

RESEARCH ARTICLE

A 45-Second Self-Test for Cardiorespiratory Fitness: Heart Rate-Based Estimation in Healthy Individuals

Francesco Sartor^{1*}, Matteo Bonato², Gabriele Papini^{1,3,4}, Andrea Bosio⁵, Rahil A. Mohammed^{6,7}, Alberto G. Bonomi¹, Jonathan P. Moore⁶, Giampiero Merati², Antonio La Torre², Hans-Peter Kubis⁶

1 Personal Health, Philips Research, Eindhoven, The Netherlands, **2** Department of Biomedical Sciences for Health, Università degli Studi di Milano, Milan, Italy, **3** Department of Information Engineering, University of Pisa, Pisa, Italy, **4** Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands, **5** Mapei Sport, Olgiate Olona (Varese), Italy, **6** College of Health and Behavioural Sciences, Bangor University, Bangor, United Kingdom, **7** School of Physical Education, University of Sulaimani, Sulaimani, Iraq

* francesco.sartor@philips.com



OPEN ACCESS

Citation: Sartor F, Bonato M, Papini G, Bosio A, Mohammed RA, Bonomi AG, et al. (2016) A 45-Second Self-Test for Cardiorespiratory Fitness: Heart Rate-Based Estimation in Healthy Individuals. *PLoS ONE* 11(12): e0168154. doi:10.1371/journal.pone.0168154

Editor: Yoshihiro Fukumoto, Kurume University School of Medicine, JAPAN

Received: June 13, 2016

Accepted: November 26, 2016

Published: December 13, 2016

Copyright: © 2016 Sartor et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All files are available from the 45 sec CRF test Study database (accession number <http://dx.doi.org/10.5061/dryad.vh610>).

Funding: Dr. Alberto G. Bonomi, Mr. Gabriele Papini and Francesco Sartor work for Philips Research; Dr. Andrea Bosio works for Mapei Sport. The funders provided support in the form of salaries for authors A.G.B., G.P., F.S. and A.B., but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of

Abstract

Cardio-respiratory fitness (CRF) is a widespread essential indicator in Sports Science as well as in Sports Medicine. This study aimed to develop and validate a prediction model for CRF based on a 45 second self-test, which can be conducted anywhere. Criterion validity, test re-test study was set up to accomplish our objectives. Data from 81 healthy volunteers (age: 29 ± 8 years, BMI: 24.0 ± 2.9), 18 of whom females, were used to validate this test against gold standard. Nineteen volunteers repeated this test twice in order to evaluate its repeatability. CRF estimation models were developed using heart rate (HR) features extracted from the resting, exercise, and the recovery phase. The most predictive HR feature was the intercept of the linear equation fitting the HR values during the recovery phase normalized for the height² ($r^2 = 0.30$). The Ruffier-Dickson Index (RDI), which was originally developed for this squat test, showed a negative significant correlation with CRF ($r = -0.40$), but explained only 15% of the variability in CRF. A multivariate model based on RDI and sex, age and height increased the explained variability up to 53% with a cross validation (CV) error of $0.532 \text{ L} \cdot \text{min}^{-1}$ and substantial repeatability (ICC = 0.91). The best predictive multivariate model made use of the linear intercept of HR at the beginning of the recovery normalized for height² and age²; this had an adjusted $r^2 = 0.59$, a CV error of $0.495 \text{ L} \cdot \text{min}^{-1}$ and substantial repeatability (ICC = 0.93). It also had a higher agreement in classifying CRF levels ($\kappa = 0.42$) than RDI-based model ($\kappa = 0.29$). In conclusion, this simple 45 s self-test can be used to estimate and classify CRF in healthy individuals with moderate accuracy and large repeatability when HR recovery features are included.

these authors are articulated in the 'author contributions' section. All other authors have no conflict of interest to disclose.

Competing Interests: The authors have read the journal's policy and the authors of this manuscript have the following competing interests: F.S. and A.G.B. work for Philips Research; G.P. was doing an internship at Philips Research during his contribution in this work. A.B. works for Mapei Sport. All other authors have no conflict of interest. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

Introduction

Maximal oxygen uptake ($\dot{V}_{O_{2max}}$) estimation is essential in sports, clinical and home settings [1] because it is considered the best index of cardio-respiratory fitness (CRF) [2], which expresses the maximal aerobic power of an individual [3]. The importance of $\dot{V}_{O_{2max}}$ as predictor of cardiovascular disease and mortality has been recently highlighted by a scientific statement issued by the American Heart Association [4]. Remarkably, individuals with low cardiorespiratory fitness had a 70% higher risk of all cause-mortality and a 56% risk of cardiovascular disease mortality [4]. The inclusion of cardio-respiratory fitness in the traditional cardiovascular disease risk score is currently being evaluated [5]. Nevertheless, the direct assessment of $\dot{V}_{O_{2max}}$ implies a higher risk of adverse cardiovascular events, costly equipment for the measurement of the expired gases, and trained personnel. Moreover, lack of motivation, general discomfort, reduced exercise tolerance, and the occurrence of symptoms can hamper its correct assessment [2]. Maximal exercise criteria are often unmet in non-athlete populations [6, 7]. For these reasons a large number of submaximal protocols have been designed to estimate $\dot{V}_{O_{2max}}$. We have recently reviewed submaximal exercise protocols employing different modalities such as cycling, walking, running, treadmill, stepping and many others, which can be performed in laboratories as well as in health/fitness centres or at school [1]. However, most of them require some sort of equipment, a large space, and a 3 to 6 minute exercise protocol. Only a few can be executed anywhere and in a short period of time. We have identified the 45 s squat test, also known as the Ruffier-Dickson test, as a suitable protocol to provide an easy and fast self-evaluation of $\dot{V}_{O_{2max}}$.

