes were evaluated in three visits. The WRT was initially performed at a speed of 5 km·h^{-1} and 4% grade. The grade was increased by 2% until 12%. The speed was increased to 9 km·h^{-1} and the grade reduced to 1% if subjects did not reach 65% of the heart rate reserve (HR_{RES}). Once HR_{RES} was attained, the effort was maintained for 6 minutes. The SSIT consisted in 4 minutes of free intensity selection (speed and/or grade) and 6 minutes for HR monitoring. The Student's t-test, ICC and typical error were used to test reliability. ANOVA and a graphical Bland-Altman analysis were used for validity. Only speed was different in WRT between test-retest.
Results: The main results achieved in the research should be commented on.
Discussion: Differences were found between the measured $\dot{V}O_{2MAX}$ (49.5 ± 7.9 mL·kg^{-1}·min^{-1}) and predicted $\dot{V}O_{2MAX}$ in WRT (43.2 ± 7.6 mL·kg^{-1}·min^{-1}) and SSIT (40.6 ± 9.8 mL·kg^{-1}·min^{-1}).
Conclusions: The two tests proved to be reproducible, especially for the WRT; however, both may present an underestimation of $\dot{V}O_{2MAX}$.

Keywords

Physical fitness; exercise test; oxygen consumption; adult; reproducibility of results .

Resumen

Objetivo: Verificar la fiabilidad y validez de dos pruebas submáximas -prueba de marcha-carrera (WRT) y prueba de intensidad autoseleccionada (SSIT)- para predecir la $\dot{V}O_{2MAX}$ en adultos sanos.
Metodología: Se evaluó a un total de 20 adultos de ambos sexos en tres visitas. El WRT se realizó inicialmente a una velocidad de 5 km·h^{-1} y un grado del 4%. El grado se incrementó en un 2% hasta el 12%. La velocidad se aumentó a 9 km·h^{-1} y el grado se redujo al 1% si los sujetos no alcanzaban el 65% de la reserva de frecuencia cardiaca (RFC). Una vez alcanzada la RFC, el esfuerzo se mantenía durante 6 minutos. El SSIT consistió en 4 minutos de selección libre de la intensidad (velocidad y/o grado) y 6 minutos para la monitorización de la FC. Para comprobar la fiabilidad se utilizaron la prueba t de Student, el CCI y el error típico. Para la validez se utilizaron ANOVA y un análisis gráfico de Bland-Altman. Sólo la velocidad fue diferente en WRT entre test-retest.
Resultados: Se encontraron diferencias entre el $\dot{V}O_{2MAX}$ medido (49,5 ± 7,9 mL·kg^{-1}·min^{-1}) y el $\dot{V}O_{2MAX}$ predicho en WRT (43,2 ± 7,6 mL·kg^{-1}·min^{-1}) y SSIT (40,6 ± 9,8 mL·kg^{-1}·min^{-1}).
Discusión: Se deben contrastar los resultados de la investigación con los de otras investigaciones encontradas en la literatura.
Conclusiones: Las dos pruebas demostraron ser reproducibles, especialmente para la WRT; sin embargo, ambas pueden presentar una subestimación de la $\dot{V}O_{2MAX}$.

Palabras clave

Aptitud Física; prueba de esfuerzo; consumo de oxígeno; adulto; reproducibilidad de los resultados.

Introduction

Cardiorespiratory fitness is considered an important health-related component, and therefore low levels of cardiorespiratory fitness have been associated with an increased risk of sudden death, being important to improve or maintain it through life (Aune et al., 2020; ACSM, 2014; Tauda et al., 2024). Maximal oxygen consumption ($\dot{V}O_{2MAX}$) is accepted as a criterion for the assessment of cardiorespiratory fitness (Tauda, 2024). It is possible to measure or estimate it directly and/or indirectly under maximum and/or submaximal conditions and through questionnaires (Väisänen et al., 2024; Yang et al., 2021; Sartor et al., 2013). Direct measurement of $\dot{V}O_{2MAX}$ requires voluntary exhaustion, a high degree of motivation, specialized personnel, sophisticated equipment, and an appropriate location, in addition to medical-cardiological supervision in adverse health conditions (Sartor et al., 2013; Arena et al., 2007; Guazzi et al., 2017).

Given these aspects, several submaximal tests were developed to minimize the barriers that the maximum tests establish, above all, the exposure time and operational costs (Dugas et al., 2023; Evans et al., 2015; Yang et al., 2021). Åstrand and Rhyming (1954) developed a single-stage submaximal test to predict $\dot{V}O_{2MAX}$ (Åstrand & Rhyming, 1954). This test considered the linear relationship between the increase in heart rate (HR) and oxygen consumption ($\dot{V}O_2$) from submaximal tests on a bench, treadmill, and cycle ergometer. On the other hand, other tests with multiple stages were proposed to evaluate the HR response in different submaximal loads, extrapolating the results based on the HR predicted by the age of the subject, for example in the test of the Young Men's Christian Association (YMCA) (ACSM, 2014). However, Swain and Wright (1997) found that despite the good correlation (r = 0.79 and 0.81), the method overestimated $\dot{V}O_{2MAX}$ by an average of 28%, suggesting that short stages reduce the HR for that intensity, increasing the $\dot{V}O_{2MAX}$ estimate (Swain & Wright, 1997). In this context, Swain et al. (2004) proposed a test with progressive stages (up to 55% of HR_{RES}) followed by a 6-minute stage in steady state between 65 to 75% of HR_{RES}, and found a correlation of 0.89 between the $\dot{V}O_{2MAX}$ estimated with the direct measurement of $\dot{V}O_{2MAX}$ and a standard error of estimative (SEE) of 4.0 $mL\cdot min^{-1}\cdot kg^{-1}$ (≅11%) (Swain et al., 2004). This concept of HR stabilization for 6 minutes was subsequently adapted for running on track and treadmill (Santos et al., 2004; Mainardi, 2016; Beltrame et al., 2017).

