

# Validity of the revised Ekblom Bak cycle ergometer test in adults

Frida Björkman<sup>1</sup>  · Elin Ekblom-Bak<sup>1</sup> · Örjan Ekblom<sup>1</sup> · Björn Ekblom<sup>1</sup>

Received: 5 March 2016 / Accepted: 5 June 2016 / Published online: 16 June 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

## Abstract

**Purpose** To further develop the Ekblom Bak-test prediction equation for estimation of  $\text{VO}_2\text{max}$  from submaximal cycle ergometry.

**Methods** The model group (117 men and 100 women, aged  $48.3 \pm 15.7$  and  $46.1 \pm 16.8$  years,  $\text{VO}_2\text{max}$   $46.6 \pm 11.1$  and  $40.4 \pm 9.6$   $\text{mL kg}^{-1} \text{min}^{-1}$ , respectively) and the cross-validation group (60 men and 55 women, aged  $40.6 \pm 17.1$  and  $41.6 \pm 16.7$  years,  $\text{VO}_2\text{max}$   $49.0 \pm 12.1$  and  $43.2 \pm 8.9$   $\text{mL min}^{-1} \text{kg}^{-1}$ , respectively) performed 4 min of cycling on a standard work rate (30 W) directly followed by 4 min on a higher work rate. Heart rate (HR) at each work rate was recorded. Thereafter, participants completed a graded maximal treadmill test for direct measurement of oxygen uptake. The new prediction equation was cross-validated and accuracy compared with the original Ekblom Bak equation as well as by the Åstrand test method.

**Results** The final sex-specific regression models included age, change in HR per-unit change in power ( $\Delta\text{HR}/\Delta\text{PO}$ ), the difference in work rates ( $\Delta\text{PO}$ ), and HR at standard work rate as independent variables. The adjusted  $R^2$  for the final models were 0.86 in men and 0.83 in women. The coefficient of variation (CV) was 8.7 % and SEE  $0.28 \text{ L min}^{-1}$ . The corresponding CV and SEE values for the EB-test<sub>2012</sub> and the Åstrand tests were 10.9 and 18.1 % and 0.35 and  $0.48 \text{ L min}^{-1}$ , respectively.

**Conclusion** The new EB-test prediction equation provides an easy administered and valid estimation of  $\text{VO}_2\text{max}$  for a wide variety of ages (20–86 years) and fitness levels ( $19\text{--}76 \text{ mL kg}^{-1} \text{min}^{-1}$ ).

**Keywords** Cardiorespiratory fitness · Estimation · Public health · Oxygen uptake · Submaximal test · Validity

## Abbreviations

bpm	Beats per minute
EB-test <sub>2012</sub>	The Ekblom-Bak test with the original prediction equation (Epub 2012 Nov 6)
EB-test <sub>new</sub>	The Ekblom Bak-test with the new prediction equation
HR	Heart rate
RPE	Rate of perceived exertion
rpm	Revolutions per minute
$\text{VO}_2$	Oxygen uptake
$\text{VO}_2\text{max}$	Maximal oxygen uptake
$\Delta\text{HR}$	The difference in HR between the high and low work rate
$\Delta\text{PO}$	The difference in work rate between the high and low work rate

## Introduction

Cardiorespiratory fitness, assessed as maximal oxygen consumption ( $\text{VO}_2\text{max}$ ), is a key factor in physical performance (Bassett and Howley 2000) and a strong independent predictor of health and longevity (Blair et al. 1989; Kodama et al. 2009).  $\text{VO}_2\text{max}$  is determined during maximal physical effort by indirect calorimetry using special laboratory equipment. The methodology is time consuming, expensive, and dependent on physiological expertise.

Communicated by Jean-René Lacour.

✉ Frida Björkman  
frida.bjorkman@gih.se

<sup>1</sup> Åstrand Laboratory of Work Physiology, The Swedish School of Sport and Health Sciences, Box 5626, 114 86 Stockholm, Sweden

Furthermore, the  $\text{VO}_2\text{max}$  measurements require a maximal effort from an individual, which is not always suitable in the general non-athlete population. Even if the maximal exercise test per se is a relatively safe procedure, there are also a number of contraindications that limits the possibility to evaluate  $\text{VO}_2\text{max}$  in many settings (American Thoracic and American College of Chest 2003). To enable some evaluation of cardiorespiratory fitness in a population, in which the direct determination of  $\text{VO}_2\text{max}$  is not possible, several tests, such as different step tests (Brouha et al. 1943; McArdle et al. 1972; Bennett et al. 2015), walk tests (Kline et al. 1987; Solway et al. 2001), treadmill (Swank et al. 2001), and cycle ergometer tests (Åstrand and Ryhming 1954; Legge and Bannister 1986; Golding et al. 1989; Hartung et al. 1993; Beekley et al. 2004), have been developed to estimate  $\text{VO}_2\text{max}$  from submaximal performance.

In 2012, a submaximal cycle ergometer test (the original Ekblom-Bak test, EB-test<sub>2012</sub>) was presented, which estimates  $\text{VO}_2\text{max}$  based on sex, age, and heart rate (HR) difference between a standard, low standard work rate, and a high individually chosen work rate (Ekblom-Bak et al. 2014). In a mixed population with regard to sex, age, and physical activity status, there was a strong correlation between estimated and actually measured  $\text{VO}_2\text{max}$ ,  $r = 0.91$ , with a corresponding coefficient of variation (CV) of 9.3 %. This was a significantly improved precision level compared with one of the most commonly used submaximal cycle ergometer tests and the Åstrand test (Åstrand 1960), reported by the authors to have a CV of 15 %. This is similar to other validation reports (Jessup et al. 1977; Siconolfi et al. 1982; Ekblom et al. 2007) and to the values found in our study population in the 2012 publication ( $r = 0.68$  and CV 18.1 %) (Ekblom-Bak et al. 2014).

Although the relatively high precision of estimating  $\text{VO}_2\text{max}$  by the EB-test<sub>2012</sub>, there was a strong, significant correlation between the error of measured and estimated  $\text{VO}_2\text{max}$  and absolute  $\text{VO}_2\text{max}$  level (spearman  $\rho = 0.42$  in the total population,  $\rho = 0.51$  in women, and  $\rho = 0.81$  in men). This means that individuals with high  $\text{VO}_2\text{max}$  were underestimated and individuals with low  $\text{VO}_2\text{max}$  overestimated with the EB-test<sub>2012</sub>. Moreover, during the development of the EB-test<sub>2012</sub> prediction equation, we found that the size of the study population limited the possibility of sex-specific equations, something that probably would enhance the prediction equation as sex was included as independent variable in the prediction equation. In addition, by including a sample having a greater range of  $\text{VO}_2\text{max}$  (currently for the EB-test<sub>2012</sub> prediction equation 1.56–3.73 L min<sup>-1</sup> in women and 2.75–4.49 L min<sup>-1</sup> in men) and age (currently 21–65 years), the test would be applicable on a greater proportion of the general population.