The Ruffier-Dickson test was originally designed by J.E. Ruffier and modified by J. Dickson, who developed the Ruffier-Dickson index (RDI), and it has been used to classify cardio-respiratory fitness [8]. However, the only data available to compare $\dot{V}_{O_{2max}}$ with RDI was published in Rugby players [9]. Using those rugby players data we have shown a moderately valid model for estimating $\dot{V}_{O_{2max}}$ in these players (see Sartor et al. [1]). Nevertheless, this model was developed in a specific population of athletes with a reasonably high $\dot{V}_{O_{2max}}$, it was not cross validated, and it was based on a small sample size of 22 players [1, 9]. Furthermore, to our knowledge there is no evidence to judge how well the CRF levels are classified by the RDI, although this was one of the main reasons this index was designed for.

The aim of the present study was to validate as well as cross-validate the 45-second squat test in a larger population with a wider range of $\dot{V}_{O_{2max}}$. The secondary objective was to evaluate the correlation between the original Ruffier-Dickson index (RDI) and $\dot{V}_{O_{2max}}$, and its CRF classification value. Moreover, the repeatability of the model developed was evaluated via a test re-test design.

Materials and Methods

Participants

This study has been conducted in four different centres: Bangor University in the United Kingdom, Philips Research in The Netherlands, and Università degli Studi di Milano (Milan, Italy), and Mapei-Sport in Italy. The volunteers were recruited via posters and newsletters. Volunteers were excluded if they suffered from any chronic disease; if they had functional or cognitive impediments; or if they were taking any medication affecting the hormonal and metabolic systems. In total 94 healthy volunteers were included in the study data from 13 subjects were excluded because of incompleteness or technical difficulties (e.g. HR data were too noisy for various reasons and could not be improved by low-pass filtering). The characteristics of the

Table 1. Subjects characteristics.

	Males	Females	All
	Means ± SDs (Min ÷ Max)	Means ± SDs (Min ÷ Max)	Means ± SDs
	n = 63	n = 18	n = 81
Age (years)	29 ± 8 (18 ÷ 67)	27 ± 9 (19 ÷ 46)	29 ± 8
Height (m)	1.76 ± 0.06 (1.66 ÷ 1.88)	1.70 ± 0.05 (1.61 ÷ 1.76)	1.74 ± 0.06
Weight (kg)	74.5 ± 9.8 (58 ÷ 100)	68.0 ± 10.7 (52.5 ÷ 86.95)	73.1 ± 10.3
BMI	24.1 ± 2.9 (18.8 ÷ 33.6)	23.5 ± 3.0 (19.0 ÷ 29.22)	24.0 ± 2.9
HR _{rest} (beats·min ⁻¹)	66 ± 10 (46 ÷ 93)	70 ± 11 (47 ÷ 96)	67 ± 11
$\dot{V}_{O_{2max}}$ (L·min ⁻¹)	3.15 ± 0.72 (1.96 ÷ 5.22)	2.27 ± 0.47 (1.35 ÷ 3.18)	2.95 ± 0.77
$\dot{V}_{O_{2max}}$ (mL·kg ⁻¹ ·min ⁻¹)	43.4 ± 11.7 (27.0 ÷ 77.8)	33.8 ± 6.9 (23.7 ÷ 43.4)	41.3 ± 11.5
HR _{max} (beats·min ⁻¹)	187 ± 10 (150 ÷ 209)	184 ± 13 (157 ÷ 202)	186 ± 11

doi:10.1371/journal.pone.0168154.t001

remaining 81 subjects are reported in Table 1. The study protocol was reviewed and approved by the Departmental Ethics Committees of Bangor’s School of Sport Health and Exercise Sciences, the Internal Committee of Biomedical Experiments of Philips Research and Institutional Ethics Review Committee of the Università degli Studi di Milano in conformity with the Declaration of Helsinki. All subjects have signed the written informed consent prior to testing.

Submaximal and maximal cardiorespiratory fitness assessment

Bare-footed subjects’ stature was measured using a wall-mounted stadiometer (Bodycare Products, Southam, United Kingdom, used in Bangor, Seca 217, Vogel & Halke, Hamburg, Germany, used in Eindhoven, Varese, and Milan). A digital scale (Seca; Vogel & Halke, Hamburg, Germany, used in Bangor and Milan, BC-533, Tanita, Amsterdam, the Netherlands, used in Eindhoven, RB-L 200, and Wunder, Trezzo sull’Adda (MI), Italy, used in Varese) was used to measure body weight of subjects in their underwear.

