Validity and reliability of seismocardiography for the estimation of cardiorespiratory fitness

Mikkel T. Hansen, MSc,*[1] Tue Rømer, MSc,*[1] Amalie Højgaard, BM,*
Karina Husted, PhD,* Kasper Sørensen, PhD,†‡ Samuel E. Schmidt, PhD, AP,†‡
Flemming Dela, MD,*§ Jørn W. Helge, PhD*

*From the *Xlab, Center for Healthy Aging, Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, †Department of Health Science and Technology, Aalborg University, Aalborg, Denmark, ‡VentriJect ApS, Hellerup, Denmark, and §Department of Geriatrics, Bispebjerg University Hospital, Copenhagen, Denmark.*

BACKGROUND Low cardiorespiratory fitness (ie, peak oxygen consumption [$\dot{V}O_2$peak]) is associated with cardiovascular disease and all-cause mortality and is recognized as an important clinical tool in the assessment of patients. Cardiopulmonary exercise test (CPET) is the gold standard procedure for determination of $\dot{V}O_2$peak but has methodological challenges as it is time-consuming and requires specialized equipment and trained professionals. Seismofit is a chest-mounted medical device for estimating $\dot{V}O_2$peak at rest using seismocardiography.

OBJECTIVE The purpose of this study was to investigate the validity and reliability of Seismofit $\dot{V}O_2$peak estimation in a healthy population.

METHODS On 3 separate days, 20 participants (10 women) underwent estimations of $\dot{V}O_2$peak with Seismofit ($\times$2) and Polar Fitness Test (PFT) in randomized order and performed a graded CPET on a cycle ergometer with continuous pulmonary gas exchange measurements.

RESULTS Seismofit $\dot{V}O_2$peak showed a significant bias of –3.1 $\pm$ 2.4 mL·min^{-1}·kg^{-1} (mean $\pm$ 95% confidence interval) and 95% limits of agreement (LoA) of $\pm$10.8 mL·min^{-1}·kg^{-1} compared to CPET. The mean absolute percentage error (MAPE) was 12.0%. Seismofit $\dot{V}O_2$peak had a coefficient of variation of 4.5% $\pm$ 1.3% and an intraclass correlation coefficient of 0.95 between test days and a bias of 0.0 $\pm$ 0.4 mL·min^{-1}·kg^{-1} with 95% LoA of $\pm$1.6 mL·min^{-1}·kg^{-1} in test–retest. In addition, Seismofit showed a 2.4 mL·min^{-1}·kg^{-1} smaller difference in 95% LoA than PFT compared to CPET.

CONCLUSION The Seismofit is highly reliable in its estimation of $\dot{V}O_2$peak. However, based on the measurement error and MAPE >10%, the Seismofit $\dot{V}O_2$peak estimation model needs further improvement to be considered for use in clinical settings.

KEYWORDS Seismocardiography; VO_2max estimation; Nonexercise VO_2peak equation; Cardiorespiratory fitness test; Method agreement

(Cardiovascular Digital Health Journal 2023;4:155–163)

Introduction

Cardiorespiratory fitness is a powerful independent predictor of cardiovascular disease (CVD) and all-cause mortality.[1–4] Measurement of peak oxygen consumption ($\dot{V}O_2$peak) during a graded cardiopulmonary exercise test (CPET) is considered the gold standard method for the determination of cardiorespiratory fitness.[1,5] Routine assessment of $\dot{V}O_2$peak in the clinical setting is recommended by the American Heart Association (AHA)[1] because it significantly improves CVD risk classification compared to traditional CVD risk models.[6,7] Furthermore, even a small increment in $\dot{V}O_2$peak is found to be associated with reduced mortality risk[8,9] and reduced risk of future CVD readmission in cardiac patients.[10] The AHA recommendation has not yet been successfully adopted, as it might not be possible or feasible to perform exercise testing during most patient encounters.[11] Nonexercise estimation of $\dot{V}O_2$peak has been proposed as a pragmatic alternative to CPET.[1,11–14] Due to recent advantages in heart rate monitoring technology that allow estimation of $\dot{V}O_2$peak to be performed in the resting condition, Polar Electro Oy (Kempele, Finland) created the Polar Fitness Test (PFT), which uses age, body weight, height, and self-reported weekly training hours together with resting heart rate and heart rate variability to estimate

The study was conducted at Xlab, Center for Healthy Aging, Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [1]Denotes shared first authorship. ClinicalTrials.gov Identifier: NCT05356871. **Address reprint requests and correspondence:** Mikkel Thunestvedt Hansen, Xlab, Department of Biomedical Sciences, Office 07-2-56, Blegdamsvej 3B, DK-2200 Copenhagen, Denmark. E-mail address: mikkel.hansen@sund.ku.dk.

2666-6936/ https://doi.org/10.1016/j.cvdhj.2023.08.020

KEY FINDINGS

- The reliability of the Seismofit peak oxygen consumption ($\dot{V}O_2$peak) estimation was high both within and between test days.
- The Seismofit underestimated $\dot{V}O_2$peak compared to the gold standard measurement and the mean absolute percentage error was above the predefined level of 10% for clinical relevance.
- The reliability indicates that the Seismofit is dependable in recording and processing the seismocardiography signal. Thus, it is expected that Seismofit $\dot{V}O_2$peak estimation will improve when more data are available for training the algorithm.

$\dot{V}O_2$peak. However, the validity of the PFT is still considered inadequate for application in clinical health care and in sports.[15–17] Seismocardiography (SCG) is a technique for noninvasive evaluation of cardiac activity through the measurement of precordial vibrations using an accelerometer.[18–20] The complexity of SCG signals has previously been a challenge; however, the new advantages of low-cost lightweight sensors, signal processing, feature extraction, and machine learning methods have made the SCG methodology relevant for clinical application.[20] It is widely accepted that maximal cardiac output (ie, maximal heart rate and stroke volume) is the primary factor limiting $\dot{V}O_2$peak in the healthy population,[21] and this is highly dependent on diastolic function.[22] A faster relaxation and increase in diastolic filling would increase cardiac output and thus $\dot{V}O_2$peak (ie, the Frank-Starling law). Fiducial points in the resting SCG signal have previously been shown to correlate with characteristic events in the cardiac cycle[23] and provide information on ventricular performance.[24] In addition, a high correlation between the diastolic SCG peak-to-peak value (fiducial point C_d to D_d) and the diastolic relaxation parameter e´ measured by echocardiography in the resting condition has recently been established.[25] Functionally, a faster relaxation would lead to a more rapid drop in left ventricular pressure, thereby creating a greater pressure difference between the left ventricle and ascending aorta, which would result in a larger amplitude of the C_d to D_d fiducial point.[26] A $\dot{V}O_2$peak estimation model using both diastolic and systolic SCG information at rest has been proposed[26] and validated against CPET.[27,28] This led to the development of a new wireless sensor device containing an accelerometer with a compatible smartphone app and a cloud-based solution for automated SCG signal processing (Seismofit system).[29] The Seismofit is an interesting novel solution and potentially is a clinically applicable alternative to the standard CPET. However, thorough assessment of the validity and reliability of the Seismofit estimation of $\dot{V}O_2$peak compared with the CPET is needed to evaluate the applicability of the method. This study aimed to examine the validity and reliability of the Seismofit $\dot{V}O_2$peak estimation compared with CPET in a healthy population sample. It was hypothesized that the Seismofit would prove valid and reliable for estimation of $\dot{V}O_2$peak, with a mean absolute percentage error (MAPE) $\leq$10% between methods and a within-method coefficient of variation (CoV) $\leq$5%.