Therefore, in the present paper, we aimed to further develop the EB-test<sub>2012</sub> prediction equation by including

additional participants to the study population on which the 2012 prediction equation was developed. This enabled us to develop sex-specific prediction equations and to expand the valid age and  $\text{VO}_2\text{max}$  range for the test. By identifying individual characteristics and physiological responses associated to  $\text{VO}_2\text{max}$ , we also aimed to reduce the estimation bias seen with high  $\text{VO}_2\text{max}$  level in the original test. The new prediction equation (EB-test<sub>new</sub>) was then validated internally in the model group and in an external cross-validation sample and was also compared with the Åstrand and the EB-test<sub>2012</sub>.

## Materials and methods

### Participants

After public announcement and word-of-mouth in the region of Stockholm, Sweden, we included 74 additional participants to the 2012 model group (Ekblom-Bak et al. 2014). This additional sample consisted of 52 men and 22 women, mean age 58.8 (20–86) years, with a mean  $\text{VO}_2\text{max}$  of 3.10 (1.33–5.97) L min<sup>-1</sup> and 39.2 (18.0–76.4) mL kg<sup>-1</sup> min<sup>-1</sup>. Inclusion criteria for this group were an age, and  $\text{VO}_2\text{max}$  mainly identified to be outside the valid range for the 2012 prediction equation (Ekblom-Bak et al. 2014). These participants were thereafter pooled to the model group of the 2012 prediction equation. The pooled sample was used to create the new prediction equation for the EB-test. The same recruitment method as for the participants to the model group was applied for the cross-validation sample. Recruitment of the latter sample consisted of a mixture of men and women of different ages and fitness levels, to enable cross validation in both the total sample and in subgroups. Inclusion criteria were men and women above 20 years of age, with no known diseases or disabilities. The characteristic of both the new model group ( $n = 217$ ) as well as the cross-validation sample ( $n = 115$ ) is presented in Table 1.

All participants were free from all the types of diseases that limit the physical work capacity and stated themselves as healthy on the test day. Furthermore, they were not taking any medications that could influence the relationship between HR and  $\text{VO}_2$ . Exclusion criteria were smoking, snuff use, and medication with beta blockers or asthmatic medicine. Participants visited the test laboratory on one occasion to perform the submaximal EB-test and a maximal treadmill test to assess actual  $\text{VO}_2\text{max}$ . Before the visit, the participants were asked to refrain from smoking and vigorous physical activity the day before and on the test day, and to not consume a heavy meal less than 3 h before the test. All participants were fully informed about the details of the study and provided written consent.

**Table 1** Characteristics of the study samples (means  $\pm$  standard deviations)

	Model group for the EB-test <sub>new</sub> equation		Cross-validation group for the EB-test <sub>new</sub> equation		Comparison group <sup>a</sup>	
	Men ( <i>n</i> = 117)	Women ( <i>n</i> = 100)	Men ( <i>n</i> = 60)	Women ( <i>n</i> = 55)	Men ( <i>n</i> = 27)	Women ( <i>n</i> = 44)
Age (years)	48.3 (15.7)	46.1 (16.8)	40.6 (17.1)	41.6 (16.7)	36.0 (12.8)	36.9 (13.7)
Height (cm)	180.4 (6.9)	166.2 (6.2)	180.4 (6.0)	167.3 (6.9)	179.0 (5.6)	167.2 (7.2)
Body Mass (kg)	80.7 (9.0)	63.7 (8.5)	82.5 (15.4)	63.9 (8.8)	81.8 (17.2)	63.5 (7.9)
VO <sub>2</sub> max (L min <sup>-1</sup> )	3.73 (0.86)	2.55 (0.58)	3.95 (0.89)	2.75 (0.59)	3.70 (0.55)	2.85 (0.53)
VO <sub>2</sub> max (mL kg <sup>-1</sup> min <sup>-1</sup> )	46.6 (11.1)	40.4 (9.6)	49.0 (12.1)	43.2 (8.9)	46.4 (8.4)	45.1 (7.9)
HR <sub>max</sub> (beats min <sup>-1</sup> )	178 (17)	179 (12)	185 (15)	182 (12)	192 (11)	185 (11)

<sup>a</sup> Including those participants in the cross-validation group with a valid EB-test<sub>2012</sub> and Åstrand test

### Submaximal and maximal test

All tests were performed in climate-controlled laboratory environment. When arriving to the test centre, body mass (measured while wearing in light-weight clothes to the nearest 0.1 kg) and height (to the nearest 0.1 cm) were measured. The participants were informed about the test procedure and equipped with a HR monitor (Polar Electro, Kempele, Finland). After individual adjustments of the seat and handlebar of the cycle ergometer and an introduction of the Borg's scale of perceived exertion (RPE) (Borg 1970), the participant performed an EB-test according to the original 2012 test procedure (Ekblom-Bak et al. 2014). The test was performed on a mechanically braked cycle ergometer (Monark model 828E, Varberg, Sweden). Test procedure included 4 min of cycling on a standard and low work rate of 0.5 kilopond (kp) with a pedal frequency of 60 rpm ( $\approx 30$  W when 1 W = 6.116 kpm/min), directly followed by 4 min of cycling on a higher individually chosen work rate (aiming at a RPE of  $\approx 14$  on the Borg scale). Mean steady-state HR during the last minute on the low and high work rates, respectively, was recorded by taking the mean of the observed HR at 3:15, 3:30, 3:45, and 4:00 min at each work rate. In addition, VO<sub>2</sub>max was also estimated by the Åstrand test method by applying the work rate and HR of the high work rate to the Åstrand nomogram (Åstrand and Ryhming 1954) and associated age-correction factors (Åstrand 1960). The same way of obtaining Åstrand test results from the EB-test procedure was used in the original publication of the first EB-test prediction equation, and is further described and discussed in the previous article (Ekblom-Bak et al. 2014). Direct measurement of VO<sub>2</sub> during the submaximal cycle test was conducted in a subsample (*n* = 110) in the model group, using a computerised metabolic system (Jaeger Oxycon pro, Hoechberg, Germany) connected to a face mask worn by the participant. Before each test, ambient temperature, humidity, and barometric pressure were measured with built-in automatic procedures and a handheld instrument (HygroPalm, Rotronic,

Bassersdorf, Schweiz). Gas analyzers and inspiratory flow-meter were calibrated with the metabolic system's built-in automatic procedures, where high-precision calibration gases (15.00  $\pm$  0.01 % O<sub>2</sub> and 6.00  $\pm$  0.01 % CO<sub>2</sub>, Air Liquid, Kungsängen, Sweden), and ambient indoor air was used for the gas analyses.

