Validity of the Actiheart step test for the estimation of maximum oxygen consumption in endurance athletes and healthy controls

Juliane Heydenreich, Yves Schutz, Bengt Kayser & Katarina Melzer

1 Swiss Federal Institute of Sport, Magglingen, Switzerland
2 Institute of Sport Sciences, University of Lausanne, Lausanne, Switzerland
3 Department of Physiology, University of Fribourg, Fribourg, Switzerland

+ Shared last authorship
* Corresponding author: Prof. Dr. Bengt Kayser, Institute of Sport Sciences, University of Lausanne, Synathlon Uni-Centre, 1015 Lausanne, Switzerland
Tel: +41 (0)21 692 37 95, Fax: +41 (0)21 692 32 93
Email: bengt.kayser@unil.ch

ABSTRACT

Submaximal step tests are often used for estimation of maximum oxygen consumption (VO$_2$max) in humans. The validity of the Actiheart step test for VO$_2$max estimation was not fully studied yet. Therefore, purpose of the study was to estimate VO$_2$max using the Actiheart step test and to compare the data with measured VO$_2$max in endurance trained athletes (ATH) and healthy non-athletes (CON). 68 ATH (54% men, 28.0±5.4 yrs, 20.9±1.7 kg∙m$^{-2}$) and 63 CON (46% men, 27.6±5.1 yrs, 22.1±1.7 kg∙m$^{-2}$) performed the Actiheart step test and a spiroergometry for assessment of VO$_2$max. In addition, resting metabolic rate (RMR; indirect calorimetry), maximum heart rate (HRmax; heart rate monitoring system during spiroergometry), and sleeping heart rate (SHR; Actiheart 6-day long term measurement) were determined. Validity of two different Actiheart software entry modes was assessed: (1) AHraw (estimated RMR [Schofield] and HRmax [Tanaka], SHR = 70 bpm) and (2) AHcomplete (measured RMR, HRmax, and SHR). Validity was investigated using linear regression ($R^2$ and standard error of the estimate (SEE)) and repeated-measures ANOVA with a Bonferroni post-hoc correction. The level of significance was set to $\alpha=0.05$.

VO$_2$max estimated by AHraw was significant related to measured VO$_2$max in women CON ($R^2=0.22$; $p<0.05$), whereas when VO$_2$max was estimated by AHcomplete the relation was significant in women ATH and CON, and in men CON ($R^2=0.17-0.24$; $p<0.05$). AHraw significantly underestimated VO$_2$max in the total sample by 8% (51.4±10.2 vs. 55.9±7.6 ml∙kg$^{-1}$∙min$^{-1}$; $p<0.0001$), whereas no significant difference between AHcomplete and the criterion method was found (57.0±11.1 vs. 55.9±7.6 ml∙kg$^{-1}$∙min$^{-1}$; $p=0.26$).

The Actiheart step test is an acceptable tool for the estimation of VO$_2$max if an error within 8% can be tolerated. However, accuracy of the VO$_2$max prediction is much improved when entering measured variables, such as RMR, SHR, and HRmax, into the software.

Keywords:
Maximum oxygen consumption – athletes – cardiorespiratory fitness – exercise testing – metabolic equivalent

Citation:
Introduction

Maximum oxygen consumption (VO$_{2\text{max}}$), a key indicator of cardiorespiratory fitness, is used in both athletic and health settings, as a determinant of physical performance (Bassett & Howley, 2000) or as a predictor of health risk and longevity (Kodama et al., 2009). VO$_{2\text{max}}$-testing is usually performed on a treadmill or a cycle ergometer while oxygen uptake and expired carbon-dioxide are measured with a respiratory gas analyzer. The workload is progressively increased until the participant reaches volitional exhaustion (after a minimum of 5 min (treadmill) or 7 min (cycle), up to 26 min (Midgley, Bentley, Luttikholt, McNaughton, & Millet, 2008)). However, maximum exercise tests are time consuming, expensive and depend on physiological expertise (Björkman, Ekblom-Bak, Ekblom, & Ekblom, 2016). In addition, some individuals from the general population cannot achieve the maximal effort required for the determination of VO$_{2\text{max}}$. Furthermore, there are contraindications for maximum exercise tests (American Thoracic Society & American College of Chest Physicians, 2003). These contraindications include uncontrolled asthma, syncope, acute myocardial infarction, and respiratory failure. Therefore, several indirect methods to estimate VO$_{2\text{max}}$ have been developed, where VO$_{2\text{max}}$ is predicted from submaximal exercise results. Most of these tests use the linear relationship between heart rate (HR) and oxygen uptake (VO$_{2}$). The VO$_{2\text{max}}$ is then predicted by extrapolation of the submaximal values to an estimated maximum heart rate (HRmax).