Methods

Study outline and objectives

Three identical test days separated by at least 48 hours and within 14 days were conducted to assess the validity and reliability of the Seismofit $\dot{V}O_2$peak estimation. On each test day, $\dot{V}O_2$peak was estimated twice with Seismofit (Seismofit 1 and 2) and once with PFT before being measured directly with the gold standard CPET using indirect calorimetry. CPET was used as a reference method to assess the agreement of Seismofit and PFT and to compare the reliability against the different estimation methods. The repeated Seismofit measurement within each test day was used to assess repeatability. The present study only compares estimated and measured $\dot{V}O_2$peak values performed in a clinical setting and is not addressing the algorithms behind the estimation models. Testing was conducted between June 2020 and November 2020. The study conformed to the Helsinki Declaration and was approved by the Science Ethical Committee of the Greater Region of Copenhagen, Denmark (H-19081375). The study was registered as ClinicalTrials.gov Identifier NCT05356871.

Study population

Twenty healthy participants were recruited (10 women, 10 men). A priori recruitment strategy including 4 trained and 4 untrained between the ages of 18 to 39 years, and 6 trained and 6 untrained between the ages of 40 to 75 years with equal distribution between sexes was completed. A detailed description of the recruitment strategy is provided in the Supplementary Materials (Supplemental Table 1). Inclusion criteria were healthy adults between the ages of 18 and 75 years. Exclusion criteria were a history of previous or current CVD, pregnancy, and/or conditions that prevent maximal exercise testing. Before volunteering and signing a written consent form, the participants received oral and written information about the study and its possible risks.

Experimental design

Measurements of anthropometrics and blood pressure together with collection of a resting blood sample were obtained on each test day, and body composition was determined on the first test day (Table 1). A detailed description of the measurements, experimental procedures, and equipment used is given in the Supplemental Material.

Estimation of $\dot{V}O_2$peak in resting condition

The Seismofit system was used for recording of the SCG signal and estimation of $\dot{V}O_2$peak at supine rest.[29] The Seismofit system consists of a small medical device containing a

Table 1 Measured and estimated values from three separate test days

	Day 1	Day 2	Day 3	Within-subject SD	Within-subject CoV%	ICC	P-value
Measurements							
Age (yrs.)	47 ± 7	-	-	-	-	-	-
Height (cm)	176 ± 4	-	-	-	-	-	-
Weight (kg)	73.7 ± 5.1	73.7 ± 5.1	73.8 ± 5.4	0.5	0.6 ± 0.1	1.00	0.837
Body fat, (%)	26.1 ± 3.3	-	-	-	-	-	-
Systolic BP (mmHg)	132 ± 9	128 ± 7	128 ± 8	5	3 ± 1	0.91	0.063
Diastolic BP (mmHg)	80 ± 5	77 ± 5*	76 ± 4*	4	5 ± 1	0.85	**0.009**
HRrest (beats·min^{-1})	55 ± 4	54 ± 4	54 ± 4	3	5 ± 2	0.85	0.280
Haemoglobin (mmol/L)	9.2 ± 0.4[a]	8.9 ± 0.4[b] *	8.9 ± 0.4[b] *	0.2	2.3 ± 0.7	0.94	**0.031**
Haematocrit (%)	45.5 ± 2.0[a]	44.0 ± 2.0[b] *	44.2 ± 2.1[b] *	1.0	2.2 ± 0.6	0.95	**0.028**
Non-exercise $\dot{V}O_2$peak estimation							
Seismofit 1 (ml·min^{-1}·kg^{-1})	36.6 ± 3.4	36.6 ± 3.1	36.6 ± 3.4	1.7	4.5 ± 1.3	0.95	0.907
Seismofit 2 (ml·min^{-1}·kg^{-1})	36.6 ± 3.3	36.7 ± 3.1	36.5 ± 3.5	1.5	3.8 ± 1.2	0.96	0.872
Polar Fitness Test (ml·min^{-1}·kg^{-1})	43.5 ± 4.8	44.9 ± 4.8	43.9 ± 4.6	1.8	3.0 ± 1.4	0.97	0.168
Cardiopulmonary exercise test							
$\dot{V}O_2$peak (ml·min^{-1}·kg^{-1})	39.6 ± 2.7	39.7 ± 3.1	39.2 ± 3.2[a]	1.5	3.4 ± 0.8	0.95	0.907
$\dot{V}O_2$peak (ml·min^{-1})	2916 ± 293	2930 ± 315	2837 ± 289[a]	116	3 ± 1	0.97	0.808
$V_E$peak (L·min^{-1})	124 ± 14	127 ± 15	124 ± 14[a]	8	5 ± 2	0.94	0.246
HRpeak (beats·min^{-1})	173 ± 7	172 ± 6	172 ± 7[a]	3	2 ± 0	0.96	0.150
Watt peak (W)	251 ± 28	252 ± 30	242 ± 29[a]	8	2 ± 1	0.99	0.896

Data are presented as mean ± 95% CI. n=20.
[a]n=18.
[b]n=17 $\dot{V}O_2$, oxygen consumption; V_E, minute ventilation; HR, heart rate; Seismofit, $\dot{V}O_2$peak estimation model using seismocardiography version 4.3; Polar Fitness Test, $\dot{V}O_2$peak estimation using the Polar Fitness Test. SD; standard deviation. CoV; coefficient of variation. ICC; interclass correlation coefficient. The Watt peak is calculated using the last completed Watt stage + (Seconds / 60 seconds * watt increments in the representative protocol). The p-value represent the fixed effect; *test day* from a mixed-effect model fitted using Restricted Maximum Likelihood (REML) for assessment of systematically differences.
*denotes significant different from day 1 (p<0.05).