After a short rest, a 5 min warm-up on the treadmill preceded a graded maximal treadmill test to measure VO<sub>2</sub>max. The individually designed protocol for the VO<sub>2</sub>max test started off at 1° incline and a velocity corresponding to approximately 60–65 % of the participant's estimated VO<sub>2</sub>max (usually the speed that the participant felt comfortable with during the warm-up). The speed increased 1 km/h during the first 3 to 4 min of the test, and thereafter, there was an increase in incline with +1° every minute until voluntary exhaustion. For some of the well-trained participants, running to an incline of 5°–6°, there was an additional increase in speed (+1 km h<sup>-1</sup> per minute) to avoid too steep inclination on the treadmill. Direct measurements of VO<sub>2</sub> were obtained during the test with the same computerised system as mentioned above (Jaeger Oxycon pro). Criteria for acceptance of the VO<sub>2</sub>max measurement were levelling off of VO<sub>2</sub> despite an increase in speed or incline, a respiratory exchange ratio >1.1, RPE above 16, work time above 6 min, supported by a maximal HR within  $\pm 15$  beats min<sup>-1</sup> (bpm) from age-predicted maximal HR (ref Åstrand Rodahl). A test was accepted as VO<sub>2</sub>max when a minimum of three out of the five criteria was achieved. In the model group, nine participants were tested but later excluded due to non-fulfilling the requirements for acceptance of test (five participants failed the VO<sub>2</sub>max test and four participants had non-valid EB test). The corresponding values in the cross-validation group were four excluded participants in total, two with non-valid VO<sub>2</sub>max test and two with non-valid EB-test.

VO<sub>2</sub>max (L min<sup>-1</sup>) and maximal HR (bpm) were recorded into 30 and 5 s epochs, respectively. We have previously shown that there is no mean difference and a small variation (CV: 2.7 %) between test–retest of VO<sub>2</sub>max

according to the above procedure in a mixed population (Ekblom-Bak et al. 2014), indicating no need for a second  $\text{VO}_2\text{max}$  test on a separate test day to verify the first accepted measurement.

### Development of the new EB-test prediction equation

In the EB-test<sub>2012</sub> prediction equation,  $\text{VO}_2\text{max}$  is predicted by entering sex (women = 0, men = 1) and age (years), and the difference in HR between the high and low work rate ( $\Delta\text{HR}$ ) divided by the difference in work rate between the high and low work rate ( $\Delta\text{PO}$ ) into the equation  $\text{VO}_2\text{max} = 4.98196 - 2.88618 (\Delta\text{HR}/\Delta\text{PO}) + 0.65015 (\text{sex}) - 0.01712 (\text{age})$ . For the development of  $\Delta\text{HR}/\Delta\text{PO}$ , see the paper of the EB-test<sub>2012</sub> (Ekblom-Bak et al. 2014). In line with the development of the EB-test<sub>2012</sub> equation, but now based on the extended model group population, linear regression modelling was used to develop the EB-test<sub>new</sub> prediction equation. Apart from performing separate prediction equations for men and women, the natural logarithm of  $\text{VO}_2\text{max}$  ( $\ln \text{VO}_2\text{max}$ ) was used for a better fit and for avoiding extrapolation to infinity in both ends of the regression. Since both the  $\Delta\text{PO}$  variable (representing the high work rate) and the steady-state HR obtained on the low, the standard work rate was identified as highly associated with  $\text{VO}_2\text{max}$  in both men and women in the EB-test<sub>2012</sub> equation (for  $\Delta\text{PO}$ ,  $r = 0.76$  in women and  $0.79$  in men; for HR at the standard work rate,  $r = -0.57$  in women and  $r = -0.26$  in men;  $p < 0.001$  for all); we chose to include these as separate variables in the new model. The model construction is described in detail in the statistical analysis section below.

### Validation, cross validation, and comparison groups

The EB-test<sub>new</sub> prediction equation presented in the results section was internally validated in the model group as well as in the external cross-validation sample. Moreover, the validity and precision of the estimated  $\text{VO}_2\text{max}$  by the EB-test<sub>new</sub> was also compared with both the EB-test<sub>2012</sub> and estimated  $\text{VO}_2\text{max}$  derived from the Åstrand test method. This group is called the comparison group and consisted of only participants from the cross-validation group. Since both the EB-test<sub>2012</sub> and the Åstrand test method have narrower valid age and  $\text{VO}_2\text{max}$  ranges compared with the new prediction equation, only those fulfilling the prerequisites (i.e., participants aged 20–65 years with an HR on the high work rate applicable to the Åstrand nomogram, namely, 120–170 bpm) for all three tests were included in these comparison analyses (see Table 1 for the characteristics of participants in the validation group, cross-validation group, and comparison group).

$\text{VO}_2\text{max}$  estimated by the EB-test<sub>2012</sub> was obtained by entering the corresponding values for each participant into the equation presented above. Results are also reported for quartiles in fitness level for absolute ( $\text{L min}^{-1}$ ) as well as relative ( $\text{mL kg}^{-1} \text{min}^{-1}$ )  $\text{VO}_2\text{max}$ .

### Statistical analysis

A multiple linear regression with the forward method (probability of  $F = 0.05$  for entry, and  $0.10$  for removal) was used, identifying age,  $\Delta\text{HR}/\Delta\text{PO}$ ,  $\Delta\text{PO}$ , and HR at the low standard work rate to be included in the final model as independent predictors of  $\ln \text{VO}_2\text{max}$ . Since we found significant gender-HR and gender- $\Delta\text{HR}/\Delta\text{PO}$  interaction effects (both  $p < 0.001$ ) for the prediction of  $\ln \text{VO}_2\text{max}$ , regressions were performed for men and women separately. The final model was checked for homoscedasticity and equal variance. Tolerance values  $>0.2$  for all independent variables indicated low probability of multicollinearity. The 95 % confidence intervals (95 % CI) for the regression coefficients are given. Paired  $t$  tests were used to study the difference between measured and estimated  $\text{VO}_2\text{max}$ . Bland–Altman plots, including limits of agreement (LoA), were produced for the cross-validation sample. The coefficient of variation variable (Tables 2, 3, 4) was calculated by dividing the standard deviation (SD) of the difference between measured and estimated  $\text{VO}_2\text{max}$  with the mean of the measured and estimated  $\text{VO}_2\text{max}$ . The adjusted  $R^2$  and standard errors of the estimate (SEE) were obtained by linear regression, entering measured  $\text{VO}_2\text{max}$  as the dependent variable and calculated  $\text{VO}_2\text{max}$  as the independent variable. Two-sided statistical significance was set to  $p < 0.05$  for all analyses. The SPSS statistical software version 21.0 (SPSS Inc., Chicago, IL, USA) was used for the statistical analyses.

## Results

Subject characteristics in the model group, cross-validation group, and the comparison groups are shown in Table 1. The age and  $\text{VO}_2\text{max}$  ranges for the participants in the model group were 21–86 years and  $1.33$ – $3.94 \text{ L min}^{-1}$  ( $18.9$ – $61.9 \text{ mL kg}^{-1} \text{min}^{-1}$ ) in women, respectively, and 20–84 years and  $1.67$ – $5.97 \text{ L min}^{-1}$  ( $23.5$ – $76.4 \text{ mL kg}^{-1} \text{min}^{-1}$ ) in men, respectively.