Submaximal tests thus allow evaluation of cardiorespiratory fitness in a population, in which the direct determination of VO$_{2\text{max}}$ is not possible (Björkman et al., 2016).

The Actiheart is a lightweight (10 g), waterproof combined HR and movement sensor (accelerometer) designed to non-invasively assess daily physical activity levels (Cambridge Neurotechnology Ltd., Papworth, UK). The Actiheart was shown to give accurate estimations of activity energy expenditure against indirect calorimetry during a wide range of activities in men and women in both laboratory (Thompson, Batterham, Bock, Robson, & Stokes, 2006) and field (Crouter, Churilla, & Bassett, 2008) settings. A good level of agreement between Actiheart and doubly labelled water measured daily total energy expenditure was found in adult men and women (Brage et al., 2015), in children and adolescents (Butte et al., 2010), and in lean and overweight men of various fitness levels (Villas et al., 2012). The Actiheart needs to be individually calibrated for each person with a standard step test, a built-in function of the Actiheart software. During stepping, the relationships between actimetry and HR vs. work rate are assessed. The step test also yields estimated VO$_{2\text{max}}$ values, obtained by extrapolating the HR vs. work rate regression line to the estimated or measured HRmax for the individual. No studies assessed the validity of the VO$_{2\text{max}}$ estimation using the Actiheart step test. The purpose of the study was to estimate VO$_{2\text{max}}$ using the Actiheart step test in AHRaw and AHcomplete modes (see “Methods”) and to compare the results with measured VO$_{2\text{max}}$ over a range of aerobic capacities. The results were expected to fill a gap in the knowledge regarding the practical use and precision of the Actiheart modes for estimating VO$_{2\text{max}}$ in comparison to the VO$_{2\text{max}}$ ergometer measurements for athletic and non-athletic populations.

Methods

Participants

We recruited by advertisement 68 competitive endurance athletes (ATH; 31 women, 37 men; regular endurance training volume ≥ 300 min · wk$^{-1}$ and participation in competitions) and 63 healthy, non-endurance-trained nonsmoking controls (CON; 34 women, 29 men; max. 150 min · wk$^{-1}$ moderate endurance training). Most but not all participants were workers, students or athletes located on the campus of the Swiss Federal Institute of Sports Magglingen. All participants had to be weight-stable (< 2 kg of weight difference in the last 3 months) with a Body Mass Index (BMI) between 18.5 and 25 kg · m$^{-2}$, and aged between 18 and 40 years. Participants were excluded if they were smoking, pregnant, lactating, dieting, suffering from metabolic disease and/or eating disorders, or taking medication (apart from contraceptives). Athletes were excluded from the study if they had changed their training habits within the last four weeks before the experiments (e.g., due to injury or disease). All experimental procedures were approved by the Regional Ethics Committee in Berne, Switzerland (KEK-number 090/15), and the study was carried out according to the recommendations of the Helsinki Declaration. Written informed consent of the participants was obtained before any testing.

Study design

The participants were recruited by advertisement. They came to the testing center on two separate testing days, having refrained from strenuous physical activity ≥ 24 h. After providing written informed consent, on the first day the participants completed, in a fasted state (≥ 12 h absence of any food or fluid intake, ≥ 36 h absence of alcohol or caffeine intake) measurements in the following order: 1) anthropometry and body composition; 2) resting metabolic rate (RMR); and 3) a step test using the in-built function of the Actiheart. One week after the first measurement day, the participants did an incremental exercise test (VO$_{2\text{max}}$) in a non-fasted state. In the days between the two testing days the participants wore an Actiheart for at least 7 days. All tests were carried out in Magglingen (Switzerland) at an altitude of 950 m.

Estimation of VO$_{2\text{max}}$ with the Actiheart

The Actiheart was clipped onto two standard ECG electrodes (3M™ Red Dot™ Electrode 2560; 3M Health Care, St. Paul, USA) on the chest of the participant according to manufacturer’s instructions and worn day and night (Brage et al., 2006). The de-
vice was individually calibrated for each participant using the standard step test, a built-in function of the Actiheart software version 4.0.92 (Cambridge Neurotechnology Ltd., Papworth, UK). In short, the participants stepped up and down a 195 mm high step for eight minutes while stepping rate was ramped linearly from 15 step cycles ∙ min⁻¹ (one step cycle is “up, up, down, down”) to 33 step cycles ∙ min⁻¹ at the end (rate of change 2.5 body lifts ∙ min⁻¹). Participants were advised to stop the test if they felt uncomfortable. After the end of the test, participants were requested to stand still for two minutes with the Actiheart still fitted while not speaking during which recovery HR was assessed. The mechanical power of the step test (mass-specific lift work rate) was calculated as 9.81 m ∙ s⁻² ∙ step height (m) ∙ lift frequency (number of body weight lifts ∙ min⁻¹) and expressed in J ∙ kg⁻¹ ∙ min⁻¹. The Actiheart software uses linear regression to model the relationship between work rate and HR during stepping. To estimate the VO₂max, the regression line was extrapolated to the estimated HRmax for each participant. The predicted VO₂max in L ∙ kg⁻¹ ∙ min⁻¹ was then calculated as a maximum mechanical power in J ∙ kg⁻¹ ∙ min⁻¹ divided by a factor of 20.2 (assuming an invariant net efficiency of 20%). The HRmax was defined in two ways: (1) estimated by use of the Tanaka equation (Tanaka, Monahan, & Seals, 2001); and (2) measured during a maximum exercise test (see section “Measurement of VO₂max”).