3-axis digital output accelerometer, a smartphone app, and a cloud solution for signal processing. The SCG recording was performed with the Seismofit device mounted on the sternum with double adhesive tape 2 cm proximal to the xiphoid process (Supplemental Figure 1). Before estimation, the age, height, weight, and sex of the participant were entered into the app. $\dot{V}O_2$peak estimation lasted approximately 5 minutes, which included SCG recording, data transfer to the app, and signal processing. Subsequently, the result of the $\dot{V}O_2$peak estimation was shown in the app. The newest version, at the time of the study, of the $\dot{V}O_2$peak estimation model was used (SCG Version 4.3), which has previously been described.[27] The $\dot{V}O_2$peak estimation model is reproduced in Equation (1):

$\dot{V}O2$ peak estimation

$$\begin{aligned} &= (-15.108) + (-0.122 \cdot S2\,FrequencySpec) \\ &+ (0.371 \cdot S2\,Morphology) \\ &+ (0.001 \cdot RR) + (0.247 \cdot S1\,FrequencySpec) \\ &+ (0.701 \cdot FRIENDS\ study\ prediction\ model) \\ &+ (143.4 \cdot amp\_Dd) + (-159.45 \cdot amp\_Ks) \\ &+ (-0.042 \cdot SYSRV\_STD) \\ &+ (-87.583 \cdot amp\_Ls) + (42.993 \cdot amp\_Fs) \end{aligned} \tag{1}$$

S2FrequencySpec is analysis of the frequency spectrum of the average SCG diastolic complex quantified using principal component analysis. S2Morphology is the morphology of the average SCG diastolic complex quantified using principal component analysis. RR is the average duration of a heartbeat in milliseconds. S1FrequencySpec is a frequency measure of the average SCG systolic complex quantified using principal component analysis of the frequency spectrum of the systolic complex. The FRIENDS study prediction model is an algorithm based on sex, age, and body weight for prediction of $\dot{V}O_2$peak.[12] SYSRV_STD is the variation in respiratory rate estimated from heart rate variability. Amp_Dd, Amp_Ks, Amp_Ls, and Amp_Fs are the amplitude of Dd, Ks, Ls, and Fs fiducial points, respectively (Figure 1). The definition of the fiducial points has previously been reported.[23] Visualization of the filtered SCG signals from 3 different test days of 1 participant is shown in Figure 1.

The patented "Polar Fitness Test"[30] was performed at supine rest using a chest-strapped heart rate sensor (Polar H10) connected with a Polar Vantage M watch. The PFT uses demographic data (age, weight, height, sex, and training background) together with a recording of heart rate and heart rate variability.[30] The estimation lasted approximately 5 minutes after entry of the participant's demographic data in the watch settings. The categories of training background were 0–1, 1–3, 3–5, 5–8, 8–12, and 12+ hours per week.

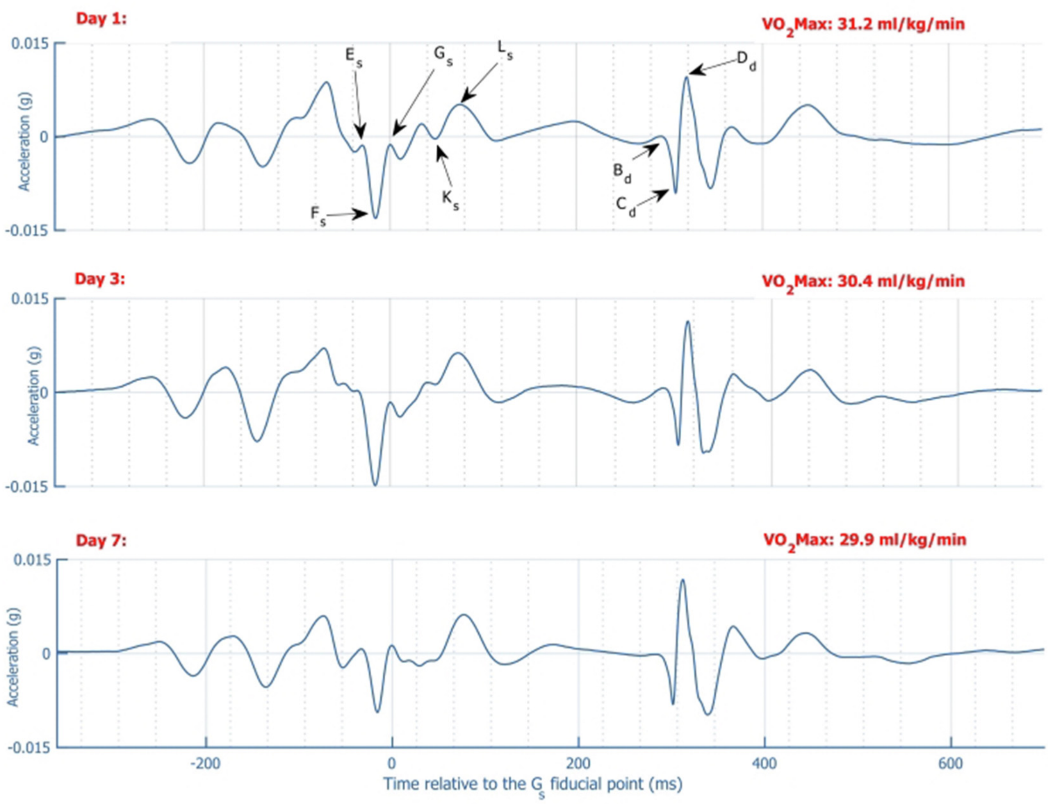

Figure 1 Filtered seismocardiography (SCG) signals from a participant on 3 separate test days together with SeismoFit peak oxygen consumption ($\dot{V}O_2$peak) estimation. *Arrows* indicate the different fiducial points used in the SeismoFit $\dot{V}O_2$peak estimation model with annotation of fiducial points following what is previously used in a normal SCG signal. B_d = aortic valve closure; C_d = aortic valve closure minimum point; D_d = aortic valve closure maximum point; E_s = mitral valve closure; F_s = aortic valve opening minimum point; G_s = aortic valve opening maximum point; K_s = systolic outflow minimum point; L_s = systolic outflow maximum point.