### New equation based on data from the model group

The final sex-specific regression models (with 95 % CI for the independent variables) for EB-test<sub>new</sub> were

Men:  $\ln \text{VO}_2\text{max} = 2.04900$  (95 % CI  $1.83517$ – $2.26282$ )  $- 0.00858$  (95 % CI  $-0.00987$  to  $0.00728$ )

**Table 2** Validity of the EB-test<sub>new</sub> equation (in the model group)

	Measured VO <sub>2</sub> max (L min <sup>-1</sup> )	Estimated VO <sub>2</sub> max (L min <sup>-1</sup> )	Difference measured vs estimated VO <sub>2</sub> max (L min <sup>-1</sup> )	Coefficient of variation (%)	R <sup>2</sup> adjusted	SEE (L min <sup>-1</sup> )
	Mean (SD)	Mean (SD)	Mean (95 % CI)			
All (n = 217)	3.18 (0.95)	3.17 (0.92)	-0.01 (-0.05 to 0.03)	8.7	0.91	0.28
Men (n = 117)	3.73 (0.86)	3.72 (0.83)	-0.01 (-0.07 to 0.05)	8.4	0.87	0.31
Women (n = 100)	2.55 (0.58)	2.53 (0.52)	-0.01 (-0.06 to 0.03)	9.2	0.84	0.24
Age						
<35 (n = 59)	3.57 (0.91)	3.63 (0.91)	0.06 (-0.02 to 0.13)	7.9	0.90	0.28
35–49 (n = 61)	3.65 (0.86)	3.59 (0.83)	-0.06 (-0.12 to 0.01)	6.8	0.92	0.25
50–64 (n = 49)	3.03 (0.71)	2.94 (0.64)	-0.09 (-0.18 to -0.01)	10.3	0.81	0.31
≥65 (n = 48)	2.27 (0.55)	2.32 (0.51)	0.05 (-0.03 to 0.12)	10.9	0.80	0.25
Absolute VO <sub>2</sub> max level (according to L min <sup>-1</sup> )						
Q1 (n = 55)	2.24 (0.53)	2.34 (0.53)	0.10 (0.04 to 0.16)	9.9	0.82	0.22
Q2 (n = 54)	2.95 (0.58)	2.95 (0.60)	-0.00 (-0.08 to 0.09)	11.0	0.72	0.31
Q3 (n = 54)	3.47 (0.66)	3.46 (0.73)	-0.01 (-0.07 to 0.05)	6.5	0.91	0.20
Q4 (n = 54)	4.09 (0.84)	3.95 (0.87)	-0.14 (-0.21 to -0.06)	6.9	0.90	0.27
Relative VO <sub>2</sub> max level (according to mL kg <sup>-1</sup> min <sup>-1</sup> )						
Q1 (n = 54)	2.27 (0.57)	2.36 (0.57)	0.09 (0.02 to 0.16)	10.5	0.82	0.24
Q2 (n = 54)	3.02 (0.62)	2.96 (0.63)	-0.06 (-0.14 to 0.01)	9.3	0.81	0.27
Q3 (n = 55)	3.48 (0.78)	3.44 (0.73)	-0.03 (-0.12 to 0.05)	8.8	0.85	0.31
Q4 (n = 54)	3.96 (0.87)	3.92 (0.90)	-0.04 (-0.11 to 0.03)	6.6	0.92	0.25

All measured and estimated values are in expressed as L min<sup>-1</sup>

Quartiles (L min<sup>-1</sup>), women; <2.14, 2.14–2.59, 2.59–2.93, >2.93; Men; <3.07, 3.07–3.75, 3.75–4.33, >4.33

Quartiles (mL kg<sup>-1</sup> min<sup>-1</sup>), women; <32.5, 32.5–40.2, 40.2–47.4, >47.4; men; <37.5, 37.5–47.6, 47.6–54.8, >54.8

SD standard deviation, SEE standard error of the estimate, Q quartile

(age) - 0.90742 (95 % CI -1.11676 to -0.69808) (ΔHR/ΔPO) + 0.00178 (95 % CI 0.00127–0.00228) (ΔPO) - 0.00290 (95 % CI -0.00438 to -0.00141) (HR at standard work rate).

Women: ln VO<sub>2</sub>max = 1.84390 (95 % CI 1.53151–2.15628) - 0.00673 (95 % CI -0.00812 to 0.00534) (age) - 0.62578 (95 % CI -0.81368 to -0.43789) (ΔHR/ΔPO) + 0.00175 (95 % CI 0.00056–0.00295) (ΔPO) - 0.00471 (95 % CI -0.00674 to -0.00268) (HR at standard work rate).

After entering the corresponding values into the equation, VO<sub>2</sub>max (in L min<sup>-1</sup>) was estimated by putting in the obtained value (x) as an exponent in the natural logarithm. R<sup>2</sup> adjusted for the final models were 0.86 in men and 0.83 in women.

Table 2 presents the validity of the new equation in the model group. The systematic error (the difference between measured and estimated VO<sub>2</sub>max) was assessed in the full sample, as well as in subgroups for age and fitness level (presented as the quartiles for relative as well as absolute VO<sub>2</sub>max). The systematic error in the different subgroups ranges from an underestimation at the most of

0.14 L min<sup>-1</sup> to an overestimation of 0.10 L min<sup>-1</sup>. The coefficient of variation was 8.7 % in the full sample, ranging between 6.5 and 11.0 % in the subgroups. In addition, the explained variance of the measured VO<sub>2</sub>max by the estimated value was 91 % in the full sample, and the SEE values were 0.28 L min<sup>-1</sup> for the full sample and ranging between 0.20 and 0.31 L min<sup>-1</sup> in the subgroups.

The differences between measured and estimated VO<sub>2</sub>max by the EB-test<sub>new</sub> equation in the model group were not associated with maximal HR in women (Spearman ρ = 0.10, p = 0.33) or men (ρ = 0.14, p = 0.13), and not with deviation for age-predicted maximal HR (ρ = -0.17, p = 0.08) in women, but in men (ρ = -0.33, p < 0.001). In both women and men, the differences were associated with VO<sub>2</sub>max level, ρ = 0.40, p < 0.001, and ρ = 0.25, p = 0.006, respectively.

**Cross validation**

In the cross-validation sample, the estimation of VO<sub>2</sub>max by the EB-test<sub>new</sub> equation was analysed in the full sample and thereafter stratified into the same subgroups as