A recent study has investigated several models of $\dot{V}O_{2MAX}$ estimation for active/recreational athletes, and observed an overestimation of $\dot{V}O_{2MAX}$ for lower trained subjects (Wiecha et al., 2023). Recently, Dugas et al. (2023) have also observed an overestimation of $\dot{V}O_{2MAX}$ when the HR response during submaximal running was tested (Dugas et al., 2023). This probably happened because 75% of the HR_{RES} was close to the transition from walking to running in less conditioned subjects, which may overestimate the physical capacity of the subject when they choose to run at the same intensity (alteration of k1 and k2) (ACSM, 2014; Swaim et al., 2004; Mainardi, 2016). Therefore, increasing the treadmill slope before speed would avoid possible overestimation in $\dot{V}O_{2MAX}$. Thus, the present study proposes a new protocol with the inclusion of an initial walking phase (with subsequent running, if necessary) to enable evaluating subjects from low to moderate levels of conditioning.

Despite practical applicability to submaximal tests using HR, other studies have proposed to increase exercise intensity based on rate perceived exertion (RPE), and have shown high estimation errors (≈15%) which might be dependent on the adaptation of the subjects to the scale (Ferri Marini et al., 2024; Olher et al., 2019; Coquart et al., 2016; Shea et al., 2022). On the other hand, in self-selected intensity exercises based on sensations (regardless of RPE), subjects usually select intensities below the first ventilatory threshold (HR_{RES} below 75%), which follows what is suggested by ACSM (2014) (ACSM, 2014; Ekkekakis, 2009). Therefore, it would be theoretically possible to estimate $\dot{V}O_{2MAX}$ in a self-selected intensity protocol based on sensations.

Therefore, the first objective of the present study was to verify the reliability and validity of a submaximal walking and running test based on FC_{RES} to predict $\dot{V}O_{2MAX}$, and the second objective of the study was to verify the reliability and predictive validity of a submaximal test with intensities based on self-selection to predict $\dot{V}O_{2MAX}$ in healthy adults.

Method

This section explains how the research was done. The design of the same is described and it is explained how it was put into practice, justifying the choice of the methods used. This section should contain the type of quantitative research, the scope or depth of the research (exploratory, correlational and/or explanatory), population and sample, and the techniques used should be added. This section is fundamental, because it is the one that will allow the scientific community to reproduce the result. Most of this section should be written in the past tense, in a descriptive style.

Participants

A total of 20 adults participated in the present study: 14 men (mean ± standard deviation [SD]: age, 24.9 ± 4.0 years; body mass, 78.1 ± 10.5 kg; height 1.78 ± 0.07 m; body mass index [BMI] 24.8 ± 32 kg·m-2; $\dot{V}O_{2MAX}$ 53.1 ± 6.1 mL·kg-1·min-1) and six women (age, 21.8 ± 2.0 years; body mass, 66.3 ± 3.5 kg; height, 1.66 ± 0.06 m; BMI 24.1 ± 1.5 kg·m^{-2}; $VO_{2MÁX}$ 41.1 ± 4.4 mL·kg^{-1}·min^{-1}). The subjects were selected by non-probabilistic convenience. Subjects who did not present signs and symptoms of cardiovascular, pulmonary, renal, and metabolic diseases and risk factors for cardiovascular diseases were included. In addition, participants did not use medication, did not have any type of physical limitation that restricted their ability to walk or run on a treadmill and were not athletes (professional or amateur).

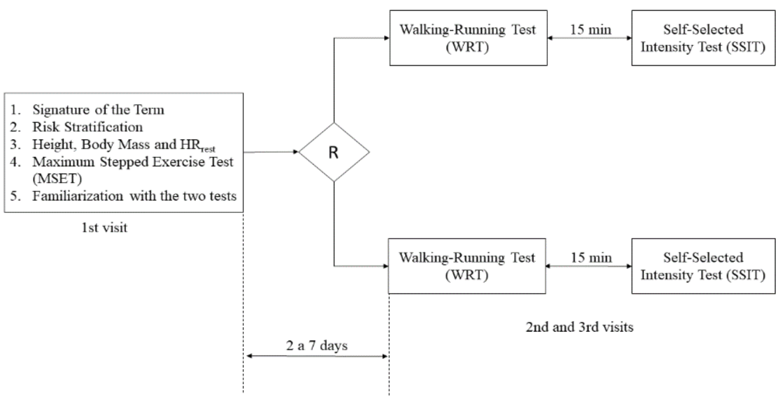
Individuals who did not perform all the scheduled visits or who had injuries during the collection period were excluded. In addition, it is suggested that the subjects did not perform any vigorous exertion 24 h of each test, nor ingest food or stimulants (3 and 6 h, respectively) before the maximal test, based on self-declaration.