Total energy expenditure (TEE) and activity energy expenditure were estimated by analyzing 6 full-day (24 h) recordings of HR and body movement with a 15-second averaging epoch setting. During the recording period participants were requested to continue their habitual life routine and physical activities. Sleeping heart rate (SHR) was measured as a mean of the 120 lowest HR recordings during 24 h, the SHR for a full six-day period was determined as the mean value of the six SHR recordings. TEE was calculated as sum of RMR (measured by indirect calorimetry, see section “Resting metabolic rate”), activity energy expenditure, and diet-induced thermogenesis (estimated as 10% of TEE (Manore & Thompson, 2010)). Physical activity level (PAL) was calculated as TEE/RMR.

For the validity testing, two inbuilt scenarios of the Actiheart step test were looked at: (1) “AHraw”. In this mode SHR is set to 70 bpm, RMR is estimated by the Schofield equation (Schofield, 1985), and HRmax is estimated by use of the Tanaka formula (Tanaka et al., 2001); and (2) “AHcomplete”. In this mode SHR, measured during the long-term recordings of the Actiheart, RMR, measured using indirect calorimetry, and HRmax, measured during a maximum exercise test, are manually entered into the Actiheart software.

Measurement of VO₂max

Before exercise testing each participant filled out the German (Marti, Villiger, Hintermann, & Lerch, 1998) or French (Société canadienne de physiologie de l’exercice, 2002) version of the Physical Activity Readiness Questionnaire (PAR-Q). Only if all items were answered with “no”, participants were allowed to start the exercise testing. The test was performed on a treadmill (women: model mercury; men: model venus, h/p/Cosmos Sports & Medical GmbH, Traunstein, Germany). Treadmill inclines were set at 4° throughout the test. After a 5-min warm-up jog, non-athletic participants began running at 7 km ∙ h⁻¹, whereas participants from the athlete group started at 9 km ∙ h⁻¹. The speed was increased by 1 km ∙ h⁻¹ every minute for the first 3 minutes of the test and thereafter by 0.5 km ∙ h⁻¹ every 30 s until exhaustion (Steiner & Wehrlin, 2011). Gas exchange was measured breath-by-breath with an open-circuit system (Quark CPET; COSMED Srl, Rome, Italy), which was calibrated before each test according to manufacturer’s instructions. VO₂ data was processed using 10-second time averages. VO₂max was determined as the highest 30-second VO₂ average for the test (Robergs, Dwyer, & Astorino, 2010). Heart rate was continuously registered with a wireless HR monitoring system (model SZ990; COSMED Srl, Rome, Italy). If the primary criteria of a plateau in oxygen uptake (defined as an increase of VO₂ < 2.1 ml ∙ kg⁻¹ ∙ min⁻¹) was not reached by the participant (n = 4), then the secondary criteria of a RQ value ≥ 1.10, and a HR close (± 10 bpm) to the age-predicted HRmax (Tanaka et al., 2001) were used to determine whether the participant reached maximal effort and thus VO₂max. VO₂max-tests were carried out at a mean room temperature of 21.8 ± 1.0°C, a humidity of 39.8 ± 10.1%, and an air pressure of 914 ± 7 hpa. In general, a temperature range of 20 to 22°C in a cool, dry environment (< 50% humidity) is considered comfortable for exercise testing (Myers et al., 2009).