Subjects lied down (supine) for 5 minutes in order to rest. At the end of the 5 minutes rest subjects stood up. The researcher monitored the HR until the HR stabilized in the standing position. Once the HR reached steady state in the standing position (this usually required only a few seconds), subjects started squatting. Subjects performed 30 squats in 45 s, following the tempo set by a metronome (80 beats·min⁻¹), executing a down and up movement. The squatting movement consisted of flexion of the knees to 90°, keeping the back straight and the arms extended frontally. Complete squatting was avoided in order to make this test feasible to a large range of people. At the end of the 45 seconds the subjects lied down (supine) and recovered for 3 minutes. Heart rate was recorded in inter-beat-interval mode throughout the entire test including the resting and recovery phases by means of a chest strap HR monitor (Polar, RS800CX; Polar Electro, Kempele, Finland). Nineteen non-athlete subjects repeated the squat test on a different day at least one week after the first test.

A few minutes after having completed the Ruffier-Dickson squat test when HR was ≤ initial resting HR (and in any case not earlier than 10 min) the subjects moved to the stationary cycle ergometer (Lode Excalibur Sport PFM, Groningen, The Netherlands, used in Bangor, Eindhoven and Varese; and a Monark Ergomedic 839 E, Monark, Varberg, Sweden, used in Milan) and were prepared for the incremental test to exhaustion to determine their maximum oxygen uptake. Prior to the test, the subjects warmed-up for 10 min on the cycle ergometer with no resistance at a pedalling rate between 90 and 100 rpm. The incremental exercise began at a workload of 75 W for 3 min to increase by 25 W every minute until volitional exhaustion, which was defined as the inability to maintain a cadence of 90 rpm for more

than 5 s despite strong verbal encouragement. During the test, \dot{V}_{O_2} , \dot{V}_{CO_2} , ventilation and respiratory rate were continuously monitored breath by breath (MetaLyzer 3B, Leipzig, Germany, used in Bangor; Oxycon Pro Metabolic Cart, Carefusion, California, USA, used in Eindhoven; Vmax29, SensorMedics, Yorba Linda, CA, USA, used in Varese; and Cosmed Quark CPET, Cosmed, Roma, Italy, used in Milan.). HR was measured continuously using the same heart rate monitor worn during the Ruffier-Dickson squatting test.

In addition to the 94 healthy volunteers tested to validate the squat test 10 more subjects (age = 24 ± 5 , BMI = 23.5 ± 4.1 , 60% males) were recruited in order to evaluate the aerobic component of the 45 seconds squat test. These subjects were asked to perform the same test while wearing a breath by breath indirect calorimeter (k4b2, Cosmed, Albano Laziale, Italy), calibrated before each test according to the maker's requirements.

Multiple regression models

All data collected in this study were processed by using Matlab software (R2013b, Matworks). The raw beat-to-beat HR intervals, in milliseconds, were converted into HR, in $\text{beats} \cdot \text{min}^{-1}$, which was re-sampled to 1 Hz. Peaks were removed according with the following logic operation. For each HR sample the short term variation, defined as the standard deviation in an interval of 3 samples, was compared to the long term variation, defined as the standard deviation in an interval of 21 samples. When the short term standard deviation was at least 10 times greater than the long term standard deviation, the HR sample was replaced with the average of the previous and following HR samples. Furthermore, the signal was low-pass filtered using a 5 samples moving average. Each signal was visually inspected to ensure that this cleaning procedure did not alter the morphology of the signal. A total of 25 features were extracted from the HR signal during the resting, squatting and recovery phase. These HR features together with subjects' characteristics (e.g. age, height and weight, etc) were used to build the prediction models. The following HR features were considered. Rest HR mode was defined as the mode HR value during 5 min of resting in the supine position, excluding the first and the last minute of measurement. Squat peak was the maximum HR observed during the squatting exercise. The recovery HR linear intercept (named for simplicity here, start HR_{rec}) was the intercept (b) of the linear fit $y = m x + b$ (where m was the angular coefficient of the linear fit) of the HR signal corresponding to the complete 180 s recovery phase. Recovery exponential coefficient was the coefficient (d) of the exponential fit ($y = d \cdot e^{fx}$) of the HR decay during the recovery phase. Recovery HR end was the final value of the complete recovery phase. Finally RDI was calculated as previously described by Piquet et al [9] that is,

$$\text{RDI} = \frac{(P_1 - 70) + 2(P_2 - P_0)}{10}$$

where P_0 is 15 s average resting HR, in our case, from 3 min 45 s to 4 min, P_1 is the maximum HR recorded during the first 15 s of recovery, and P_2 is the 15 s average after the 1st minute of recovery, that is from 1 min and 00 s and 1 min and 15 s.

Statistical analysis

The statistical analysis was conducted in Matlab. The prediction models were built by using stepwise forward multiple linear regressions. In order to assess the repeatability of the $\dot{V}_{O_{2\text{max}}}$ prediction using squat test HR derived features and models, 19 subjects performed a test re-test protocol from which we have calculated the Pearson correlation coefficient, and the intra-

class coefficient (ICC), by using the following equation

$$ICC = \frac{MSr - MSe}{MSr}$$

where MSr is the mean square of the scores in test 1 and test 2, and MSe is the mean square error. The leave one out method was used to cross-validate the models. This means that the dataset was divided randomly into a number of groups equal to the number of subjects. Any iteration data among each subject was removed from the dataset and the model was tested against the remaining observations. Moreover, a simulation for evaluating the stability of the model features was performed. This was done by calculating the grand average of the root mean squared error (RMSE) of each model, developed in this study, as subject's number increased (Fig 1). The RMSE was calculated from a random "testing" subset (10% of the full dataset) using a model trained on an increasing number of subjects (1 to 73, that is, full dataset—size of testing data set) belonging to a fictitious "training" dataset. Subjects were randomly included in this "training" dataset to remove any possible order effect. This procedure was repeated in 100 iterations (arbitrary number) per number of subjects included in the "training" dataset and the grand average RMSE was computed.