Procedure

Experimental design

A cross-sectional study was conducted over three visits. Anthropometric measurements of body mass, height and HRREST were performed on the first visit, after signing the informed consent (ICF) and risk stratification form. Participants subsequently performed a graded exercise test followed by a familiarization with the two submaximal tests (Submaximal Walk/Run Test - WRT and Self-selected Submaximal Test - SSIT). The submaximal tests were performed randomly during the second and third visits with an interval of 15 minutes between them for the same day and two to seven days between visits. They were performed at the same time of day (8:00-12:00 am and 13:00-17:00 pm) (Figure 1). The study protocol was previously approved by the research ethics committee of the local University (protocol #553,851). All tests were performed in an environment with controlled temperature (20 to 22°C) and relative humidity between 40 and 60% (Dry and Humid Bulb Analog Thermo-Hygrometer, Incoterm, Brazil).

Figure 1. Experimental design



Procedures

Heart rate and RPE. Following the recommendations of the ACSM (2014), the resting heart rate (HR_{REST}) was checked in the supine position for 5 min by a frequency meter (FT1, Polar, Finland) prior to the maximal and submaximal tests, as well as being monitored during all tests by a sensor coupled to the gas analyzer . The CR10 Scale was used to measure RPE in the final ten seconds of each stage (Rodriguez-Fuentes et al., 2024).

Anthropometry. Body mass measurements (digital scale, Filizola®, Brazil), height (stadiometer, Sanny®, Brazil) were determined with a precision of 0.1 kg and 0.01 cm, respectively.

Maximum Stepped Exercise Test (MSET). The treadmill exercise protocol (Super ATL, InbraSport®, Brazil) had a three-minute warm-up at 5 km·h-1, with an initial speed of 8 km·h-1 and increments of 1 km ·h-1 every two minutes until maximum voluntary exhaustion. The slope was set at 1%.

Walking-Running Test (WRT). The initial speed was set at 5 km·h^{-1} and the slope at 4%, and each minute the slope increased by 2% until reaching 12%, with the aim of participants reaching 65% of HR_{RES}. If the HR_{RES} value was not reached, the speed was increased to 9 km·h^{-1} and the slope was reduced to 1% with successive increments of 1 km·h^{-1} every minute until reaching 65% of the HR_{RES}. Upon reaching 65% of HR_{RES}, the speed and/or slope were maintained for six minutes to allow HR stabilization at an intensity of 75% HR_{RES}. If the subject did not reach 75% HR_{RES} in the stabilization period, a new increment of 1 km·h^{-1} was administered. The warm-up speed was 5 km·h^{-1} at 1% slope for three minutes.

Self-Selected Intensity Test (SSIT). The SSIT protocol had a total duration of ten minutes. Four minutes were designated for free selection of intensity (speed and/or slope) and the subsequent six minutes were allocated to the stabilization phase at the previously self-selected intensity. Any change in the stabilization phase implied a 6-min time restart. Only the elapsed time was mentioned, and no information regarding the selected intensity was given. The instructions relevant to the selection of intensities followed the basic text below:

"The test you will take is intended to find out what your heart rate responses are at an intensity (speed and/or incline) of your choice for 6 minutes without change. The intensity should be one you prefer that you would do regularly for 20-min or more.

Aim for an intensity at the limit of your comfort. The activity should not cause extreme tiredness. Your breathing and heart rate must be kept under control.

You will do a 4-min warm-up to adjust the intensity before the main activity.

You can modify the selected intensity of the activity whenever you deem necessary, however, remember that if there is any change in the 6 min period, the timer will be restarted.

You will not have access to any relevant information regarding the selected intensity. "

Respiratory exchanges. The $\dot{V}O_{2MAX}$ was determined from the direct monitoring of respiratory gas exchange variables ($\dot{V}O_2$, $\dot{V}CO_2$ and VE) through a VO2000 gas analyzer (Medical Graphics, St.Paul, USA) with an average every ten seconds. The equipment was calibrated before each test based on the manufacturer's instructions. The discontinuation of maximal testing followed the criteria suggested by the ACSM (2014) for non-diagnostic testing (ACSM, 2014). In addition, the test was considered maximum when interrupted by the participant motivated by severe manifestations of fatigue, in addition to HR_{MAX} > 85% of predicted, plateau in oxygen consumption < 150 mL·min-1, respiratory quotient > 1.1 and perception of effort > 9 on the 0-10 scale.

$\dot{V}O_{2MAX}$ prediction. The equation proposed by Swain et al. (2004) (Equation 1) was used to estimate the $\dot{V}O_{2MAX}$ taking into account the average values of HR at the 5th and 6th min, in addition to the velocity and the percentage of slope during the steady state [14].