**Table 3** Cross validation of the EB-test<sub>new</sub> equation (in the external cross-validation group)

	Measured VO <sub>2</sub> max (L min <sup>-1</sup> )	Estimated VO <sub>2</sub> max (L min <sup>-1</sup> )	Difference measured vs estimated VO <sub>2</sub> max (L min <sup>-1</sup> )	Coefficient of variation (%)	R <sup>2</sup> adjusted	SEE (L min <sup>-1</sup> )
	Mean (SD)	Mean (SD)	Mean (95 % CI)			
All ( <i>n</i> = 115)	3.37 (0.97)	3.39 (1.02)	0.02 (−0.04 to 0.08)	9.4	0.90	0.30
Men ( <i>n</i> = 60)	3.95 (0.89)	4.06 (0.89)	0.11 (0.02 to 0.20)	8.3	0.86	0.33
Women ( <i>n</i> = 55)	2.75 (0.59)	2.66 (0.54)	−0.09 (−0.16 to −0.01)	10.0	0.79	0.27
Age						
<35 ( <i>n</i> = 55)	3.75 (0.77)	3.77 (0.90)	0.01 (−0.08 to 0.11)	9.1	0.86	0.29
35–49 ( <i>n</i> = 28)	3.53 (1.07)	3.49 (1.07)	−0.04 (−0.15 to 0.08)	8.5	0.92	0.30
50–64 ( <i>n</i> = 15)	2.87 (0.81)	2.95 (0.89)	0.08 (−0.11 to 0.27)	11.8	0.84	0.33
≥65 ( <i>n</i> = 17)	2.34 (0.49)	2.39 (0.56)	0.06 (−0.07 to 0.19)	10.9	0.77	0.23
Absolute VO <sub>2</sub> max level (according to L min <sup>-1</sup> )						
Q1 ( <i>n</i> = 24)	2.28 (0.47)	2.44 (0.59)	0.16 (0.03 to 0.29)	12.8	0.73	0.24
Q2 ( <i>n</i> = 21)	2.96 (0.51)	3.02 (0.71)	0.06 (−0.07 to 0.18)	9.3	0.90	0.16
Q3 ( <i>n</i> = 23)	3.41 (0.67)	3.38 (0.80)	−0.03 (−0.14 to 0.09)	8.0	0.89	0.22
Q4 ( <i>n</i> = 47)	4.10 (0.81)	4.05 (0.96)	−0.05 (−0.15 to 0.05)	8.5	0.88	0.29
Relative VO <sub>2</sub> max level (according to mL kg <sup>-1</sup> min <sup>-1</sup> )						
Q1 ( <i>n</i> = 23)	2.58 (0.70)	2.68 (0.66)	0.10 (−0.01 to 0.21)	9.3	0.87	0.25
Q2 ( <i>n</i> = 19)	3.00 (0.82)	3.02 (0.91)	0.03 (−0.10 to 0.15)	8.8	0.91	0.24
Q3 ( <i>n</i> = 31)	3.32 (0.80)	3.36 (0.93)	0.04 (−0.10 to 0.19)	11.9	0.82	0.34
Q4 ( <i>n</i> = 42)	4.02 (0.86)	3.97 (0.99)	−0.05 (−0.15 to 0.04)	7.7	0.91	0.26

All measured and estimated values are in expressed as L min<sup>-1</sup>

Quartiles (L min<sup>-1</sup>), women; <2.14, 2.14–2.59, 2.59–2.93, >2.93; men; <3.07, 3.07–3.75, 3.75–4.33, >4.33

Quartiles (mL kg<sup>-1</sup> min<sup>-1</sup>), women; <32.5, 32.5–40.2, 40.2–47.4, >47.4; men; <37.5, 37.5–47.6, 47.6–54.8, >54.8

SD standard deviation, SEE standard error of the estimate, Q quartile

the model group. Data were found to be homoscedastic. Results are shown in Table 3. The cross-validation analyses showed similar results as in the internal validation sample, with a non-significant mean systematic difference of 0.02 (95 % CI −0.04 to 0.08) and a similar variation, 9.4 %, in the full sample, ranging from −0.09 to 0.16 L min<sup>-1</sup> and 7.7 % to 12.8 %, respectively, in the different subgroups. Bland–Altman plots of the estimated and measured VO<sub>2</sub>max (in L min<sup>-1</sup> and mL kg<sup>-1</sup> min<sup>-1</sup>) are given in Figs. 1 and 2, respectively. LoA for absolute values were −0.54 to 0.76 L min<sup>-1</sup> in men and −0.61 to 0.44 L min<sup>-1</sup> in women. The corresponding values for relative VO<sub>2</sub>max were −7.2 to 10.3 mL kg<sup>-1</sup> min<sup>-1</sup> in men, and −9.7 to 7.1 mL kg<sup>-1</sup> min<sup>-1</sup> in women, respectively. SEE for the absolute values were 0.33 L min<sup>-1</sup> for men and 0.27 min<sup>-1</sup> for woman, and the corresponding relative values were 4.17 and 4.26 mL kg<sup>-1</sup> min<sup>-1</sup>, respectively.

### Comparison group

There were overall larger systematic errors and variations in estimations of VO<sub>2</sub>max by the EB-test<sub>2012</sub> equation and

the Åstrand method, compared with the EB-test<sub>new</sub> equation (Table 4). Mean error was significantly lower for the EB-test<sub>new</sub> compared with the EB-test<sub>2012</sub> ( $p < 0.001$ ). The adjusted R<sup>2</sup> between estimated and measured values was significantly higher in EB-test<sub>new</sub> compared with the Åstrand test. Compared with the new test equation, the larger coefficient of variation and SEE and lower R<sup>2</sup> adjusted indicated a larger variability in the EB-test<sub>2012</sub> and Åstrand test, respectively. In the full sample and in most subgroups, the difference between measured and estimated VO<sub>2</sub>max was lower for the new test equation compared with the other two tests.

In addition, difference between measured and estimated VO<sub>2</sub>max for the EB-test<sub>new</sub> and the EB-test<sub>2012</sub> equation was correlated with measured VO<sub>2</sub>max level in the comparison group. The associations between estimation error and VO<sub>2</sub>max level for the EB-test<sub>2012</sub> were  $\rho = 0.37$  in the full sample, and  $\rho = 0.44$  and  $\rho = 0.83$  (all  $p < 0.01$ ) for women and men, respectively. For the EB-test<sub>new</sub>, these correlations were lower and only significant in woman; the corresponding values were  $\rho = 0.074$  ( $p = 0.53$ ) in the full sample, and  $\rho = 0.43$  ( $p = 0.004$ ) and  $\rho = 0.36$  ( $p = 0.069$ ) in women and men, respectively.