Determination of $\dot{V}O_2$peak with CPET

Graded CPET until voluntary exhaustion was performed on a cycle ergometer to verify whether the Seismofit $\dot{V}O_2$peak estimation could accurately determine $\dot{V}O_2$peak in healthy adults. The warm-up and initial test workload of the CPET protocol differed between sexes and training status and increased by 20 and 25 W every minute for women and men, respectively (Supplemental Table 2). The different protocols were applied to achieve a test duration of 8 to 12 minutes.[31] Pulmonary gas exchange was continuously obtained breath by breath during the CPET (Quark CPET, Cosmed, Rome, Italy) and sampled into 10-second intervals by an automatic online system (PFT Suite, Cosmed). Heart rate was continuously measured with a chest-strapped sensor (H10, Polar, Finland). Directly after test termination, the participants rated the perceived exhaustion on the Borg 6–20 scale.[32] $\dot{V}O_2$peak was calculated as the highest value measured over consecutive 30 seconds. The primary criterion for achieving $\dot{V}O_2$peak was a plateau in $\dot{V}O_2$, defined as <2.1 $mL \cdot min^{-1} \cdot kg^{-1}$ increment in $\dot{V}O_2$ with increasing workload.[33] The secondary criteria were respiratory exchange ratio >1.10, maximal heart rate within 5 bpm of the highest measured value on the 3 test days, and ≥ 18 on the Borg 6–20 scale.[34,35] $\dot{V}O_2$peak was accepted if the primary criterion or 2 of 3 of the secondary criteria were met. CPET data from the third test of 2 participants were excluded due to a calibration error. Data and the premises for exclusion are given in Supplemental Figure 2. A detailed description of the equipment calibration and evaluation of the $\dot{V}O_2$peak plateau is presented in the Supplemental Material.

Statistical analysis

All data were treated as normally distributed. Data are given as mean ± 95% confidence interval (CI). Agreement and prediction error of Seismofit 1 and PFT compared with CPET was determined by the following statistical analyses: Bland-Altman plot with 95% limits of agreement (BA-plot), standard error of estimate (SEE), MAPE, and CoV. Pearson correlation coefficient r was used to evaluate the relation between Seismofit 1 and CPET, and between PFT and CPET.

Bland-Altman plots show the difference between the intra-method means plotted on the y-axis and the CPET mean plotted on the x-axis when CPET is the reference method.[36] Otherwise, the average of the 2 intra-method means is plotted on the x-axis.[37]

SEE was calculated using $SEE = \sqrt{\sum \frac{(Y-Y')^2}{n-2}}$, with Y representing CPET $\dot{V}O_2$peak and Y' representing Seismofit or

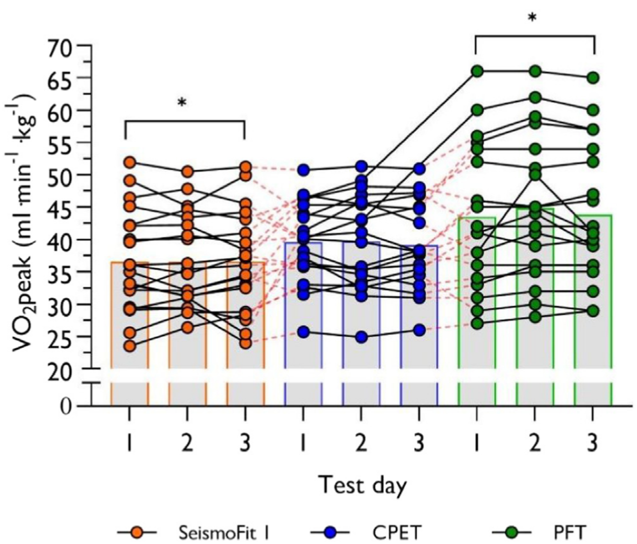

Figure 2 Individual values of peak oxygen consumption ($\dot{V}O_2$peak) on the 3 different test days were estimated using the seismocardiography model (Seismofit) and the Polar Fitness Test (PFT) and directly measured during a graded cardiopulmonary exercise test with pulmonary gas exchange measurements (CPET). Mixed-effect model was applied, with the method and test day as fixed effects. Seismofit vs CPET: method (p = .022), test day (P = .907), interaction of method and test day (P = .949). PFT vs CPET: method (P = .007), test day (P = .169), interaction of method and test day (P = .135). *Significantly different from CPET (P <.05).

PFT $\dot{V}O_2$peak values. Interpretation of SEE was very good <4.0 mL·min^{-1}·kg^{-1}; good 4.0–6.0 mL·min^{-1}·kg^{-1}; fair 6.0–8.0 mL·min^{-1}·kg^{-1}; poor 8.0–10.0 mL·min^{-1}·kg^{-1}; and very poor >10.0 mL·min^{-1}·kg^{-1}. The interpretation was based on previous studies using nonexercise estimations models.[12,13,38] MAPE ≤10% was used as a criterion level for clinical relevance. Interpretation of Pearson r was very high >0.90; high 0.70–0.90; moderate 0.50–0.70; low 0.30–0.50; and little, if any, 0.00–0.30.[39]

The reliability of the test methods was assessed by the interclass correlation coefficient (ICC), within-subject SD, and CoV. Interpretation of ICC was poor <0.50; moderate 0.50–0.75; good 0.75–0.90; and excellent >0.90.[40] within-subject SD was calculated by taking the square root of the within-subject variance. In addition, a mixed-effect model was applied to assess systematic differences between the test methods and test days for Seismofit 1 and CPET, and for PFT and CPET, respectively. Test-retest validity of Seismofit 1 and Seismofit 2 was assessed with the following analyses: BA-plot, Pearson's r and a mixed-effect model. The mixed model uses a compound symmetry covariance matrix, and is fitted using restricted maximum likelihood. This method gives the same P value and multiple comparison tests as repeated measures analysis of variance in the absence of missing values. In the presence of missing values (missing completely at random), the results can be interpreted as repeated measures analysis of variance. Statistical analyses were performed, and figures were constructed in GraphPad Prism 9.3.1 (GraphPad Software, Boston, MA) and Microsoft Excel (Microsoft Corporation, Redmond, WA).

Results

CPET

After the exclusion of 2 recordings with measurement errors, excellent ICC and no statistically significant difference in measures of pre-exercise body weight or peak values of $\dot{V}O_2$, minute ventilation, heart rate, and workload were observed between test days for CPET (Table 1). A difference was observed for pre-exercise hematocrit and hemoglobin (Table 1). CPET displayed a low within-subject SD and CoV (Table 1). Individual values from the CPET are shown in Supplemental Figure 3.

Validity of $\dot{V}O_2$peak estimations

Seismofit 1 estimated $\dot{V}O_2$peak was, on average, 7% lower compared to CPET $\dot{V}O_2$peak (Figure 2). The agreement between methods showed a relatively large negative bias together with substantial 95% limits of agreement (LoA) (Figure 3C). A fair SEE (Figure 3A) together with a moderate CoV and MAPE above the predefined clinical limit for relevance were found between Seismofit 1 and CPET (Table 2). A high Pearson r was found between methods (Table 2).