Equation 1

$$VO2MAX = \frac{\{[(k1\ \times\ vel) + (k2\ \times\ \%inc\ \times\ vel)\ +\ 3.5]\ 3.5\}}{[(HRload - HRrest)/(HRmax - HRrest)]} + 3.5$$

K1 – constant: walking 0.1 or running 0.2;
K2 – constant: walking 1.8 or running 0.9;

$\dot{V}O_{2MAX}$ – maximum oxygen consumption in mL.kg-1.min-1;
HR_{LOAD} – average HR between 5 and 6 min in steady state;
HR_{REST} – HR measured during rest;
HR_{MAX}–highest HR achieved in the maximum progressive test or estimated by age.
Vel – Velocity selected to perform the test on the treadmill;
%Inc - % of the selected slope to perform the test on the treadmill;

Data analysis

The characteristics of the subjects were presented by the mean and standard deviation. The Shapiro-Wilk test was adopted for data normality. A paired Student's t-test between the variables was used in order to compare the test and re-test variables. reliability was analyzed using the intraclass correlation coefficient (ICC), a two-way random model, with mean measurements for absolute agreement [ICC (2, k)], with the following interpretation: values < 0.50 (low reliability), > 0.50 to 0 .75 (moderate reliability), > 0.75 to 0.90 (high reliability) and values > 0.90 (excellent reliability) (Rodriguez-Fuentes et al., 2024; Koo & Li, 2016). In addition to the ICC, the typical measurement error was determined (TME = SDmean × ($\sqrt{1}$ - ICC).

Criterion validity was evaluated by simple linear regression, establishing the coefficient of determination (r^2) and standard error of the estimate (SEE) between the measured maximum consumption values and those estimated by both submaximal tests. The sphericity was evaluated by the Mauchly test with Grennhouse-Geisser correction when necessary, and an ANOVA of repeated measures was performed to compare the values, with Bonferroni posthoc when necessary. In addition, the Bland-Altman plot (1995) was visually analyzed to show the agreement between the prediction models, seeking to reduce the interpretation bias associated only with the regression analysis (Bland & Altman, 1986). All analyzes were performed using the IBM SPSS version 23.0 program and figures were made using Graph Pad Prism version 5.01. The significance level adopted was $p < 0.05$.

Results

Reliability

Table 1 shows the descriptive values referring to mean, standard deviation and the Student's t-test, as well as the values referring to test re-test reliability (ICC, ETM) in the variables HR 5-6min, speed, and $\dot{V}O_{2MAX\_PRED}$ for both tests. Only the stabilization speed showed a significant difference ($t(19) = -2.65$; $p = 0.02$) in the WRT. The ICC values for $\dot{V}O_{2MAX\_PRED}$ in both tests were considered excellent.

Validity

A difference ($F(2, 38) = 16.357$; $p < 0.001$) was found between the measured $\dot{V}O_{2MAX}$ (49.5 ± 7.9 mL·kg^{-1}·min^{-1}) and the $\dot{V}O_{2MAX\_PRED}$ in the WRT (43.2 ± 7.6 mL·kg^{-1}·min^{-1}; $p < 0.0001$) and SSIT (40.6 ± 9.8 mL·kg^{-1}·min^{-1}; $p < 0.0001$). The regression analysis (Figure 2) presents the values of the correlation coefficient (r), the standard error of the estimate (SEE) and the correction equation, in addition to the residual plot to identify homoscedasticity. The WRT $\dot{V}O_{2MAX\_PRED}$ value showed a higher correlation ($F(1.18) = 22.633$; $p < 0.000$; $r = 0.75$; $r2 = 0.56$; SEE = 5.4 mL·kg^{-1}·min^{-1}) with the standard measurement compared to SSIT ($F (1.18) = 8.983$; $p < 0.008$; $r = 0.58$; $r2 = 0.33$; SEE = 6.6 mL·kg^{-1}·min^{-1}). Figure 3 presents the absolute values of the mean differences from the Bland-Altman plot analysis. It was found that there is a high variability among subjects, but they do not show heteroscedasticity. The mean difference for the WRT was 6.3 mL·kg^{-1}·min^{-1} and it was 8.9 mL·kg^{-1}·min^{-1} for the SSIT. The limits of agreement were lower in WRT (-4.5 to 17.1 mL·kg^{-1}·min^{-1}) compared to SSIT (-7.4 to 25.2 mL·kg^{-1}·min^{-1}).