Anthropometric data and body composition

Body weight was measured to the nearest 0.1 kg on a calibrated beam scale (Seca 877, Seca, Hamburg, Germany) and body height was measured to the nearest 0.5 cm with a height rod (Seca 213, Seca, Hamburg, Germany), with the participants in light clothing and without shoes. Body composition was assessed using Lunar iDXA (GE Healthcare, Madison, WI, USA). Calibration of the iDXA was performed and checked on a daily basis before testing using a calibration phantom. Participants voided their bladder before the scan. The participants were in underwear and all metal artefacts were removed. During measurement, participants were in supine position on the scanning table with their ankles and legs fixed using supports. Arms were positioned to the side with the palms flat on the table. Participants were requested to stay still during the measurement. Whole body scans were performed according to the manufacturer’s instructions, and adipose tissue mass, lean tissue mass, and bone mineral content were derived (enCore software v. 11.10; GE Healthcare, Madison, WI, USA). Total body composition estimates with the Lunar iDXA have been reported to be excellent in other studies (Carver, Christou, & Andersen, 2013; Hind, Oldroyd, & Truscott, 2011).
Resting metabolic rate

RMR was assessed by indirect calorimetry using a ventilated hood system (Quark CPET; COSMED Srl, Rome, Italy). Calibrations of the gas analyzer and flowmeter were carried out before each test according to manufacturer’s instructions. After acclimatization and relaxing for 30 min on a bed, a ventilated hood was placed over the participant’s head and measurements were started. VO₂ and carbon dioxide production (VCO₂) were measured for 30 min at 10-second intervals with participants remaining motionless in supine position in a thermoneutral environment (20 - 25 °C (Compher, Frankenfeld, Keim, Roth-Yousey, & Evidence Analysis Working Group, 2006)). The first 5 min were eliminated as acclimatization artifact. From the remaining 25 min an interval of 5 consecutive minutes with a < 10% coefficient of variation of VO₂ and VCO₂ was considered as steady state. VO₂ and VCO₂ were then used to calculate RMR using the abbreviated Weir equation (Weir, 1949). Since respiratory quotient (RQ, defined as the ratio of VCO₂ and VO₂) measures at rest < 0.70 and > 1.00 suggest protocol violations or inaccurate gas measurement (Compher et al., 2006), participants with RQ-values outside of the plausible range should be excluded from data analysis (n = 1). The obtained RQ for all measurements was 0.76 ± 0.04 and the mean coefficients of variation of VO₂ and VCO₂ were 3.8 ± 1.5% and 5.0 ± 1.9%. RMR measurements took place at a mean temperature of 21.9 ± 1.2°C, 40.0 ± 10.2% humidity, and an air pressure of 914 ± 7 hpa.

Statistics

Statistical analyses were performed using the statistical software SPSS version 24 for MS-Windows (IBM Corp., Chicago, IL, USA). Mean values and standard deviations (SD) were calculated and data was checked for normality using the Shapiro-Wilk-test. All parameters were normal distributed with the exception of age, body fat (%), fat-free mass (kg), RMR (kcal ∙ d⁻¹), VO₂max (L ∙ min⁻¹), recovery HR, and PAL. Group differences were tested by independent t-tests and Mann-Whitney-U-tests (α = 0.05). The validity of the Actiheart step test was first investigated using Pearson’s Product moment correlation analysis. The correlation coefficients (r) were classified according to Cohen (Cohen, 1988). An r between 0.10-0.29 was considered a small, between 0.30-0.49 a moderate and between 0.50-1.0 a strong association. The data were further analyzed using a repeated-measures ANOVA with a Bonferroni post-hoc correction to compare measured VO₂max, estimated VO₂max by AHraw, and estimated VO₂max by AHcomplete. In addition, the mean absolute error (MAE) and the mean absolute percentage error (MAPE) of AHraw and AHcomplete compared to the criterion measure were computed. Since no standardized threshold exists for high or low MAPE, we considered a MAPE ≥ 10% as an indicator of inaccuracy as suggested by other authors (Boudreaux et al., 2018; Lee, Kim, & Welk, 2014; Nelson, Kaminsky, Dickin, & Montoye, 2016; Roos, Taube, Beeler, & Wyss, 2017). The R² and standard error of the estimate (SEE) were calculated by linear regression, where measured VO₂max was entered as dependent variable and estimated VO₂max as independent variable. Levels of agreement between Actiheart and measured values were further expressed as mean difference with Limits of Agreement (mean difference ± 1.96 SD) (Bland & Altman, 1999).

Results

Participants

Data of in total 50 women (24 ATH and 26 CON) and 41 men (22 ATH and 19 CON) with valid Actiheart, VO₂max, and RMR measurements were included into the analysis. The Table 1 presents the anthropometric data, body composition, RMR and VO₂max of the participants. The VO₂max ranged from 39.1 – 65.4 ml ∙ kg⁻¹ ∙ min⁻¹ in women and 42.8 – 78.4 ml ∙ kg⁻¹ ∙ min⁻¹ in men.