In comparison, PFT estimated $\dot{V}O_2$peak was 11% higher than CPET $\dot{V}O_2$peak (Figure 2). The agreement between PFT and CPET showed a relatively large positive bias together with large 95% LoA (Figure 3D). A poor SEE (Figure 3B), a moderate CoV, and a MAPE well above the predefined clinical limit for relevance were found between PFT and CPET (Table 2). A high Pearson r was found between PFT and CPET.

Reliability of $\dot{V}O_2$peak estimations

Between test days, both Seismofit 1 and PFT demonstrated a low within-subject SD and CoV, together with an excellent ICC (Table 1).

Agreement analysis of the Seismofit test–retest revealed a nonsignificant bias of –0.0 ± 0.4 mL·min^{-1}·kg^{-1} (P = .993) with an acceptable 95% LoA (Figure 4B). A very high Pearson r of 0.99 was found between Seismofit 1 and Seismofit 2 (Figure 4A).

Discussion

The present study provides novel insight into the validity and reliability of a nonexercise estimation of $\dot{V}O_2$peak using SCG performed with a chest-mounted device—Seismofit—in a healthy population. The main finding was that even though Seismofit underestimated VO_2max significantly, Seismofit had a high correlation with the gold standard CPET and proved highly reliable in the estimation of $\dot{V}O_2$peak both within test days and between test days. In addition, the Seismofit displayed an overall better agreement with CPET than PFT.

Validity of Seismofit $\dot{V}O_2$peak estimation

When clinical variables are measured, some degree of error is always present. Therefore, when comparing methods, when neither method provides an unequivocally correct

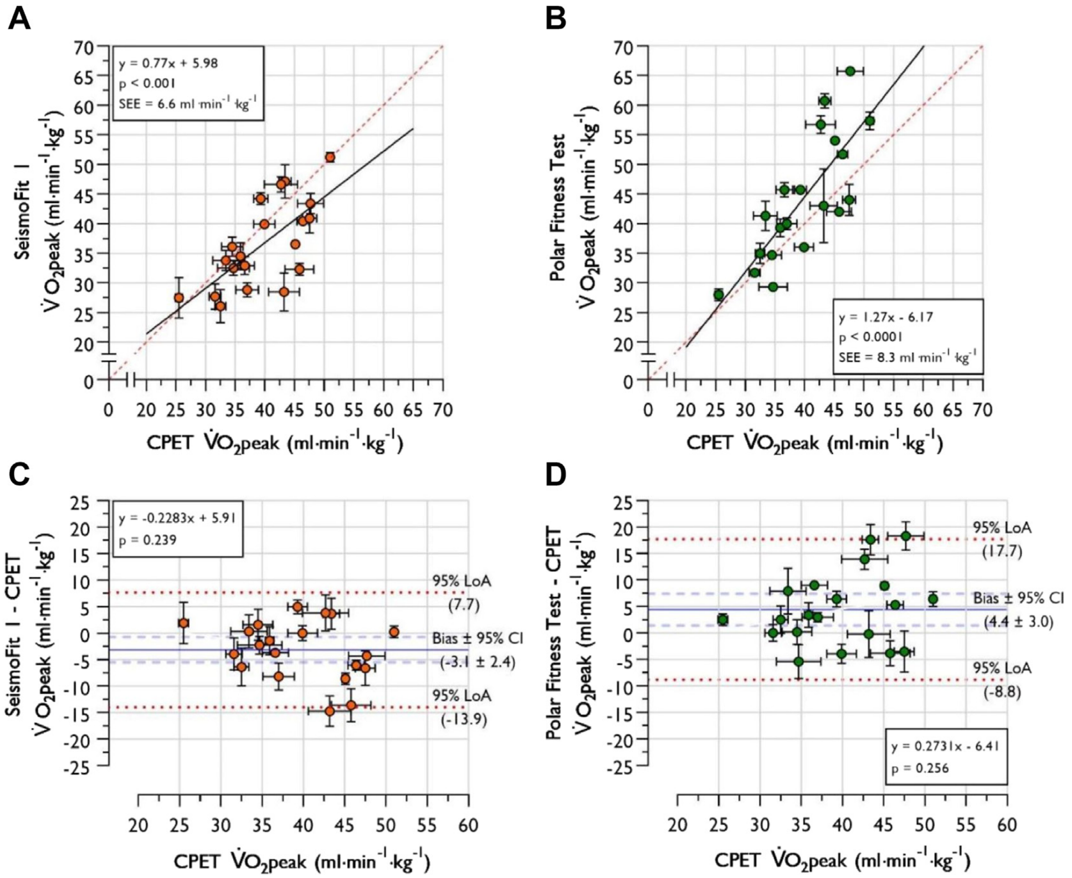

Figure 3 Data ($n = 20$) from of 3 separate tests for each participant (mean ± 95% confidence interval [CI]). **A:** Scatter plot with linear regression between $\dot{V}O_2$peak estimated with a nonexercise model using seismocardiography (SeismoFit 1) and directly measured during a graded cardiopulmonary exercise test with pulmonary gas exchange measurements (CPET). **B:** Scatter plot with linear regression between $\dot{V}O_2$peak estimated with the nonexercise Polar Fitness Test (PFT) and CPET. For **A** and **B,** *red dotted line* represents the line of identity ($r = 1.0$). **C:** Bland-Altman plot of the agreement between SeismoFit 1 and CPET. **D:** Bland-Altman plot of the agreement between PFT and CPET. *Blue line* represents the bias. *Blue dotted line* represents the 95% CI of the bias. *Red dotted line* represents the 95% LoA. Note: Two participants only have data from 2 separate tests included in the analysis. LoA = limits of agreement; SEE = standard error of estimate.

measurement, the preference is to assess the degree of agreement between methods.[41] Furthermore, evaluation of whether the methods agree sufficiently should be decided depending on the context in which the measurement will be used.[42] To the best of our knowledge, there is no accepted measurement error (systematic and random) that indicates when the agreement of a $\dot{V}O_2$peak estimation is acceptable

Table 2 Accuracy between non-exercise $\dot{V}O_2$peak estimations and a gold standard cardiopulmonary exercise test (CPET)

	Pearson r	Bias ml·min^{-1}·kg^{-1}	SEE ml·min^{-1}·kg^{-1}	MAPE %	CoV %
SeismoFit 1	0.70	-3.1 ± 2.4*	6.6	12.0	9 ± 3
Polar Fitness Test	0.79	4.4 ± 3.0*	8.3	14.8	10 ± 3

Data are presented as mean ± 95% CI. n=20. SeismoFit 1, $\dot{V}O_2$peak estimation model using seismocardiography version 4.3; Polar, $\dot{V}O_2$peak estimation using the Polar Fitness Test; SEE, standard error of estimate; MAPE, mean absolute percentage error. CoV; coefficient of variation.
*significantly different from CPET ($p<0.05$).