**Table 4** Comparison of the three tests (in the comparison group)

	Measured $\dot{V}O_2\text{max}$ ( $\text{L}\cdot\text{min}^{-1}$ )		Estimated $\dot{V}O_2\text{max}$ ( $\text{L}\cdot\text{min}^{-1}$ )		Difference measured vs. estimated $\dot{V}O_2\text{max}$ ( $\text{L}\cdot\text{min}^{-1}$ )				
	Mean (SD)	EB-test <sub>new</sub> Mean (SD)	EB-test <sub>2012</sub> Mean (SD)	$\dot{A}\text{strand}$ Mean (SD)	EB-test <sub>new</sub> Mean (95 % CI)		EB-test <sub>2012</sub> Mean (95 % CI)		$\dot{A}\text{strand}$ Mean (95 % CI)
					Mean (SD)	Mean (SD)	Mean (95 % CI)	Mean (95 % CI)	
All ( $n = 71$ )	3.17 (0.67)	3.14 (0.71)	3.30 (0.66)	3.10 (0.79)	-0.03 (-0.10 to 0.04)	0.12 (0.04 to 0.21)	-0.07 (-0.21 to 0.06)		
Men ( $n = 27$ )	3.70 (0.55)	3.79 (0.53)	3.89 (0.33)	3.29 (0.75)	0.09 (-0.03 to 0.21)	0.20 (0.05 to 0.34)	-0.41 (-0.61 to -0.20)		
Women ( $n = 44$ )	2.85 (0.53)	2.74 (0.47)	2.93 (0.54)	2.98 (0.80)	-0.11 (-0.19 to -0.03)	0.08 (-0.02 to 0.18)	0.13 (-0.02 to 0.28)		
Age									
<35 ( $n = 41$ )	3.41 (0.50)	3.37 (0.56)	3.54 (0.43)	3.37 (0.66)	-0.04 (-0.13 to 0.06)	0.14 (0.03 to 0.24)	-0.04 (-0.23 to 0.14)		
35–49 ( $n = 17$ )	2.97 (0.67)	2.91 (0.72)	3.06 (0.56)	2.81 (0.80)	-0.06 (-0.21 to 0.09)	0.10 (-0.04 to 0.24)	-0.15 (-0.48 to 0.17)		
50–64 ( $n = 12$ )	2.75 (0.86)	2.77 (0.89)	2.86 (1.01)	2.69 (0.85)	0.02 (-0.19 to 0.23)	0.11 (-0.23 to 0.45)	-0.06 (-0.36 to 0.24)		
Absolute $\dot{V}O_2\text{max}$ level (according to $\text{L}\cdot\text{min}^{-1}$ )									
Q1 ( $n = 10$ )	2.19 (0.43)	2.36 (0.55)	2.52 (0.74)	2.24 (0.55)	0.16 (-0.05 to 0.37)	0.32 (0.00 to 0.64)	0.05 (-0.33 to 0.42)		
Q2 ( $n = 15$ )	3.00 (0.47)	3.08 (0.68)	3.37 (0.62)	2.67 (0.62)	0.08 (-0.08 to 0.24)	0.37 (0.21 to 0.53)	-0.33 (-0.62 to -0.05)		
Q3 ( $n = 23$ )	3.41 (0.67)	3.38 (0.80)	3.49 (0.69)	3.22 (0.67)	-0.03 (-0.14 to 0.09)	0.08 (-0.07 to 0.23)	-0.18 (-0.39 to 0.02)		
Q4 ( $n = 23$ )	3.47 (0.41)	3.27 (0.44)	3.40 (0.35)	3.63 (0.62)	-0.20 (-0.31 to -0.10)	-0.08 (-0.16 to 0.01)	0.15 (-0.12 to 0.43)		
Relative $\dot{V}O_2\text{max}$ level (according to $\text{mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ )									
Q1 ( $n = 11$ )	2.84 (0.79)	2.93 (0.72)	3.25 (0.77)	2.61 (0.55)	0.09 (-0.07 to 0.25)	0.40 (0.25 to 0.56)	-0.24 (-0.73 to 0.25)		
Q2 ( $n = 12$ )	2.94 (0.82)	2.98 (0.89)	3.08 (0.96)	2.84 (0.87)	0.03 (-0.14 to 0.21)	0.14 (-0.10 to 0.38)	-0.10 (-0.37 to 0.16)		
Q3 ( $n = 25$ )	3.21 (0.62)	3.20 (0.79)	3.34 (0.65)	3.01 (0.75)	-0.01 (-0.16 to 0.14)	0.13 (-0.02 to 0.28)	-0.19 (-0.42 to 0.03)		
Q4 ( $n = 23$ )	3.41 (0.50)	3.25 (0.50)	3.38 (0.40)	3.56 (0.67)	-0.15 (-0.25 to -0.06)	-0.03 (-0.16 to 0.11)	0.15 (-0.08 to 0.39)		

Table 4 continued

	Coefficient of variation				R <sup>2</sup> adjusted				SEE			
	EB-test <sub>new</sub> (%)		EB-test <sub>2012</sub> (%)		Åstrand (%)		EB-test <sub>new</sub>		EB-test <sub>2012</sub>		Åstrand	
	EB-test <sub>new</sub> (%)	EB-test <sub>2012</sub> (%)	Åstrand (%)	EB-test <sub>new</sub>	EB-test <sub>2012</sub>	Åstrand	EB-test <sub>new</sub>	EB-test <sub>2012</sub>	Åstrand			
All (n = 71)	9.4	10.9	18.1	0.83	0.74	0.50	0.28	0.35	0.48			
Men (n = 27)	8.4	9.7	14.8	0.68	0.56	0.50	0.31	0.36	0.39			
Women (n = 44)	9.3	11.8	17.2	0.75	0.62	0.61	0.26	0.32	0.33			
Age												
<35 (n = 41)	8.8	9.5	17.3	0.71	0.57	0.25	0.27	0.33	0.44			
35–49 (n = 17)	9.9	9.1	21.7	0.83	0.82	0.38	0.28	0.28	0.52			
50–64 (n = 12)	11.9	19.0	17.2	0.85	0.70	0.69	0.33	0.47	0.47			
Absolute VO <sub>2</sub> max level (according to L·min <sup>-1</sup> )												
Q1 (n = 10)	12.8	18.8	23.9	0.69	0.67	0.08	0.24	0.24	0.41			
Q2 (n = 15)	9.7	9.2	18.2	0.86	0.78	0.29	0.17	0.22	0.39			
Q3 (n = 23)	8.0	10.1	14.5	0.89	0.75	0.53	0.22	0.34	0.46			
Q4 (n = 23)	7.3	5.7	17.7	0.69	0.76	0.06	0.23	0.20	0.40			
Relative VO <sub>2</sub> max level (according to mL·kg <sup>-1</sup> ·min <sup>-1</sup> )												
Q1 (n = 11)	8.2	7.8	26.9	0.91	0.90	0.12	0.25	0.25	0.75			
Q2 (n = 12)	9.3	12.5	14.6	0.90	0.84	0.75	0.27	0.33	0.41			
Q3 (n = 25)	11.3	10.9	17.4	0.79	0.70	0.47	0.28	0.34	0.45			
Q4 (n = 23)	6.6	9.2	15.5	0.81	0.59	0.34	0.22	0.32	0.41			

All measured and estimated values are in expressed as L·min<sup>-1</sup>

Quartiles (L·min<sup>-1</sup>), women: <2.14, 2.14–2.59, 2.59–2.93, >2.93; men: <3.07, 3.07–3.75, 3.75–4.33, >4.33

Quartiles (mL·kg<sup>-1</sup>·min<sup>-1</sup>), women: <32.5, 32.5–40.2, 40.2–47.4, >47.4; men: <37.5, 37.5–47.6, 47.6–54.8, >54.8

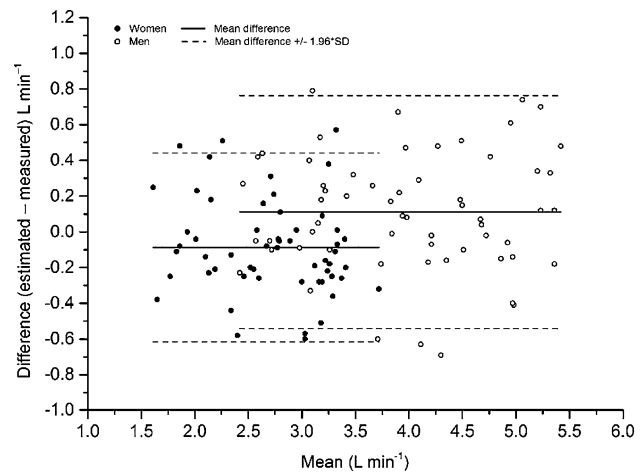
SD standard deviation, SEE standard error of the estimate, Q quartile