Validity of the Actiheart step test for estimation of VO₂max

The Pearson’s correlation analyses revealed moderately correlated data of AHraw and the criterion method only in women CON (r = 0.46, p < 0.05), whereas the correlation between AHcomplete and measured VO₂max was significant in women ATH, women CON, and men CON (r = 0.41 – 0.49, all p < 0.05; Table 2), and r was interpreted as moderate association. The MAE, MAPE, and SEE for all groups are presented in Table 2. In women and men ATH the AHraw significantly underestimated VO₂max (p < 0.05), whereas in the CON groups no significant differences between estimated and measured VO₂max were found. For AHcomplete, VO₂max was significantly underestimated in women ATH (p = 0.03), and overestimated in men CON (p = 0.03), whereas no significant differences were obtained in women CON and men ATH. When looking at the total sample, significant differences between AHraw and the measured VO₂max (51.4 ± 10.2 vs. 55.9 ± 7.6 ml ∙ kg⁻¹ ∙ min⁻¹; p < 0.0001) were observed, whereas the difference between AHcomplete and VO₂max was non-significant (57.0 ± 11.1 vs. 55.9 ± 7.6 ml ∙ kg⁻¹ ∙ min⁻¹; p = 0.26). The range of the MAPE across all groups was 11.4 – 17.7% and 10.8 – 14.7% for AHraw and AHcomplete, respectively (Table 2). The absolute difference between measured and estimated VO₂max by use of AHraw was smaller than 5 ml ∙ kg⁻¹ ∙ min⁻¹ in 29% (n = 7), 54% (n = 14), 27% (n = 6), and 58% (n = 11) in women ATH, women CON, men ATH, and men CON, respectively. When VO₂max was estimated by AHcomplete, the absolute difference between measured and estimated VO₂max was smaller than 5 ml ∙ kg⁻¹ ∙ min⁻¹ in 42% (n = 10), 39% (n = 10), 50% (n = 11), and 37% (n = 7) of women ATH, women CON, men ATH, and men CON, respectively. The results of the linear regression analyses are presented in Table 3. In women CON a statistically significant relationship between VO₂max estimated by AHraw and the measured val-
Table 1: Participant characteristics of the men and women athletes (ATH) and controls (CON).

<table>
<thead>
<tr>
<th></th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATH ($n=24$)</td>
<td>CON ($n=26$)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>30.0 ± 5.5</td>
<td>27.4 ± 4.9</td>
</tr>
<tr>
<td>Body mass (kg)</td>
<td>57.0 ± 5.9$^a$</td>
<td>61.6 ± 6.6</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>168 ± 5</td>
<td>168 ± 7</td>
</tr>
<tr>
<td>BMI (kg · m$^{-2}$)</td>
<td>20.3 ± 1.6$^b$</td>
<td>21.8 ± 1.6</td>
</tr>
<tr>
<td>Body fat (%)</td>
<td>23.3 ± 4.8$^a$</td>
<td>26.0 ± 4.8</td>
</tr>
<tr>
<td>Fat-free mass (kg)</td>
<td>44.4 ± 4.3</td>
<td>45.7 ± 5.2</td>
</tr>
</tbody>
</table>

RMR
- kcal · d$^{-1}$: 1436 ± 121
- kcal · kg$^{-1}$ · d$^{-1}$: 25.4 ± 2.5
- kcal · kg$^{-1}$ · h$^{-1}$: 1.06 ± 0.10

VO$_{2}$max
- L · min$^{-1}$: 3.2 ± 0.3$^a$
- ml · kg$^{-1}$ · min$^{-1}$: 56.1 ± 4.5$^c$

Maximum metabolic equivalent of task
- 15.2 ± 1.4$^c$

HRmax (bpm)
- 181 ± 9$^a$

SHR (bpm)
- 49 ± 7$^a$

Recovery HR (bpm)
- 87 ± 14

PAL
- 2.1 ± 0.3$^b$

Legend: Results are expressed as mean ± standard deviation.$^a$ significantly different from CON of the same sex ($p < 0.05$),$^b$ significantly different from CON of the same sex ($p < 0.01$),$^c$ significantly different from CON of the same sex ($p < 0.0001$)

Table 2: Concurrent validity (tested mode vs. criterion measure) of the two Actiheart modes to estimate VO$_{2}$max in men and women athletes (ATH) and controls (CON).