Table 1. Physiological parameters, t-test, and reliability related to WRT, values expressed as mean and standard deviation (SD) (n = 20)

Variables	Day 1	Day 2			Reliability	
	Mean ± SD	Mean ± SD	t(19)	p	ICC (2, k) $CI_{95\%}$	TME
WRT						
$HR_{5\text{-}6min}$ (bpm)	163 ± 6.9	166 ± 9.5	-1.90	0.07	0.60 (0.05 – 0.84)	5.2
Velocity (km·h^{-1})	7.4 ± 2.2	8.3 ± 2.2	-2.65	0.02*	0.80 (0.44 – 0.92)	1.0
$\dot{V}O_{2MAX\_PRED}$ (mL·kg^{-1}·min^{-1})	42 ± 7.0	43.2 ± 7.6	-1.58	0.13	0.88 (0.70 – 0.95)	2.8
SSIT						
HR5-6min (bpm)	164 ± 14.6	167 ± 15.7	-0.83	0.42	0.83 (0.57 – 0.93)	6.3

Velocity (km·h-1)	8.4 ± 1.6	8.6 ± 1.7	-0.99	0.33	0.91 (0.77 – 0.96)	0.5
$\dot{V}O_{2MAX\_PRED}$ (mL·kg-1·min-1)	39.8 ± 10.1	40.6 ± 9.8	-0.64	0.53	0.92 (0.79 – 0.97)	2.9

WRT - walking test with transition to running; SSIT: self-selected intensity test; HR5-6min – average heart rate between the 5th and 6th minute in the stabilization phase; VO2MÁX – predicted maximum oxygen consumption; t – critical t value; p – statistical significance level $p < 0.05$; * - presence of statistical difference; ICC – intraclass correlation coefficient; CI95% - 95% confidence interval; TME – typical measurement error.

Figure 2. Linear regression and standardized residual plot in both submaximal tests. A and C: walking test with transition to running (WTR); B and D: self-selected intensity test (SSIT); r – Pearson's correlation coefficient; R – Coefficient of determination; SEE – Standard error of the estimate.

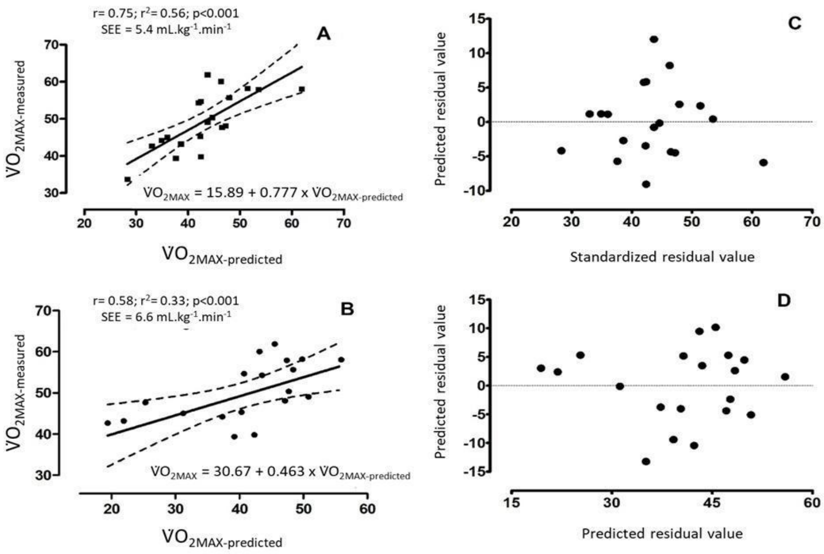
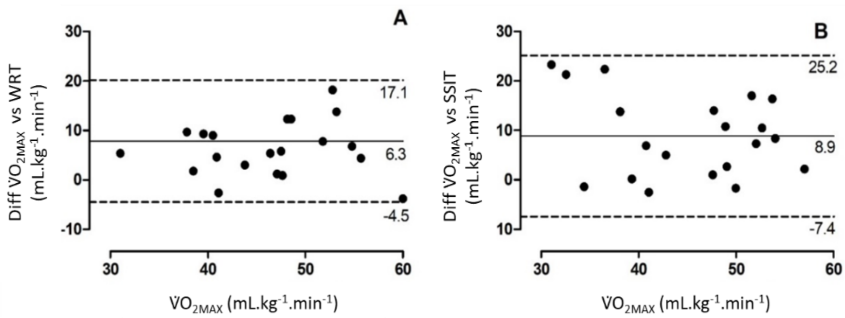


Figure 3. Bland-Altman plot for the absolute values of $\dot{V}O_{2MAX}$ in both submaximal tests. A – walking test with transition to running (WRT); B - self-selected intensity test (SSIT); Solid line – BIAS; Dashed line – 95% confidence interval (CI).



Discussion

The present study aimed to verify the reliability and predictive validity of two submaximal protocols with intensities based on $\%HR_{RES}$ and self-selection to predict $\dot{V}O_{2MAX}$. The main findings are: (i) VO_{2MAX_PRED} both estimated by WRT and SSIT showed high to excellent reliability; (ii) the TME for $\dot{V}O_{2MAX\_PRED}$ was 2.8 mL·kg-1·min-1 (6.6%) for TCSS and 2.9 mL·kg-1·min-1 (7.2%) for SSIT; (iii) a statistical difference was found between the measured and predicted $\dot{V}O_{2MAX}$ in both protocols; (iv) moderate (SSIT) to strong (WRT) correlation was found with $\dot{V}O_{2MAX}$; and (v) the Bland-Altman analysis showed high values of BIAS and limits of agreement, suggesting an underestimation of $\dot{V}O_{2MAX}$. Although the $\dot{V}O_{2MAX\_PRED}$ values are reliable and easy to apply, both underestimate $\dot{V}O_{2MAX}$ and have high limits of agreement, making their use unfeasible.