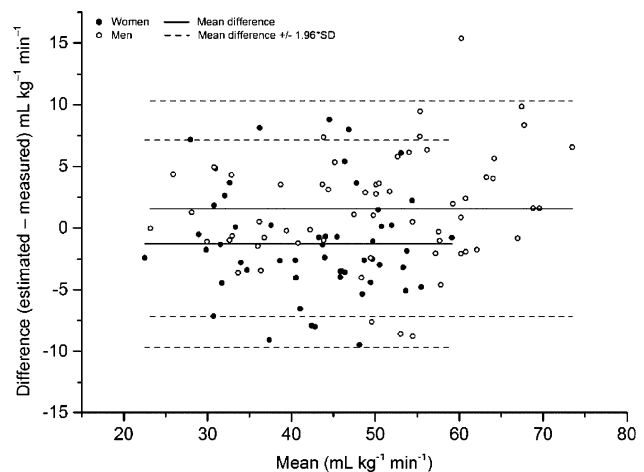
## Discussion

In the present paper, we have further developed a test, the EB-test presented in 2012, for the estimation of  $\text{VO}_2\text{max}$  from submaximal work rates on cycle ergometers by including sex-specific prediction equations and expanding the test's validity in respect to both  $\text{VO}_2\text{max}$  and age range. Compared with the internal validation for the EB-test<sub>2012</sub> (9.3 % and  $0.30 \text{ L min}^{-1}$ ), the variation between estimated and measured  $\text{VO}_2\text{max}$  and corresponding SEE was lower for the EB-test<sub>new</sub>; 8.7 % and  $0.28 \text{ L min}^{-1}$ . Cross validation in an external sample showed similar improved results. Subgroup analyses in both the internal validation sample and the cross-validation sample showed consistency in accuracy and precision between sexes, age-groups, and quartiles of  $\text{VO}_2\text{max}$ . The correlation between the estimation error and  $\text{VO}_2\text{max}$  level seen for the EB-test<sub>2012</sub> was lower for the EB-test<sub>new</sub> in the full sample and especially among men. This reduction may be partly explained by the inclusion of two new variables in the prediction equation,  $\Delta\text{PO}$  variable, and HR at standard work rate.

The objective of the EB-test<sub>new</sub>, just as with the EB-test<sub>2012</sub>, is to provide a simple, time-effective and low-risk way to enable estimation of  $\text{VO}_2\text{max}$  in settings outside the physiological laboratory. With the increased valid range for the test, it is now applicable on healthy individuals with wide ranging physical performance capacity. Another advantage of the EB-test<sub>new</sub> is the further improved precision of the estimated  $\text{VO}_2\text{max}$  compared with measured  $\text{VO}_2\text{max}$ , with low variation between sexes and different age and  $\text{VO}_2\text{max}$  groups (see Table 4). For example, by applying the variation between measured and estimated  $\text{VO}_2\text{max}$  reported in the model group (8.4 % for men and 9.2 % for women), 95 out of 100 male individuals performing the EB-test<sub>new</sub> with an actual  $\text{VO}_2\text{max}$  of  $3.0 \text{ L min}^{-1}$  will be predicted within  $\pm 0.49 \text{ L min}^{-1}$  and females within  $\pm 0.54 \text{ L min}^{-1}$ . For a male or female with an absolute  $\text{VO}_2\text{max}$  of  $2.0 \text{ L min}^{-1}$ , the corresponding estimated  $\text{VO}_2\text{max}$  is within 0.33 and  $0.36 \text{ L min}^{-1}$ , respectively. However, these results from the EB-test<sub>new</sub> rely on that the data for calculation are obtained in line with the description for test method and with the correct equipment (i.e., cycle ergometer), as described earlier. The same test manual and protocol should be applied for the EB-test<sub>new</sub> as for the EB-test<sub>2012</sub> (Ekblom-Bak et al. 2014), which is already in use. It is only the prediction equation that has been updated. The new equation is sex specific and with a wider age and  $\text{VO}_2\text{max}$  ranges, with the inclusion of two additional test variables (the  $\Delta\text{PO}$  variable and HR at the standard work rate) to reduce the estimation error and increase the precision of the estimation of  $\text{VO}_2\text{max}$ .



**Fig. 1** Bland Altman plot, including limits of agreement, for estimated and measured  $\text{VO}_2\text{max}$  ( $\text{L min}^{-1}$ ) in the cross-validation group. Filled dots women. Transparent dots men. Black line mean difference. Dashed line  $\pm 1.96 \times \text{SD}$



**Fig. 2** Bland Altman plot, including limits of agreement, for estimated and measured  $\text{VO}_2\text{max}$  ( $\text{mL kg}^{-1} \text{ min}^{-1}$ ) in the cross-validation group. Filled dots women. Transparent dots men. Black line mean difference. Dashed line  $\pm 1.96 \times \text{SD}$

There are a number of already existing submaximal tests for estimation of  $\text{VO}_2\text{max}$ , where the participant performs an amount of submaximal work in the form of a step-up exercise, walking, running, or pedalling on a cycle ergometer (Brouha et al. 1943; Åstrand and Ryhming 1954; McArdle et al. 1972; Legge and Bannister 1986; Kline et al. 1987; Golding et al. 1989; Hartung et al. 1993; Swank et al. 2001; Solway et al. 2001; Bennett et al. 2015). Compared with previously described methods, the EB-test<sub>new</sub> test has higher  $R^2$  than many step tests, which range from 0.22 to 0.90 (Perroni et al. 2013; Åstrand and Ryhming 1954; Santo and Golding 2003; McArdle et al. 1972; Francis and

Culpepper 1989; Knight et al. 2014; Chatterjee et al. 2004). Compared with other studies on the validity of the Åstrand test, the EB-test<sub>new</sub> test has similar  $R^2$  of 0.90 (Hartung et al. 1993). Validity is slightly higher for maximal treadmill running tests for the determination of  $\text{VO}_2\text{max}$ , e.g., Balke test protocol,  $R^2$  0.85 or Bruce test protocol,  $R^2$  0.77 (Pollock et al. 1976). However, maximal tests include an all-out performance, in which  $\text{VO}_2\text{max}$  is only one part. Furthermore, maximal running tests are not applicable in many situations, such as when testing older people, patients with orthopedic diagnoses, obesity or people unaccustomed to intense running. Furthermore, relative estimates or agreement, such as the correlation coefficient, are highly dependent on absolute range. In this present paper, we have the same absolute range in  $\text{VO}_2\text{max}$ , making  $R^2$  comparisons meaningful. In the comparison group, the adjusted  $R^2$  for EB-test<sub>new</sub> was 0.83, and the corresponding value for the Åstrand test method was 0.50.