or not. However, it has previously been suggested that the contextual validity of estimating $\dot{V}O_2$peak requires a measurement error within the range of 3.5 mL·kg^{-1}·min^{-1}, if the purpose is to assess improvement in the subject's health (ie, changes in $\dot{V}O_2$peak); otherwise, relevant clinical changes would be missed.[42] This is based on the association between a small increase in $\dot{V}O_2$peak (1 METS or 3.5 mL·kg^{-1}·min^{-1}) and a lower risk of all-cause mortality and incidence of CVD.[1,4] For $\dot{V}O_2$peak estimation models using either resting condition or exercise-based information, this still is not achieved as the estimation error at the individual level is too large.[42]

The present results revealed a systematic error of –3.1 ± 2.4 mL·kg^{-1}·min^{-1} with a substantial total random error span of 21.6 mL·kg^{-1}·min^{-1}. In addition, the MAPE was above the predefined level of <10% for clinical relevance. This advocates for future efforts to improve this methodology of estimating $\dot{V}O_2$peak, both at the population level and in terms of evaluating individual clinically relevant changes. This is a slight contrast to a previous study applying the same SCG algorithm version, in which the model proved valid in estimating $\dot{V}O_2$peak at a population level, based on the

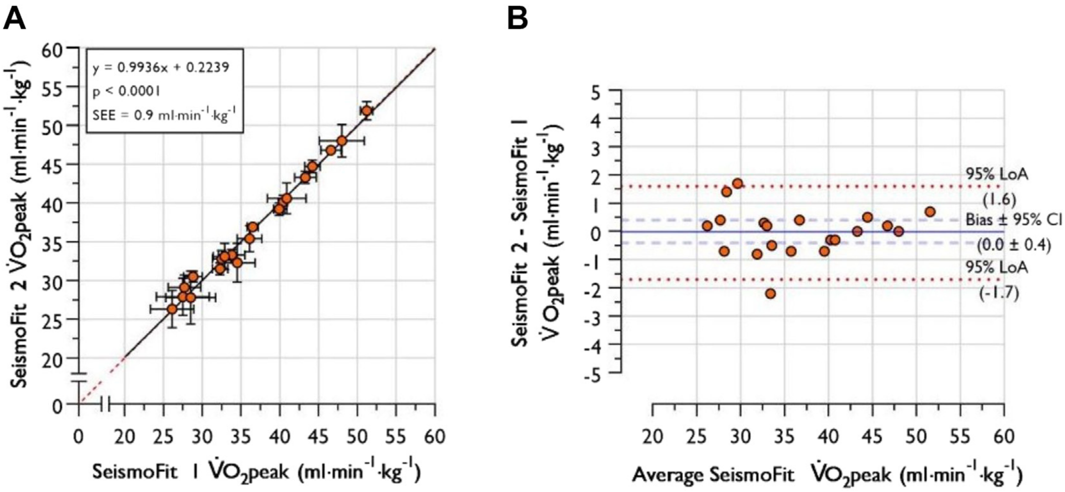

Figure 4 Data (n = 20) from 3 separate tests for each participant (mean ± 95% CI) . **A:** Scatter plot with linear regression between 2 $\dot{V}O_2$peak estimations performed with a nonexercise model using seismocardiography (SeismoFit 1 and SeismoFit 2). *Red dotted line* represents the line of identity (r = 1.0). **B:** Bland-Altman plot of the agreement between SeismoFit 1 and SeismoFit 2. *Blue line* represents the bias. *Blue dotted line* represents the 95% CI of the bias. *Red dotted line* represents the 95% LoA. Abbreviations as in Figure 3.

systematic error (n = 97; bias –1.0 ± 1.2 $mL \cdot kg^{-1} \cdot min^{-1}$).[28] Otherwise, the results were similar to the previous study (LoA –12.4 to 10.4 $mL \cdot kg^{-1} \cdot min^{-1}$; SEE 5.9 $mL \cdot kg^{-1} \cdot min^{-1}$; MAPE 12.3%; CoV 8% ± 1%; Pearson r 0.73).[28] One reason for the discrepancy in systematic error might be related to the studied population. The data used for developing the algorithm version 4.3 consisted primarily of subjects between the ages of 18 and 45 years. In addition, a proportional error was previously observed, which indicated that the SCG 4.3 version underestimates $\dot{V}O_2$peak more the better trained the subject is.[28] In the present study, the 2 participants with the largest difference between Seismofit and CPET (–13.6 and –14.7 $mL \cdot kg^{-1} \cdot min^{-1}$) (Figure 3C) were also the participants with the overall best training status, as they both were between 70 and 72 years old with rather unusual $\dot{V}O_2$peak values of 43.2 $mL \cdot kg^{-1} \cdot min^{-1}$ (female) and 45.8 $mL \cdot kg^{-1} \cdot min^{-1}$ (male). When participants >70 years were removed from the analysis (n = 3), the agreement analysis showed a bias of –1.8 ± 2.3 $mL \cdot kg^{-1} \cdot min^{-1}$ (P = .119) with 95% LoA ranging from –10.4 to 6.8 $mL \cdot kg^{-1} \cdot min^{-1}$ (Supplemental Figure 4). This implies that the current Seismofit algorithm might be restricted to estimating $\dot{V}O_2$peak at a population level with an upper age limit of 70 years and that adjustments in the model are needed for expansion of the target population to the older population. In addition, a new version of the algorithm was developed right after the current study was finalized. The new 4.5 algorithm only included training data from the previous cross-sectional study with subjects ranging between 18 and 65 years of age[28] and no data from the current study. When the 4.5 algorithm was applied to the current data, the agreement between methods was good (Supplemental Table 4 and Supplemental Figure 5).

Comparison of methods for $\dot{V}O_2$peak estimation

A recently published meta-analysis covering 8 studies validating the PFT showed an average overestimation of 2.2 ± 1.9 $mL \cdot kg^{-1} \cdot min^{-1}$ (mean ± 95% CI) with LoA ranging from –13.1 to 17.4 $mL \cdot kg^{-1} \cdot min^{-1}$ compared with CPET.[42] In addition, 3 of the 4 most recently published studies using PFT algorithms probably are most comparable to the PFT algorithm applied in this study. These studies showed bias and 95% CI between 3.0 ± 3.1 $mL \cdot kg^{-1} \cdot min^{-1}$ and 4.7 ± 3.4 $mL \cdot kg^{-1} \cdot min^{-1}$,[15–17] whereas the fourth study found a bias of –0.5 ± 2.8 $mL \cdot kg^{-1} \cdot min^{-1}$ compared with CPET.[43] Even though the PFT algorithm may have been adjusted to improve the estimate of $\dot{V}O_2$peak, these findings are in line with the present bias of 4.4 ± 3.1 $mL \cdot kg^{-1} \cdot min^{-1}$ and thus further highlight that PFT overestimates $\dot{V}O_2$peak compared with CPET. When the validity of the Seismofit and PFT $\dot{V}O_2$peak estimations was compared, they both showed an inadequate agreement; however, the difference in LoA was smaller by 2.4 $mL \cdot kg^{-1} \cdot min^{-1}$ and the MAPE lower by 2.8% points for the Seismofit, indicating a better agreement than PFT. The Seismofit estimation is entirely objective, in contrast to PFT where self-reported weekly hours influence the estimated $\dot{V}O_2$peak. This is a strength in the Seismofit model as people often tend to overestimate their activity level, and perhaps this is the reason for PFT overestimating $\dot{V}O_2$peak.