In order to assess the quality of the RDI and our newly developed models in classifying CRF levels against the American College of Sports Medicine (ACSM) norms based on the directly measured \dot{V}_{O2max} [10], sensitivity and specificity were calculated as follows. According with the original RDI classification of CRF levels [8], we made 3 categories: RDIs ≤ 5 were considered good CRF, RDIs between 6 and 10 were considered fair CRF, and RDIs ≥ 11 were

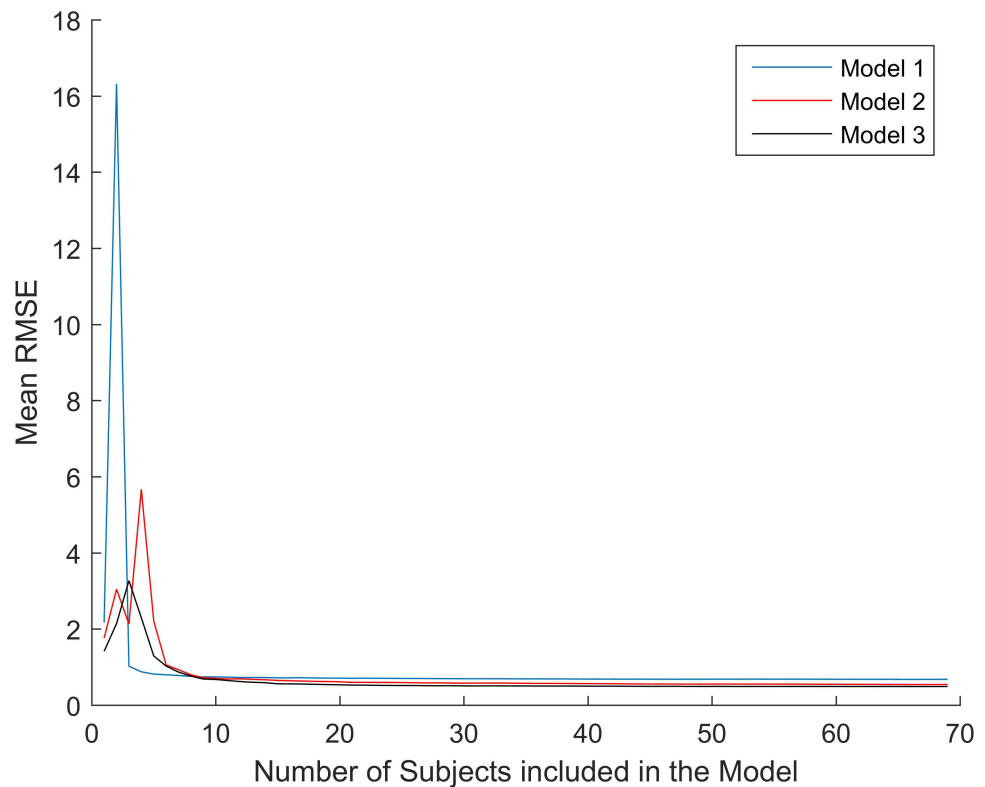


Fig 1. Features stability. These were assessed by calculating the grand average of the root mean squared error (RMSE) of each model 1, 2, and 3 as subject's number included in the evaluation increased.

doi:10.1371/journal.pone.0168154.g001

classified as poor CRF. Consistently, we have simplified the ACSM classification in the same 3 categories, pooling together very poor and poor into poor, and good excellent and superior into good. Sensitivity was defined as:

$$\text{Sensitivity} = \frac{nTP}{nTP + nFN}$$

where nTP was the number of true positive, which in our case was estimation models CRF level equal to ACSM CRF level for good and fair; and nFN was number of false negative, which was when the estimation models underestimated ACSM CRF levels. Specificity was calculated as:

$$\text{Specificity} = \frac{nTN}{nTN + nFP}$$

where nTN was the number of true negative, that is, when the models and the ACSM classifications agreed on the poor CRF levels; and nFP was the number of false positive, namely, when the models overestimated ACSM CRF levels. The false positive rate was computed as 1 – Specificity. Moreover, the agreement between the ACSM classification and the models classification was evaluated with Cohen’s kappa coefficient (κ).

Results

In Table 2 we report the linear relations between subjects’ characteristics, the HR features at rest, during squatting and recovery and the absolute $\dot{V}_{O_{2max}}$. Three multiple linear regression models to predict absolute $\dot{V}_{O_{2max}}$ were developed. Model 1 and 2 were based on the RDI, whereas models 3 made use of HR recovery features developed in this study (Fig 2). The accuracy and cross-validation errors of these models are reported in Table 3.

Nineteen individuals had repeated the squat test a second time on a different day at least one week after the first test. Intra-class correlation coefficient for Squat HR peak was 0.86 ($r' = 0.76, p < 0.001$), ICC for Start HR rec was 0.79 ($r' = 0.66, p < 0.01$) and the ICC for RDI was 0.82 ($r' = 0.70, p < 0.001$). Intraclass correlation coefficient for model 1 was the same as for

Table 2. Pearson correlation coefficients for absolute $\dot{V}_{O_{2max}}$ ($L \cdot \text{min}^{-1}$).