Reliability

The $\dot{V}O_{2MAX\_PRED}$ value showed no significant difference between days ($p > 0.05$) for both protocols (Table 1). Other studies also found no difference in this variable when two to three test days were adopted (Bennett et al., ,2016; Eng et al., 2004; Shushan et al., 2023). The ICC in both tests presented high to excellent reliability, as well as the confidence intervals are within the desired range for reliability. The

TME values of $\dot{V}O_{2MAX\_PRED}$ for WRT and SSIT were 2.8 and 2.9 $mL \cdot kg^{-1} \cdot min^{-1}$, respectively. The results are in agreement with Santos et al. (2012), who found a TME = 2.4 $mL \cdot kg^{-1} \cdot min^{-1}$when estimated by target HR by age. In assessing (direct) oxygen consumption in different tests, Eng et al. (2004) observed lower values (0.7 to 1.6 $mL \cdot kg^{-1} \cdot min^{-1}$) in different ergometers (cycle ergometer and treadmill), which can be explained by the direct measurement of $\dot{V}O_2$ during the tests (Eng et al., 2004). Bennett et al., (2016) have found in their review that the reliability of $\dot{V}O_{2MAX}$ prediction is good (mean -0.8 ± 3.7 $mL \cdot kg^{-1} \cdot min^{-1}$). No study found reliability by SSIT, making it difficult to compare the values found.

No significant differences were found between the days in analyzing the HR5-6min, however a trend ($p = 0.07$) can be observed for the HR5-6min when the WRT was adopted, probably due to the higher speed on the second day. The ICC values for this variable ranged from moderate (0.60 for WRT) to high (0.83 for SSIT) reliability. Lamberts et al. (2004) observed in an intermittent submaximal test an HR ICC between 0.96 and 0.99. This difference between the ICC and TME between the tests may be associated with the fact that the HR_{RES} was prefixed to 75% in the WRT and in the SSIT there was no pre-established limit.

The velocity variable in the WRT test was the only one that showed a significant difference between days. On the other hand, the ICC values for SSIT showed moderate to excellent reliability when compared to the WRT values, which ranged from low to excellent. Also, the TME was between 0.5 and 1 $km \cdot h^{-1}$ for the SSIT and WRT, respectively. In this same context, Wang et al. (2010) adopted a ten-minute SSIT, and observed that the speed (8.6 ± 2.2 $km \cdot h^{-1}$) presented excellent reliability values (ICC = 0.98; ETM = 0.4 $km \cdot h^{-1}$). To some extent, such divergences can be explained by how the intensity of the exercise was controlled. There was the possibility of establishing the exercise mode (walking or running) in both protocols in the present study, however the percentage in the WRT was imposed, while this intensity was perceived in the SSIT. Olher et al., (2019) have used a rating of perceived exertion to estimate $v\dot{V}O_{2MAX}$ (through heart rate cost); however, the authors did not test the reliability of the estimation. The good levels of reliability of $\dot{V}O_{2MAX\_PRED}$ in both protocols enable the use of the test to evaluate overall physical fitness in different trials over time (considering the typical error). However, it must be pointed out, that $\dot{V}O_{2MAX\_PRED}$ cannot be used interchangeable with $\dot{V}O_{2MAX}$.

Validity

Although the intensities and stabilization time of this study (WRT and SSIT) were similar to what Swain et al. (2004) highlighted as the best configuration (6-min stages and intensity close to 75% of HR_{RES}), an underestimation of $\dot{V}O_{2MAX}$ of 6.3 $mL \cdot kg^{-1} \cdot min^{-1}$ (13%) for the WRT and 8.9 $mL \cdot kg^{-1} \cdot min^{-1}$ (18%) for SSIT was observed in the present study. Such results differ from those of the study by Swain et al. (2004), who did not observe a difference for $\dot{V}O_{2MAX}$, and showed lower standard error of estimative (4.0 $mL \cdot kg^{-1} \cdot min^{-1}$). Therefore, including an increase in slope before running, or performing a self-selected intensity (as proposed in the present study) underestimates $\dot{V}O_{2MAX}$. Therefore, the $\dot{V}O_{2MAX\_PRED}$ for WRT and SSIT, should be used with caution. Recently, Dugas et al. (2023) found that ACSM equation and the submaximal test using HR overestimated the $\dot{V}O_{2MAX}$. However, the latter provided a more accurate prediction of $\dot{V}O_{2MAX}$ (Dugas et al., 2023). Compared with the present study, Dugas et al. (2023) found a lower mean difference for $\dot{V}O_{2MAX}$ estimation using HR at submaximal speed (3.8 $mL \cdot kg^{-1} \cdot min^{-1}$) and higher difference when using maximal speed to estimate $\dot{V}O_{2MAX}$ (9.8 $mL \cdot kg^{-1} \cdot min^{-1}$). Therefore, it seems that using HR response might be more accurate to estimate running $\dot{V}O_{2MAX}$.