Reliability of $\dot{V}O_2$peak measurement and estimation

Reliability assessment in method comparison studies is often lacking, probably due to requirements of an extended study design. However, when incorporated it provides important knowledge in the understanding of the quality of the method. Directly measured $\dot{V}O_2$ during a graded CPET is considered the most accurate method for determination of $\dot{V}O_2$peak, and reproducibility typically is reported with CoV around 5%.[44–47] Reliability of CPET in the present study was high, with excellent ICC and within-subject CoV of 3.4% between test days observed. Therefore, CPET was considered acceptable as a reference method. Based on excellent ICC and within-subject CoV of 4.5% (Seismofit 1) and 3.8% (Seismofit 2), Seismofit $\dot{V}O_2$peak estimation is highly reliable between test days. Furthermore, the test–retest also proved that Seismofit repeatability was high. The consistency in within-subject $\dot{V}O_2$peak estimations is vital because it confirms that SCG signal recording, features, and fiducial point extraction and processing by the Seismofit are reliable. This indicates that the lack of accuracy probably is algorithm-dependent and not related to Seismofit SCG signal recording. SCG algorithm version 4.3 is trained using data from $<$150 healthy subjects within a relatively small age span, who performed a resting SCG recording and a CPET cycle ergometer test.[26,27]

PFT has previously been concluded to be reliable based on Pearson r = 0.91 between 2 test days.[43] In the present study, PFT was proved highly reliable between 3 test days, with within-subject CoV of 3.0% ± 1.4% and ICC of 0.97.

Study limitations

The lack of a familiarization test day for the participants to the used equipment and exercise testing is a limitation. However, it was prioritized to include a third test day instead, as few method validation studies include more than 2 test days. There were no differences in the measured $\dot{V}O_2$peak values between test days. The recruitment of the trained elderly participants is a limitation as they might not reflect the trained elderly population in general, but rather the very top end. Lastly, an obvious limitation is missing the 2 CPET tests due to measurement error and not including a fourth test day within a reasonable timeframe.

Conclusion

In the current study, Seismofit $\dot{V}O_2$peak estimation correlated to the gold standard CPET but underestimated $\dot{V}O_2$peak with a significant systematic error. Compared to PFT, Seismofit showed better agreement based on a smaller random error and MAPE. The reliability of Seismofit $\dot{V}O_2$peak estimation was high both within test days and between test days, with CoV being low and ICC excellent, indicating that the Seismofit system is dependable in recording and processing the SCG signal. Consequently, because the $\dot{V}O_2$peak estimation model is cloud-based, it is possible that Seismofit $\dot{V}O_2$peak estimation could improve and eventually be considered applicable in clinical settings when more data are available for training the algorithm. Thus, based on the present findings, the Seismofit $\dot{V}O_2$peak estimation model needs further improvement if it is to be considered for use in a clinical setting.

Acknowledgments

All participants are thanked for their commitment and participation in the study. VentriJect Aps are thanked for providing the equipment for performing the $\dot{V}O_2$peak estimation.

Funding Sources

This work was financially supported by VentriJect Aps.

Disclosures

VentriJect Aps support two-thirds of Dr Mikkel Hansen's scholarship. Dr Samuel Emil Schmidt and Dr Kasper Sørensen hold shares in VentriJect Aps. Dr Schmidt works part-time as Chief Scientific Officer and Dr Sørensen full-time as Head of Software at VentriJect Aps. There was no restriction on publication. Declaration of Generative AI and AI-assisted Technologies in the writing process: No generative AI or AI-assisted technologies have been used in the preparation of this work.

Authorship

All authors attest they meet the current ICMJE criteria for authorship.

Patient Consent

All patients provided written consent.

Ethics Statement

The study conforms to the Helsinki Declaration and was approved by the Science Ethical Committee of the Greater Region of Copenhagen, Denmark (H-19081375).

Data availability

All data that were used in the current study are available upon reasonable request. The estimated and measured $\dot{V}O_2$peak values on the 3 test days from all the participants are given in Supplemental Table 3.

Appendix

Supplementary data

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.cvdhj.2023.08.020.

References

1. Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. Circulation 2016;134:e653–e699.
2. Laukkanen JA, Kurl S, Salonen R, Rauramaa R, Salonen JT. The predictive value of cardiorespiratory fitness for cardiovascular events in men with various risk profiles: a prospective population-based cohort study. Eur Heart J 2004; 25:1428–1437.
3. Blair SN, Kohl HW 3rd, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. JAMA 1989;262:2395–2401.
4. Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events. J Am Med Assoc 2009; 301:2024–2035.
5. Albouaini K, Egred M, Alahmar A, Wright DJ. Cardiopulmonary exercise testing and its application. Postgrad Med J 2007;83:675–682.
6. Gupta S, Rohatgi A, Ayers CR, et al. Cardiorespiratory fitness and classification of risk of cardiovascular disease mortality. Circulation 2011;123:1377–1383.
7. Myers J, Nead KT, Chang P, Abella J, Kokkinos P, Leeper NJ. Improved reclassification of mortality risk by assessment of physical activity in patients referred for exercise testing. Am J Med 2015;128:396–402.