	Males n = 63	Females n = 18	All n = 81
Sex	-	-	0.480 (p < 0.001)
Age	-0.463 (p < 0.001)	-0.326	-0.325 (p = 0.003)
Height	0.331 (p = 0.008)	0.552 (p = 0.017)	0.483 (p < 0.001)
Weight	-0.046	0.350	0.143
Rest HR mode	-0.169	-0.400	-0.262 (p = 0.018)
HR max (during $\dot{V}_{O_{2max}}$ test)	0.521 (p < 0.001)	0.300	0.459 (p < 0.001)
Squat HR peak	-0.256 (p = 0.043)	0.096	-0.279 (p = 0.011)
Start HR _{rec}	-0.381 (p = 0.002)	-0.123	-0.374 (p < 0.001)
Recovery exponential coefficient	-0.319 (p = 0.011)	0.094	-0.291 (p = 0.008)
Recovery HR end	-0.286 (p = 0.023)	-0.531 (p = 0.023)	-0.339 (p = 0.002)
RDI	-0.322 (p = 0.010)	-0.554 (p = 0.017)	-0.402 (p < 0.001)
Start HR _{rec} / Age ²	0.517 (p < 0.001)	0.289	0.238 (p = 0.032)
Start HR _{rec} / H ²	-0.505 (p < 0.001)	-0.354	-0.549 (p < 0.001)

Start HR_{rec} = Recovery HR linear intercept, H = Height

doi:10.1371/journal.pone.0168154.t002

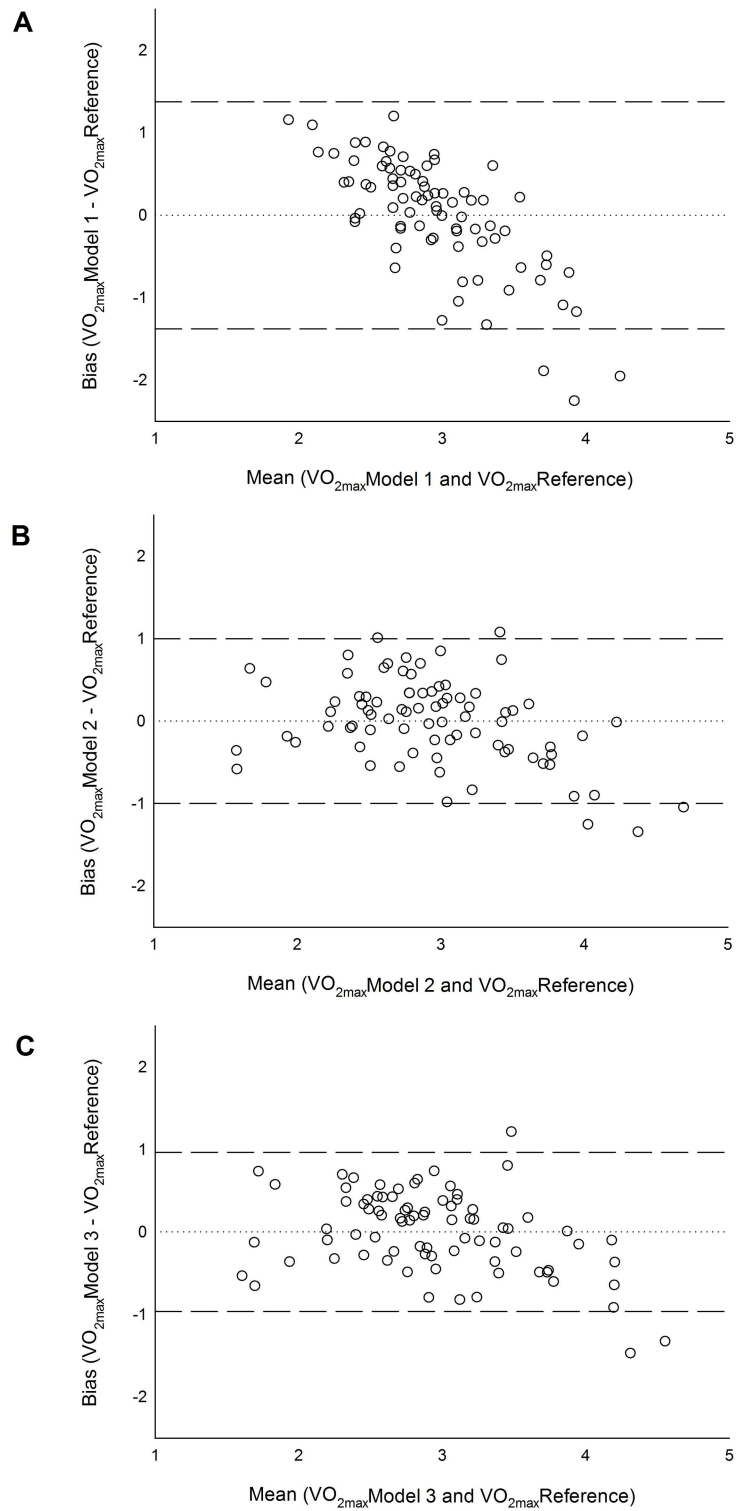


Fig 2. Bland and Altman plots. A, Model 1 based on RDI only; B, Model 2 based on RDI and subject's characteristics; C, Model 3 based on sex and Start HR_{rec} normalized for height and Age; n = 81.

doi:10.1371/journal.pone.0168154.g002

Table 3. Multiple linear regression models to predict \dot{V}_{O2max} (n = 81).