<table>
<thead>
<tr>
<th></th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATH ($n=24$)</td>
<td>CON ($n=26$)</td>
</tr>
<tr>
<td>$r$ value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHraw</td>
<td>0.225</td>
<td>0.463$^a$</td>
</tr>
<tr>
<td>AHcomplete</td>
<td>0.410$^a$</td>
<td>0.488$^a$</td>
</tr>
<tr>
<td>MAE [ml · kg$^{-1}$ · min$^{-1}$]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHraw</td>
<td>8.8 ± 9.2</td>
<td>2.5 ± 7.2</td>
</tr>
<tr>
<td>AHcomplete</td>
<td>3.0 ± 6.5</td>
<td>-1.2 ± 8.0</td>
</tr>
<tr>
<td>MAPE [%]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHraw</td>
<td>17.7 ± 13.7</td>
<td>11.4 ± 10.4</td>
</tr>
<tr>
<td>AHcomplete</td>
<td>10.8 ± 6.5</td>
<td>13.6 ± 8.9</td>
</tr>
<tr>
<td>SEE [ml · kg$^{-1}$ · min$^{-1}$]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHraw</td>
<td>4.47</td>
<td>3.98</td>
</tr>
<tr>
<td>AHcomplete</td>
<td>4.19</td>
<td>3.91</td>
</tr>
</tbody>
</table>

Legend: Results are expressed as mean ± standard deviation.$^a$ correlations significant at $p < 0.05$
ues was detected \((p < 0.05)\). The relationship between VO\(_2\)\(_{\text{max}}\) estimated by AHcomplete and the criterion method was significant in women ATH and CON and in men CON \((p < 0.05)\).

In Figure 1, the data of AHraw and AHcomplete and the reference method are presented using Bland-Altman plots. For the total sample, the mean bias \((± 1.96 \text{SD})\) was 4.5 \((-13.6; 22.6)\) ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) for AHraw, whereas for the AHcomplete no systematic errors were observed. Likewise, no systematic errors were found in the single groups for both AHraw and AHcomplete, except for women ATH when VO\(_2\)\(_{\text{max}}\) was estimated by AHraw \((p < 0.0001; \text{Figure 1})\).

**Discussion**

We evaluated the validity of the Actiheart step test for the estimation of VO\(_2\)\(_{\text{max}}\) in men and women along a range of aerobic fitness levels, comparing two different data entry modes, AHraw (estimated SHR, HRmax and RMR) and AHcomplete (measured SHR, HRmax and RMR). Based on MAPE, the Actiheart step test is not an acceptable method for estimation of VO\(_2\)\(_{\text{max}}\).

There exist several submaximal step tests for the estimation of VO\(_2\)\(_{\text{max}}\). The most common are the Chester step test, the STEP tool protocol, the modified YMCA 3-minute step test, and the Åstrand-Rhyming step test (Bennett, Parfitt, Davison, & Eston, 2016). The \(R^2\) of previous reported step tests for the estimation of VO\(_2\)\(_{\text{max}}\) ranged between 0.22 and 0.88 (Francis & Brasher, 1992; Francis & Culpepper, 1989; Hansen, Jacobs, Thijs, Dendale, & Claes, 2016; Knight, Stuckey, & Petrella, 2014; McArdle, Katch, Peciar, Jacobson, & Ruck, 1972; Perroni, Cortis, Minganti, Cignitti, & Castronaca, 2013; Petrella, Koval, Cunningham, & Paterson, 2001; Santo & Golding, 2003; Sykes & Roberts, 2004; Webb, Vehrs, George, & Hager, 2014). In our study we found a significant relationship between estimated and measured VO\(_2\)\(_{\text{max}}\) for the total sample, with \(R^2\) 0.24 for AHraw and 0.36 for AHcomplete. However, when dividing the groups by sex and experimental group (ATH and CON), the \(R^2\) ranged from 0.01 to 0.22 for AHraw (with being significant only for women CON), and 0.06 to 0.24 for AHcomplete (significant for women ATH and CON, and men CON).

Compared to other studies the absolute measures of agreement showed similar or slightly better validity of the Actiheart step test for the estimation of VO\(_2\)\(_{\text{max}}\) (SEE: 3.98 – 6.31 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) for AHraw and 3.91 – 6.15 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) for AHcomplete). Previous studies reported a SEE of 6.9 – 8.76 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) for the modified YMCA 3-minute step test in healthy men and women (Santo & Golding, 2003) and 3.9 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) for the Chester step test in healthy adults (Sykes & Roberts, 2004). In well-trained men a SEE of 0.28 L \(\cdot\) min\(^{-1}\) was detected for the Åstrand-Rhyming step test (Åstrand & Rhyming, 1954). Assuming a body mass of 70 kg (not reported in that study) this would amount to a SEE of 3.6 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\).

In the present study, we found a systematic bias of 4.5 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) for the total sample for AHraw, indicating that estimated VO\(_2\)\(_{\text{max}}\) was significantly lower compared to the measured VO\(_2\)\(_{\text{max}}\) values. When VO\(_2\)\(_{\text{max}}\) was estimated by AHcomplete, no systematic errors were observed \((-1.1\ \text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}\)\), non-significant). Knight et al. (2014) observed a systematic bias of -6.4 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) for the STEP tool protocol in 40 healthy men and women, with higher VO\(_2\)\(_{\text{max}}\) values in the predictive vs. maximal test. For the Chester step test a mean systematic bias of 2.8 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) was reported in university students,

### Table 3: Parameters of linear regression for maximum oxygen consumption (VO\(_2\)\(_{\text{max}}\); ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\)) in men and women athletes (ATH) and controls (CON). Measured VO\(_2\)\(_{\text{max}}\) was entered as dependent variable and estimated VO\(_2\)\(_{\text{max}}\) by Actiheart was the independent variable.