Despite the statistical difference, both tests correlated with the standard $\dot{V}O_{2MAX}$. Therefore, it is possible to predict the $\dot{V}O_{2MAX}$ from the equation presented in Figure 1A, taking into account the EPE of 5.3 $mL \cdot kg^{-1} \cdot min^{-1}$. Adopting a similar model, Santos et al. (2012) estimated performance in 3.6 km uphill, 10 and 21 km events from the $\dot{V}O2_{RES}$ and found a high association (r = -0.83; -0.95 and -0.96; EPE = 5.0, 1.7 and 3.9 $mL \cdot kg^{-1} \cdot min^{-1}$, respectively); this high correlation may be associated with the fact that the test was performed on the track and not on the treadmill (Santos et al., 2012). As the SSIT has a lower correlation and consequently a higher EPE of 6.6 $mL \cdot kg^{-1} \cdot min^{-1}$, its use should be adopted with caution, and further studies should evaluate a greater number of subjects of both sexes. It is important to note that, regardless of the test, the residual value is high, and perhaps the $\dot{V}O_{2MAX\_PRED}$ might be used to evaluate physical fitness instead of comparing with $\dot{V}O_{2MAX}$.

The subjects presented high values of BIAS and limits of agreement in SSIT, while the WRT showed the smallest difference between the methods. Mainardi (2016) validated the test proposed by Santos et al.

(2012), and found similar linear regression values (r = 0.72; R = 0.52; EPE 5.3 $mL\cdot kg^{-1}\cdot min^{-1}$) to the present study; however, a better agreement (BIAS = 0.58 ± 5.63 $mL\cdot kg^{-1}\cdot min^{-1}$). It is possible that the difference between the protocols influenced the higher agreement values in the present study. In a practical way, the WRT protocol showed better responses both in terms of reliability and validity, and it can be used at different levels of conditioning, as long as it adopts the same settings. However, some limitations were observed throughout the study: 1) a reduced number of subjects evaluated; 2) the overall results were analyzed without distinction of results by sex.

Conclusions

Both submaximal tests examined in the present study demonstrated acceptable reliability, albeit with limitations regarding $\dot{V}O_{2MAX}$ estimation. If the primary purpose of applying such protocols is to determine $\dot{V}O_{2MAX}$ for interindividual comparisons or normative referencing, it is recommended to employ the corresponding prediction equations (Figure 2) to reduce estimation error. Nonetheless, both protocols appear suitable for tracking $\dot{V}O_{2MAX}$ changes over time.

Financing

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

References

Aune D, Schlesinger S, Hamer M, Norat T, Riboli E. (2020). Physical activity and the risk of sudden cardiac death: a systematic review and meta-analysis of prospective studies. BMC Cardiovasc Disord, 20, 318.

ACSM. (2024) ACSM's Guidelines for Exercise Testing and Prescription, 11ª ed.; Wolters Kluwer/Lippincott Williams & Wilkins Health: Philadelphia.

Tauda, M., Cruzat Bravo, E., & Suárez Rojas, F. (2025). Entrenamiento interválico de alta intensidad versus entrenamiento continuo moderado en la rehabilitación cardiaca: Revisión Sistemática y Meta-análisis. *Retos*, *63*, 433–458. https://doi.org/10.47197/retos.v63.110286

Tauda, M. (2024). Análisis de la correlación entre consumo máximo de oxígeno la potencia de salto y parámetros fisiológicos en jugadores de baloncesto (Analysis of the correlation between maximum oxygen consumption, jumping power and physiological parameters in basketball players). *Retos*, *59*, 864–880. https://doi.org/10.47197/retos.v59.107603

Väisänen D, Ekblom B, Wallin P, Anderssen G, Ekblom-Bak E. (2024). Reference values for estimated VO2max by two submaximal cycle tests: the Åstrand-test and the Ekblom-Bak test. Eur J Appl Physiol, 124, 1747–1756.

Yang HI, Cho W, Lee DH, Suh S-H, Jeon JY. (2021) Development of a New Submaximal Walk Test to Predict Maximal Oxygen Consumption in Healthy Adults. Sensors, 21 (17) :5726.

Sartor F, Vernillo G, Morree HM, Bonomi AG, La Torre A, Kubis HP, et al. (2013). Estimation of maximal oxygen uptake via submaximal exercise testing in sports, clinical, and home settings. Sports Med, 43, 865-873.

Arena RJ, Myers MA, Williams M, Gulati P, Kligfield GJ, Balady E, et al. (2007). Assessment of functional capacity in clinical and research settings: a scientific statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention of the Council on Clinical Cardiology and the Council on Cardiovascular Nursing. Circulation,116, 329-343.

Guazzi M, Bandera F, Ozemek C, Systrom D, Arena R. (2017). Cardiopulmonary Exercise Testing: What Is its Value? J Am Coll Cardiol, 70, 1618-1636.

Dugas M-O, Paradis-Deschênes P, Simard L, Chevrette T, Blackburn P, Lavallière M. (2023). Comparison of VO2max Estimations for Maximal and Submaximal Exercise Tests in Apparently Healthy Adults. Sports, 11 (12), 235.

Evans HJ, Ferrar KE, Smith AE, Parfitt G, Eston RG. (2015). A systematic review of methods to predict maximal oxygen uptake from submaximal, open circuit spirometry in healthy adults. J Sci Med Sport, 18, 183-188.