	Coef.	SE	t	p level	r	Adj.r ²	RMSE (L·min ⁻¹)	Bias (L·min ⁻¹)	LoA (L·min ⁻¹)	LOOCV (L·min ⁻¹)
Model 1					0.402	0.151	0.707	-0.001	1.375	0.676
									-1.377	
Constant	3.867	0.247	15.650	<0.001						
RDI	-0.110	0.0281	-3.905	<0.001						
Model 2					0.747	0.535	0.523	0.001	0.999	0.532
									-0.998	
Constant	-3.788	1.798	-2.107	0.038						
Sex	0.560	0.159	3.534	<0.001						
Age	-0.0309	0.007	-4.393	<0.001						
Height	4.533	1.051	4.314	<0.001						
RDI	-0.0864	0.021	-4.014	<0.001						
Model 3					0.766	0.587	0.502	0.001	0.967	0.495
									-0.965	
Constant	4.121	0.458	9.007	<0.001						
Sex	0.787	0.147	5.357	<0.001						
Start HR _{rec} /Age ²	4.235	0.708	5.979	<0.001						
Start HR _{rec} /H ²	-0.0673	0.011	-6279	<0.001						

SE = Standard error, RMSE = root mean squared error, LoA = limits of agreement, LOOCV = leave one out cross validation error, RDI = Ruffier-Dickson index, H = height

doi:10.1371/journal.pone.0168154.t003

RDI, being this model based solely on RDI, ICC for model 2 was 0.91 ($r' = 0.83$, $p < 0.001$), finally ICC for model 3 was 0.93 ($r' = 0.88$, $p < 0.001$).

As described in the methods we have made 3 fitness classes based on the ACSM norms, poor, average and good. Out of our 81 subjects 40 belonged to the poor fitness category, 24 to the average fitness and 17 to the good fitness. Model 1 has shown a sensitivity to classify CRF levels according with the ACSM norms equal to 61%, whereas its specificity was 49%, meaning that the rate of false positive (e.g. subjects wrongly classified as fit) was 51%. There was a fair agreement between model 1 and ACSM CRF levels ($\kappa = 0.293$). Model 1 misclassified 7 individuals over 81 (9%) by two categorises.

Model 2 showed a better sensitivity 62% and specificity 63%, reducing the miss classification of poor to average into fit to 37%. Model 2 also had a fair agreement with the ACSM classification of CRF ($\kappa = 0.397$). Model 3, similarly to model 2, had 64% sensitivity and 62% specificity, with a moderate agreement with the classification of CRF as according with the ACSM norms ($\kappa = 0.417$). Model 2 and 3 improved on two categories misclassification, dropping this down to 1 over 81 observations (1%).

In order to assess the aerobic requirements of the squatting exercise breath by breath O_2 uptake measurements were performed during the squatting protocol on 10 subjects. This revealed a mean peak metabolic equivalent (MET) of 6.27 ± 2.23 . One MET was defined as an O_2 uptake of $3.5 \text{ ml} \cdot \text{kg}^{-1}$. According to these results this test can be classified as aerobically vigorous exercise in the less fit and moderate exercise in fitter individuals.

Discussion

\dot{V}_{O2max} can be estimated by executing a 45 s squat test in healthy individuals of both sexes. The RDI, which uses both resting and recovery HR information, significantly correlated with

absolute $\dot{V}_{O_{2max}}$. However, RDI was not the best HR metric to estimate CRF or $\dot{V}_{O_{2max}}$. The best HR feature in the current study was the intercept of the linear fit of the first 180 s of recovery, (called here start HR_{rec}).

The oxygen uptake measured while performing 30 squats in 45 seconds has verified the moderate to vigorous impact (~6 peak MET) of this test on the aerobic system and its associated cardiorespiratory load (128 beats · min⁻¹ peak HR). We expected individuals' height to play a key role in influencing the hemodynamic responses induced by repetitive squatting. This is because the displacement of the centre of mass is larger in taller individuals, corresponding to a greater exercise dose which elicits a greater cardiac work [11]. In fact once we accounted for height differences, dividing start HR_{rec} by height squared, the differences in HR were mainly dependent on CRF levels and less on anthropometric factors. Indeed, we observed a higher correlation with absolute $\dot{V}_{O_{2max}}$ when using start HR_{rec}/H². Similarly, we have corrected the start HR_{rec} for age, which is known to negatively relate with $\dot{V}_{O_{2max}}$ [12]; also this correction increases the correlation with CRF. Interestingly, body weight did not show an important correlation with absolute $\dot{V}_{O_{2max}}$. This is not surprising considering that body weight alone does not necessarily reflect fat free mass [13] and that $\dot{V}_{O_{2max}}$ measured on the cycle ergometer does not scale with body weight [14]. In a previous investigation we have developed a prediction equation based on squatting data gathered in 22 male rugby players by Piquet et al [9], using body weight, age and, end of squatting exercise HR as prediction variables; on that limited dataset a high correlation was observed ($r = 0.75$) [1]. However, when that equation was tested in this current dataset no correlation with $\dot{V}_{O_{2max}}$ ($r = -0.0089$) was found. This confirmed that body weight is not a good predictor of fitness for this type of submaximal test and that end exercise HR by itself is not a good predictor of CRF.