<table>
<thead>
<tr>
<th></th>
<th>Women (n=24)</th>
<th>Men (n=22)</th>
<th>Women (n=26)</th>
<th>Men (n=19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R^2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHraw</td>
<td>0.051</td>
<td>0.005</td>
<td>0.168</td>
<td>0.230</td>
</tr>
<tr>
<td>AHcomplete</td>
<td>0.168</td>
<td>0.055</td>
<td>0.047</td>
<td>0.038</td>
</tr>
<tr>
<td>(p)-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHraw</td>
<td>0.291</td>
<td>0.760</td>
<td>0.291</td>
<td>0.065</td>
</tr>
<tr>
<td>AHcomplete</td>
<td>0.047</td>
<td>0.293</td>
<td>0.11</td>
<td>0.216</td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHraw</td>
<td>0.251</td>
<td>-0.057</td>
<td>0.234</td>
<td>0.217</td>
</tr>
<tr>
<td>AHcomplete</td>
<td>0.266</td>
<td>0.153</td>
<td>-0.057</td>
<td>0.216</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHraw</td>
<td>50.914</td>
<td>67.394</td>
<td>36.881</td>
<td>44.632</td>
</tr>
<tr>
<td>AHcomplete</td>
<td>42.004</td>
<td>54.011</td>
<td>36.782</td>
<td>43.183</td>
</tr>
</tbody>
</table>

Legend: Significant \(p\)-values are highlighted in bold.
### Figure 1: Bland-Altman plots of AHraw and AHcomplete and the reference method (VO$_2$ maxmeas) for men and women athletes and controls.
indicating that predicted VO$_{\text{max}}$ was significantly lower compared to the measured VO$_{\text{max}}$ during a maximal treadmill test (Buckley, Sim, Eston, Hession, & Fox, 2004). We also found a significant negative relationship between the mean of measured and estimated VO$_{\text{max}}$ and the difference between both methods, indicating that with increasing VO$_{\text{max}}$ the difference between actual and estimated VO$_{\text{max}}$ increased ($r = -0.33$ for AHraw and $r = -0.45$ for AHcomplete, $p < 0.01$). This means that for participants with very high (such as endurance athletes) or very low aerobic capacity (untrained) the difference between estimated VO$_{\text{max}}$ by the Actiheart step test and measured VO$_{\text{max}}$ is larger. In participants with high aerobic capacity the Actiheart step test thus overestimates actual VO$_{\text{max}}$, whereas in participants with low aerobic capacity VO$_{\text{max}}$ is underestimated, independent of which entry mode is used.

In our study, the absolute difference between measured and estimated VO$_{\text{max}}$ was smaller than 5 ml ∙ kg$^{-1}$ ∙ min$^{-1}$ in 41.8% of the participants for both AHraw and AHcomplete. However, when VO$_{\text{max}}$ was estimated by use of AHraw, only 28.3% ($n = 13$) of the athletes had a smaller than 5 ml ∙ kg$^{-1}$ ∙ min$^{-1}$ absolute difference between measured and estimated VO$_{\text{max}}$, whereas when VO$_{\text{max}}$ was estimated by AHcomplete 45.7% ($n = 21$) of the participants were within 5 ml ∙ kg$^{-1}$ ∙ min$^{-1}$. These results indicate that in endurance trained participants the AHcomplete for estimating VO$_{\text{max}}$ is more accurate than the AHraw, whereas in persons with a good VO$_{\text{max}}$ the AHraw setting is sufficient to predict aerobic capacity.

We assume that the underlying equations of the manufacturer were originally developed for the estimation of VO$_{\text{max}}$ of the general population with a limited range of aerobic capacity and not especially for endurance athletes nor for persons with low aerobic capacity. Based on our results it would thus seem necessary to adapt these equations in order to provide more valid VO$_{\text{max}}$ values for a broader range of aerobic capacities.