8. Lee DC, Artero EG, Sui X, Blair SN. Mortality trends in the general population: the importance of cardiorespiratory fitness. J Psychopharmacol 2010;24:27–35.
9. Lee D, Sui X, Artero E, Lee I, Church T, McAuley P. Long-term effects of changes in cardiorespiratory fitness and body mass index on all-cause and cardiovascular disease. Circulation 2011;124:2483–2490.
10. Mikkelsen N, Cadarso-Suárez C, Lado-Baleato O, et al. Improvement in VO2-peak predicts readmissions for cardiovascular disease and mortality in patients undergoing cardiac rehabilitation. Eur J Prev Cardiol 2020;27:811–819.
11. Stamatakis E, Hamer M, O'Donovan G, Batty GD, Kivimaki M. A non-exercise testing method for estimating cardiorespiratory fitness: associations with all-cause and cardiovascular mortality in a pooled analysis of eight population-based cohorts. Eur Heart J 2013;34:750–758.
12. Myers J, Kaminsky LA, Lima R, Christle JW, Ashley E, Arena R. A reference equation for normal standards for VO2Max: analysis from the Fitness Registry and the Importance of Exercise National Database (FRIEND Registry). Prog Cardiovasc Dis 2017;60:21–29.
13. Malek MH, Berger DE, Housh TJ, Coburn JW, Beck TW. Validity of V̇O2max equations for aerobically trained males and females. Med Sci Sports Exerc 2004;36:1427–1432.
14. Houle SA, Sui X, Blair SN, Ross R. Association between change in nonexercise estimated cardiorespiratory fitness and mortality in men. Mayo Clin Proc Innov Qual Outcomes 2022;6:106–113.
15. Kraft GL, Dow M. Validation of the Polar Fitness Test. Int J Innov Educ Res 2018;6:27–34.
16. Passler S, Bohrer J, Blöchinger L, Senner V. Validity of wrist-worn activity trackers for estimating VO2max and energy expenditure. Int J Environ Res Public Health 2019;16:3037.
17. Snyder NC, Willoughby CA, Smith BK. Comparison of the Polar V800 and the Garmin Forerunner 230 to predict V̇O2max. J Strength Cond Res 2021; 35:1403–1409.
18. Mounsey P. Praecordial ballistocardiography. Br Heart J 1957;19:259–271.
19. Salerno DM, Zanetti J. Seismocardiography: a new technique for recording cardiac vibrations concept method and initial observations. J Cardiovasc Technol 1990;9:111–118.
20. Taebi A, Solar B, Bomar A, Sandler R, Mansy H. Recent advances in seismocardiography. Vibration 2019;2:64–86.
21. Bassett DR, Howley ET. Limiting factors for maximum oxygen uptake and determinants of endurance performance. Med Sci Sports Exerc 2000;32:70–84.
22. Grewal J, McCully RB, Kane GC, Lam C, Pellikka PA. Left ventricular function and exercise capacity. JAMA 2009;301:286–294.
23. Sørensen K, Schmidt SE, Jensen AS, Søgaard P, Struijk JJ. Definition of fiducial points in the normal seismocardiogram. Sci Rep 2018;8:1–11.
24. Dehkordi P, Khosrow-Khavar F, Di Rienzo M, et al. Comparison of different methods for estimating cardiac timings: a comprehensive multimodal echocardiography investigation. Front Physiol 2019;10:1057.
25. Agam A, Søgaard P, Kragholm K, et al. Correlation between diastolic seismocardiography variables and echocardiography variables. Eur Heart J Digit Health 2022;3:465–472.
26. Sørensen K, Poulsen MK, Karbing DS, Søgaard P, Struijk JJ, Schmidt SE. A clinical method for estimation of VO2max using seismocardiography. Int J Sports Med 2020;41:661–668.
27. Hansen MT, Grønfeldt BM, Rømer T, et al. Determination of maximal oxygen uptake using seismocardiography at rest. Comput Cardiol IEEE 2021;48–51.
28. Hansen MT, Husted KLS, Fogelstrøm M, et al. Accuracy of a clinical applicable method for prediction of VO2max using seismocardiography. Int J Sport Med 2023;44:1–7.
29. VentriJect. The SeismoFit System, https://ventriject.com/product/. Accessed June 1, 2023.
30. Polar Fitness Test, https://support.polar.com/en/the-what-and-how-of-polar-fitness-test?. Accessed June 1, 2023.
31. Buchfuhrer MJ, Hansen JE, Robinson TE, Sue DY, Wasserman K, Whipp BJ. Optimizing the exercise protocol for cardiopulmonary assessment. J Appl Physiol 1983;55:1558–1564.
32. Borg GA. Psychophysical bases of perceived exertion. Med Science Sports Exerc 1982;14:377–3781.
33. Taylor H, Henschel A, Buskirk E. Maximal oxygen intake as an objective measure of cardio-respiratory performance. J Appl Physiol 1955; 8:73–80.
34. Howley ET, Bassett DR Jr, Welch HG. Criteria for maximal oxygen uptake: review and commentary. Med Sci Sport Exerc 1995;27:1292–1301.
35. Beltz NM, Gibson AL, Janot JM, Kravitz L, Mermier CM, Dalleck LC. Graded exercise testing protocols for the determination of VO2max: historical perspectives, progress, and future considerations. J Sports Med 2016; 2016:1–12.
36. Krouwer JS. Why Bland-Altman plots should use X, not (Y + X)/2 when X is a reference method. Stat Med 2008;27:778–780.
37. Bland MJ, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–310.
38. Wier LT, Jackson AS, Ayers GW, Arenare B. Nonexercise models for estimating V̇O2max with waist girth, percent fat, or BMI. Med Sci Sports Exerc 2006; 38:555–561.
39. Hinkle DE, Wiersma W, Jurs SG. Correlation: a measure of relationship. In: Hinkle DE, Wiersma W, Jurs SG, eds. Applied Statistics for the Behavioral Sciences. Boston: Houghton Mifflin Company; 1994. p. 103–128.
40. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–163.
41. Giavarina D. Understanding Bland Altman analysis. Biochem Medica 2015; 25:141–151.
42. Molina-Garcia P, Notbohm HL, Schumann M, et al. Validity of estimating the maximal oxygen consumption by consumer wearables: a systematic review with meta-analysis and expert statement of the INTERLIVE Network. Sport Med 2022;52:1577–1597.
43. Cooper KD, Shafer AB. Validity and reliability of the Polar A300's fitness test feature to predict VO2max. Int J Exerc Sci 2019;12:393–401.
44. Decato TW, Bradley SM, Wilson EL, Hegewald MJ. Repeatability and meaningful change of CPET parameters in healthy subjects. Med Sci Sports Exerc 2018; 50:589–595.
45. Skinner JS, Wilmore KM, Jaskolska A, et al. Reproducibility of maximal exercise test data in the HERITAGE family study. Med Sci Sports Exerc 1999; 31:1623–1628.
46. Kleinloog JPD, Laar SPGA van, Schoffelen PFM, Plasqui G. Validity and reproducibility of VO2max testing in a respiration chamber. Scand J Med Sci Sport 2021;31:1259–1267.
47. Bensimhon D, Leifer E, Ellis S, et al. Reproducibility of peak oxygen uptake and other cardiopulmonary exercise testing parameters in patients with heart failure. Am J Cardiol 2008;102:712–717.