The Ruffier-Dickson index takes into account resting as well as recovery HR parameters. In particular it takes into account start recovery HR, which is very close to peak values, and the difference between 1 minute recovery HR and resting HR. Table 2 shows how these HR parameters well correlate with CRF. Once the RDI was entered in a linear regression model (Model 1) it was observed that although the mean bias was very close to zero, this model had the tendency to overestimate CRF in low fitness level people and conversely underestimate CRF in fit people (Fig 2A). This was statistically confirmed by a significant negative heteroscedasticity. This drawback was mitigated by including in the regression model information about sex, age and height (Model 2). Model 2, RDI plus individual's characteristics, also improved on accuracy dropping the cross-validation estimation error from 23% down to 18%. If no advance HR features extraction tools were available, Model 2 should be use to estimate CRF when using the 45 s squat protocol described in this study. This is because individual's characteristics improve in personalization and therefore in prediction and reduce the risk of fitness bias.

However, the best multiple linear regression model was based on HR recovery curve linear fit normalized as described above. This third model had only three predictive variables, sex start HR_{rec}/H², start HR_{rec}/Age². Model 3 showed the lowest cross validation error 16.8%, the best repeatability (ICC = 0.93), and the best CRF classification performance. It is important to notice that, although a moderate agreement ($\kappa = 0.417$) between gold standard and Model 3 in classifying CRF in 3 categories, poor, average and good, may seem low, the κ coefficient, in fact, accounts for random agreements. Indeed, a κ equal to 0 would be an agreement purely by chance [15]. Moreover, it also relevant to report that Model 3 as well as Model 2 had a very low two categories misclassification rate (1%). This analysis, which is rarely done when validating CRF tests, was very important when considering that the RDI was originally conceived to classify people according to their fitness level [8]. The RDI alone was observed to have poor

specificity, meaning that the rate of people placed in a higher fitness level category was very high (51%). Thus, the RDI should not be used alone for CRF classifications. We have also evaluated how the sample size could affect the accuracy of our models. In Fig 1, it is evident that the feature selected show stable and rather low errors when more than 10 random individuals are included in the models. This simulation together with the leave one out cross-validation gives us confidence that the accuracy shown in this dataset will be reproducible in other data. The error shown by model 3 (16.8%) is on par with well-known submaximal protocols such as the Rockport walk test on the treadmill (15%) [16], or the ACSM cycling test (15.5%) [17]. Moreover, model 3 proved to be as repeatable as most submaximal protocols (ICC = 0.93, and $r' = 0.88$), e.g. the Queens' College step test ($r' = 0.92$) [18], or the ACSM cycling test ($r' = 0.86$) [17].

This study has strengths and limitations. The main strength is that it provides an easy-to-use test (e.g. Model 2) to evaluate CRF in healthy people with a wide range of fitness levels. The main limitation of the current work is that only healthy people not taking heart rhythm medications were included. It is logical to assume that HR-based CRF estimations would not be appropriate in heart patients taking such medications. A lower HR value would be read as higher fitness by our prediction models instead of higher medication dose. In order to address this issue, we have elsewhere disclosed and validated a novel approach, which makes use of the same squatting protocol, but based on body movements [19]. Moreover, a 45 second protocol was used to replicate the original Ruffier-Dickson test. However, this may not have been the optimal duration for assessing CRF by means of squatting. Future, research should investigate the optimal duration for this squatting protocol.

In conclusion, a simple 45 s squat test can be used to estimate CRF or $\dot{V}_{O_{2max}}$ in healthy males and females with moderate accuracy and large repeatability when HR recovery features of the first 180 s are used. Furthermore, our newly developed model (model 3) can classify CRF levels reasonably well. Although the RDI alone significantly correlates with the absolute $\dot{V}_{O_{2max}}$ and its repeatability was acceptable, it seems to overestimate $\dot{V}_{O_{2max}}$ for individuals with low $\dot{V}_{O_{2max}}$ and underestimate it in individuals with high $\dot{V}_{O_{2max}}$. RDI performance was improved when individual's characteristics, such as sex, age, height were added to the model.

Acknowledgments

The authors would like to thank Prof. A. Veicsteinas for giving us access to the facilities of the Department of Biomedical Sciences for Health at Università degli Studi di Milano, and Mr. C. Bottrill for contributing to the data collection. F.S. and A.G.B work for Philips Research, G.P was doing an internship at Philips Research during his contribution in this work. A.B works for Mapei Sport. All other authors have no conflict of interest. No funding was received to conduct this study.

Author Contributions

Conceptualization: FS HPK.

Data curation: FS.

Formal analysis: FS MB GP AGB.

Investigation: MB RAM AB FS JPM GM ALT.

Methodology: FS HPK.

Project administration: FS HPK ALT.

Resources: HPK JPM ALT AB FS.

Software: FS GP AGB.

Supervision: FS HPK JPM GM ALT.

Validation: FS.

Visualization: FS GP.

Writing – original draft: FS.

Writing – review & editing: FS HPK JPM MB AB AGB.