Åstrand PO, Ryhming I. (1954). A Nomogram for Calculation of Aerobic Capacity (Physical Fitness) From Pulse Rate During Submaximal Work. J Appl Physiol,7, 218-221.

Swain DP, Wrigth R L. (1997). Prediction of VO2peak from submaximal cycle ergometry using 50 versus 80 rpm. Med Sci Sports Exerc, 29, 268-272.

Swain DP, Parrott JA, Bennett AR, Branch JD, Dowling EA. (2004). Validation of a new method for estimating VO2max based on VO2 reserve. Med Sci Sports Exerc 36, 1421-1426.

Santos TM, Rodrigues AI, Greco CC, Marques AL, Terra BS, Oliveira BR. (2012). Rev Bras Cineantropom Desempenho Hum, 14, 192-201.

Mainardi, FC. (2016). Validade de critério concorrente e preditiva do VO2MAX estimado em esteira pela reserva do consumo de oxigênio e frequência cardíaca. [Master Dissertation], Universidade Federal do Rio de Janeiro.

Beltrame T, Villar R, Hughson RL. (2017). Sex differences in the oxygen delivery, extraction, and uptake during moderate-walking exercise transition. Appl Physiol Nutr Metab, 42, 994–1000.

Wiecha S, Kasiak PS, Cieśliński I, Takken T, Palka T, Knechtle B, Nikolaidis PT, Malek LA, Postula M, Mamcarz A, Sliz D. (2023). External validation of VO2max prediction models based on recreational and elite endurance athletes. PLoS ONE, 18, 1: e0280897.

Ferri Marini C, Micheli L, Grossi T, Federici A, Piccoli G, Zoffoli L, Correale L, Dell'Anna S, Naldini CA, Lucertini F, Vandoni M. (2024). Are incremental exercise relationships between rating of perceived exertion and oxygen uptake or heart rate reserve valid during steady-state exercises? PeerJ, 12, e17158

Olher RR, Sales MM, Sousa CV, Sotero RC, Madrid B, Cunha RR, Moraes MR, Simões HG. (2019). Heart rate cost of running in track estimates velocity associated with maximal oxygen uptake. Physiol Behav, 205, 33-38.

Coquart J, Tabben M, Farooq A, Tourny C, Eston R. (2016). Submaximal, Perceptually Regulated Exercise Testing Predicts Maximal Oxygen Uptake: A Meta-Analysis Study. Sports Med, 46, 885-897.

Shea MG, Headley S, Mullin EM, Brawner CA, Schilling P, Pack QR. (2022). Comparison of Ratings of Perceived Exertion and Target Heart Rate-Based Exercise Prescription in Cardiac Rehabilitation: A Randomized Controlled Pilot Study. J Cardiopulm Rehabil Prev, 1, 42 (5), 352-358

Ekkekakis P. (2009). Let Them Roam Free? Physiological and psychological evidence for the potential of self-selected exercise intensity in public health. Sports Med, 39, 857–888.

Rodriguez-Fuentes, G., Campo-Prieto, P., Souto, X. C., & Cancela Carral, J. M. (2024). Realidad virtual inmersiva y su influencia en parámetros fisiológicos de personas sanas (Immersive virtual reality and its influence on physiological parameters in healthy people). *Retos*, *51*, 615–625. https://doi.org/10.47197/retos.v51.101164

Koo TK, Li MY. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 15, 155-163.

Bland JM, Altman DG. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. Lancet, 327, 307-310.

Shushan T, Lovell R, Buchheit M. Scott TJ, Barrett S, Norris D, McLaren SJ. (2023). Submaximal Fitness Test in Team Sports: A Systematic Review and Meta-Analysis of Exercise Heart Rate Measurement Properties. Sports Med - Open, 9, 21.

Eng JJ, Dawson AS, Chu KS. (2004). Submaximal exercise in persons with stroke: test-retest reliability and concurrent validity with maximal oxygen consumption. Arch Phys Med Rehabil, 85, 113-118.

Bennett H, Parfitt G, Davison K, Eston R. (2016). Validity of submaximal step tests to estimate maximal oxygen uptake in healthy adults. Sports Med, 46, 737-750.

Lamberts RP, Lemmink KA, Durandt JJ, LambertMI. (2004). Variation in heart rate during submaximal exercise: implications for monitoring training. J Strength Cond Res 18, 641-645.

Wang SC, Lin HF, Wu CF, Lin BN, Wang YS, Huang Y.J. (2010). Aerobic power assessment by using a 10 min heart rate control running on treadmill. J Sports Med Phys Fitness 50, 32-36.

Authors' and translators' details:

Thiago Barbosa Lima	thiago_lima19@hotmail.com	Author
Willemax dos Santos Gomes	willemaxsantos@gmail.com	Author
Vinícius de Oliveira Damasceno	viniciusvod@fab.mil.br	Author
Tony Meireles Santos	tony.meireles@ufpe.br	Author and Translator
Eduardo Zapaterra Campos	eduardo.zapaterracampos@ufpe.br	Author and Translator