**Strengths and limitations**

A strength of the study is the number of included men and women covering the upper range of aerobic capacities. In addition, body composition, RMR, and long-term Actiheart data (such as SHR and PAL) were available for all participants and could be entered into the Actiheart software in order to evaluate the effect on the estimation of VO$_{\text{max}}$. We also put effort into measuring "real" VO$_{\text{max}}$ (and not VO$_{\text{peak}}$) by applying current criteria for maximal exercise testing (Scharhag-Rosenberger 2010, Midgley et al. 2008), including the recommendation that the duration of VO$_{\text{max}}$ tests should last between 5 and 26 minutes (Midgley et al., 2008). The specific maximal exercise test in our study was performed on a motorized treadmill and the chosen protocol for the determination of VO$_{\text{max}}$ was previously shown to induce exhaustion in athletic and non-athletic persons after 5 - 9 min (Steiner & Wehrlin, 2011).

A common issue in studies involving maximal exercise testing is the selection bias, often including fitter individuals than the general population. In a recent study examining the aerobic capacity in the Swiss working population VO$_{\text{max}}$ values of 33 and 45 ml · kg$^{-1}$ · min$^{-1}$ for women and men were reported (Mundwiler et al., 2017). In our control group, the mean values for VO$_{\text{max}}$ were 48.4 and 56.6 ml · kg$^{-1}$ · min$^{-1}$ for women and men, indicating a higher fitness level in the included participants compared to the general population.

Step tests, in general, are prone to error. For example, performing a step test requires the person’s ability to maintain a certain stepping tempo and technique. Alterations in stepping technique can affect mechanical efficiency and therefore physiological responses in HR and oxygen consumption (Bennett et al., 2016). In addition, performing a step test with a pre-defined fixed step height, might introduce potential error since leg length, step length, body mass, and morphology vary between individuals and these differences might result in different individual physiological responses.

The mode AHcomplete requires the measurement of RMR, a long-term assessment of SHR, and a maximum exercise test for determination of HRmax. These measurements have higher financial and time costs, and require the same or even more expertise compared to the direct assessment of VO$_{\text{max}}$. However, HRmax can be also measured using a maximum exercise test without performing a spiroergometry, which makes its application possible in settings where the equipment for a respiratory gas analysis is not available. Another limitation of the AHcomplete mode is that performing a maximum exercise test (with or without respiratory gas analysis) is not possible in every population, which limits the application of this mode. In summary, the mode AHcomplete is only applicable in settings without respiratory gas analysis equipment and in subjects who can perform a maximum exercise test for determination of HRmax. We did not verify the validity of an Actiheart mode where RMR and SHR are entered as measured variables, and HRmax is estimated. However, using an estimated HRmax as endpoint for the estimation of VO$_{\text{max}}$, such as for AHraw, might introduce error since it was shown that estimations of HRmax can have a considerable prediction error (Tanaka et al., 2001). Another potential source of error is the reliance on the linear relationship between VO$_{\text{max}}$ and power output, as shown by Åstrand and Rodahl (1970). However, Zoladz, Rademaker and Sargeant (1995) found a non-linear relationship between VO$_{\text{max}}$ and power output, which may affect the prediction of VO$_{\text{max}}$ by use of submaximal tests. Although the non-linearity might occur predominantly at high intensities above the anaerobic threshold (Majerczak et al., 2012; Zoladz, Szukutnik, Majerczak, & Duda, 1998) this underlines the importance of choosing an individual work rate which is not too high. In addition, it was shown that there is not always a linear relationship between HR and VO$_{\text{max}}$ (Buckley et al., 2004).
Conclusion

Based on MAPE, in young adults with good to superior aerobic capacity the Actiheart step test was not acceptable for estimation of $VO_{2,max}$. The mode AHraw significantly underestimated $VO_{2,max}$ in men and women endurance trained athletes, and AHcomplete significantly underestimated $VO_{2,max}$ in women athletes. In endurance trained participants one should manually enter RMR, HRmax, and SHR into the Actiheart software in order to increase the accuracy of the $VO_{2,max}$ prediction. Areas of future investigation include the repetition of analyses of this study when performing the Actiheart step test with other step heights, evaluation of the reliability of the step test, inclusion of older and more sedentary persons and use in a clinical setting, and assessing factors such as HR and $VO_{2}$ during the Actiheart step test to improve the estimation of $VO_{2,max}$.

Acknowledgements

The authors would like to thank all the participants who volunteered to participate in this study.

Funding

The authors have no funding or support to report.

Competing Interests

The authors have declared that no competing interests exist.

Data Availability Statement

All relevant data are included in the paper. Further data can be made available upon request.

Abbreviations

ATH = athletes, BMI = body mass index, CON = controls, HR = heart rate, HR$_{max}$ = maximum heart rate, MAE = mean absolute error, MAPE = mean absolute percentage error, PAL = physical activity level, RMR = resting metabolic rate, RQ = respiratory quotient, SEE = standard error of the estimate, SHR = sleeping heart rate, TEE = total energy expenditure, VCO$_2$ = expired carbon dioxide, $VO_2$ = oxygen consumption, $VO_2$max = maximum oxygen consumption

References