References

1. Sartor F, Vernillo G, de Morree HM, Bonomi AG, La Torre A, Kubis HP, et al. Estimation of maximal oxygen uptake via submaximal exercise testing in sports, clinical, and home settings. *Sports Medicine*. 2013; 43(9):865–73. Epub 2013/07/04. doi: [10.1007/s40279-013-0068-3](https://doi.org/10.1007/s40279-013-0068-3) PMID: [23821468](https://pubmed.ncbi.nlm.nih.gov/23821468/)
2. Fletcher GF, Ades PA, Kligfield P, Arena R, Balady GJ, Bittner VA, et al. Exercise standards for testing and training: a scientific statement from the American Heart Association. *Circulation*. 2013; 128(8):873–934. Epub 2013/07/24. doi: [10.1161/CIR.0b013e31829b5b44](https://doi.org/10.1161/CIR.0b013e31829b5b44) PMID: [23877260](https://pubmed.ncbi.nlm.nih.gov/23877260/)
3. Bassett DR Jr., Howley ET. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Medicine and Science in Sports and Exercise*. 2000; 32(1):70–84. Epub 2000/01/27. PMID: [10647532](https://pubmed.ncbi.nlm.nih.gov/10647532/)
4. Kaminsky LA, Arena R, Beckie TM, Brubaker PH, Church TS, Forman DE, et al. The importance of cardiorespiratory fitness in the United States: the need for a national registry: a policy statement from the American Heart Association. *Circulation*. 2013; 127(5):652–62. Epub 2013/01/09. doi: [10.1161/CIR.0b013e31827ee100](https://doi.org/10.1161/CIR.0b013e31827ee100) PMID: [23295916](https://pubmed.ncbi.nlm.nih.gov/23295916/)
5. Goff DC Jr., Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB Sr., Gibbons R, et al. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2013; 127(1):e1–50. Epub 2013/11/14.
6. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake: a brief critique and recommendations for future research. *Sports Medicine*. 2007; 37(12):1019–28. Epub 2007/11/22. PMID: [18027991](https://pubmed.ncbi.nlm.nih.gov/18027991/)
7. Poole DC, Wilkerson DP, Jones AM. Validity of criteria for establishing maximal O₂ uptake during ramp exercise tests. *European Journal of Applied Physiology*. 2008; 102(4):403–10. Epub 2007/10/31. doi: [10.1007/s00421-007-0596-3](https://doi.org/10.1007/s00421-007-0596-3) PMID: [17968581](https://pubmed.ncbi.nlm.nih.gov/17968581/)
8. Jousselein E. Le test de Ruffier, improprement appelé test de Ruffier-Dickson. *Medicins du sport*. 2007; 83(4 January 2014):33–4.
9. Piquet L, Dalmay F, Ayoub J, Vandroux JC, Menier R, Antonini MT, et al. Study of blood flow parameters measured in femoral artery after exercise: correlation with maximum oxygen uptake. *Ultrasound in Medicine & Biology*. 2000; 26(6):1001–7. Epub 2000/09/21.
10. ACSM. ACSM's Guidelines for Exercise Testing and Prescription. 8th ed. Baltimore: Lippincott Williams & Wilkins; 2009.
11. Blomqvist CG, Lewis SF, Taylor WF, Graham RM. Similarity of the hemodynamic responses to static and dynamic exercise of small muscle groups. *Circulation Research*. 1981; 48(6 Pt 2):187–92. Epub 1981/06/01. PMID: [7226467](https://pubmed.ncbi.nlm.nih.gov/7226467/)
12. Astrand I. Aerobic work capacity in men and women with special reference to age. *Acta Physiologica Scandinavica*. 1960; 49(169):1–92. Epub 1960/01/01.
13. Goran M, Fields DA, Hunter GR, Herd SL, Weinsier RL. Total body fat does not influence maximal aerobic capacity. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity*. 2000; 24(7):841–8. Epub 2000/08/05.
14. Nevill AM, Ramsbottom R, Williams C. Scaling physiological measurements for individuals of different body size. *European journal of applied physiology and occupational physiology*. 1992; 65(2):110–7. Epub 1992/01/01. PMID: [1396632](https://pubmed.ncbi.nlm.nih.gov/1396632/)
15. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Family medicine*. 2005; 37(5):360–3. Epub 2005/05/11. PMID: [15883903](https://pubmed.ncbi.nlm.nih.gov/15883903/)
16. Pober DM, Freedson PS, Kline GM, McInnis KJ, Rippe JM. Development and validation of a one-mile treadmill walk test to predict peak oxygen uptake in healthy adults ages 40 to 79 years. *Canadian Journal of Applied Physiology*. 2002; 27(6):575–89. Epub 2002/12/26. PMID: [12500996](https://pubmed.ncbi.nlm.nih.gov/12500996/)

17. Greiwe JS, Kaminsky LA, Whaley MH, Dwyer GB. Evaluation of the ACSM submaximal ergometer test for estimating VO₂max. *Medicine and Science in Sports and Exercise*. 1995; 27(9):1315–20. Epub 1995/09/01. PMID: [8531631](#)
18. McArdle WD, Katch FI, Pechar GS, Jacobson L, Ruck S. Reliability and interrelationships between maximal oxygen intake, physical work capacity and step-test scores in college women. *Medicine and Science in Sports*. 1972; 4(4):182–6. Epub 1972/01/01. PMID: [4648576](#)
19. Papini G, Bonomi AG, Stut W, Kraal JJ, Kemps H, Sartor F. A 45-second self-test for cardiorespiratory fitness: accelerometry-based estimation in heart patients Forthcoming.