Absolute measures of agreement indicate similar or slightly better validity for the EB-test<sub>new</sub> (SEE: 4.2 mL kg<sup>-1</sup> min<sup>-1</sup> for men and 4.3 mL kg<sup>-1</sup> min<sup>-1</sup> for women, respectively) compared with other tests. The previous studies report an SEE of 8.9 mL kg<sup>-1</sup> min<sup>-1</sup> for the submaximal YMCA cycle ergometer test (Beekley et al. 2004) and for the Åstrand test method SEE from 4.3 mL kg<sup>-1</sup> min<sup>-1</sup> (Hartung et al. 1993) to 5.7 mL kg<sup>-1</sup> min<sup>-1</sup> (Cink and Thomas 1981). In the comparison group in the present paper, SEE for relative values based on the Åstrand test method was 5.6 mL kg<sup>-1</sup> min<sup>-1</sup>, compared with the EB-test<sub>new</sub> 4.1 mL kg<sup>-1</sup> min<sup>-1</sup> (data not shown).

The commonly found low accuracy and precision with the run- walk and step tests could be due to the inter-individual variance in morphology, body mass, gait, and mechanical efficiency, which may be induced when individuals perform these types of weight-bearing activities. The use of the cycle ergometers for submaximal testing diminishes some of the mechanical variations in performance, as the previous research has shown rather constant mechanical efficiency in mixed populations (Åstrand and Rodahl 1970).

Another factor influencing the accuracy of submaximal tests is the use of age-predicted maximal HR in the calculations, a postulation where there is large individual variations (Engels et al. 1998). For example, the Åstrand test uses a one-point work rate methodology combined with the above-mentioned assumption regarding maximal HR, which partly explains some of the measurement error (in the present study population  $r = 0.71$  and CV 18.1 %, respectively). In the present test, there are no calculations of maximal HR; therefore, this potential source of error is eliminated. We believe that the inclusion of the  $\Delta\text{PO}$  variable and HR at the low standard work rate in the EB-test<sub>new</sub>

equation explains some of the reduced variation in the low and high ends of the  $\text{VO}_2\text{max}$  range, resulting in a higher precision throughout the valid  $\text{VO}_2\text{max}$  span.

Although showing better accuracy and precision compared with the EB-test<sub>2012</sub>, there is still an estimation error left in the EB-test<sub>new</sub> equation with increased  $\text{VO}_2\text{max}$  level, albeit non-significant. One possible source of error could be the usage of the same and rather low standard work rate ( $\approx 30$  W) in a population with wide ranging physical performance capacity. This rate of work may be too low for obtaining full stroke volume and thereby contributes to a variation in HR response at the standard work rate, especially in highly trained subjects (Blomqvist and Saltin 1983). To analyse this, we examined the correlation of percentage of  $\text{VO}_2\text{max}$  on the standard work rate and measured  $\text{VO}_2$  during submaximal cycling ( $n = 110$ ) and the estimation error. The correlation was  $\rho = 0.18$  ( $p = 0.061$ ). The percentage of  $\text{VO}_2\text{max}$  on the standard work rate was therefore ruled out as a main source of the estimation error. However, other factors, such as blood flow distribution and blood pressure, related to high  $\text{VO}_2\text{max}$  level, may influence the estimation error.

Another potential source of error that may influence on the precision of the submaximal test is the reliance on the existence of a linear relationship between  $\text{VO}_2$  and power output, as previously shown by Åstrand and Rodahl (Åstrand and Rodahl 1970). This notion has been questioned due to later findings of a non-linear relation between  $\text{VO}_2$  and power output (Zoladz et al. 1995). This non-linearity may affect the ability of a submaximal test to predict  $\text{VO}_2\text{max}$ . However, the non-linearity of this relationship seems to be found at high intensities above the anaerobic threshold (Zoladz et al. 1998; Majerczak et al. 2012). Nonetheless, the issue of non-linearity may well be an important factor for the proper execution of the test, stressing the importance of not choosing an individual work rate that is too high.

## Strengths and limitations

A strength of the EB-test<sub>new</sub> equation is the consistent high accuracy and precision in the external cross-validation sample, as well as in the different subgroups (men and women, different age-groups and a wide range of  $\text{VO}_2\text{max}$  levels). This is of particular interest with regard to the fact that the new equation is based on a relatively large and heterogeneous sample, with a wide variation in age and  $\text{VO}_2\text{max}$ . We, therefore, believe that the test now is suitable for most non-diseased individuals. The use of sex specific prediction equations, rather than controlling for sex within an equation, has undoubtedly contributed to better precision.

A common issue in studies involving maximal testing is the selection bias, often including more fit individuals than

in the general population. This may limit the accuracy of the prediction equation in the general population. One way to analyse the representativeness of the present model group for the general population (with special regard to age), is to compare the decrease in  $\text{VO}_2\text{max}$  with age between the model group and previously reported values from general population samples. The age-related decline per decade in measured  $\text{VO}_2\text{max}$  for the model group sample in this study was  $-9.1\%$  per 10-year, compared with  $-6.5\%$  in 10,973 men and women, where  $\text{VO}_2\text{max}$  was obtained from maximal testing on cycle ergometer (Eriksen et al. 2015) and  $-6.9\%$  in 3678 men and women, where  $\text{VO}_2\text{max}$  was obtained from treadmill testing (Loe et al. 2013). The difference in slope may be due to a higher mean fitness among the young participants in the present paper, and may thus express a limited degree of bias. The implication of this bias on the validity of the test may be regarded as low.

Moreover, a submaximal test only estimates  $\text{VO}_2\text{max}$  based on variables obtained during the submaximal exercise. Hence, individuals who deviate in physiological characteristics from the individuals included in the model group, for example with an extremely high or low cardiorespiratory fitness, exceptional work efficiency or abnormal HR response, may obtain an estimated  $\text{VO}_2\text{max}$  further from their actual  $\text{VO}_2\text{max}$  than expected. Furthermore, any medications that may have an influence on HR, also may affect the results from the test. This prediction equation is based on  $\Delta\text{HR}$ , and the  $\Delta\text{HR}$  relation to power output, whereas medications, such as beta block, may alter these relationships. Future research should focus on the feasibility of the EB test in clinical populations. Another scope of interest is the ability of the EB test to detect a change in actual  $\text{VO}_2\text{max}$  over time, for example as a consequence of a training intervention. The use of the test in this situation has not yet been evaluated. To date, the recommendation is to keep the same high individually chosen work rate when monitoring an individual over time.

## Conclusion

In the present paper, we have further developed an easy administered, non-expensive, and accurate submaximal ergometer test for the estimation of  $\text{VO}_2\text{max}$ . The EB-test<sub>new</sub> estimates  $\text{VO}_2\text{max}$  throughout a wide range of ages and fitness levels, and can be used in health screenings and in research studies in large populations and in the general population.

**Acknowledgements** Huge thanks to all study participants and assisting co-workers at the Åstrand Laboratory of Work Physiology and The Swedish School of Sport and Health Sciences. The work was performed in memory of the recently deceased P.-O. Åstrand.

## Compliance with ethical standards

All procedures performed in the study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study design was approved by the Regional Ethics Committee in Stockholm (Ref. Number 2013/1634-32).

**Conflict of interest** All authors declare that they do not have any conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- American Thoracic S, American College of Chest P (2003) ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 167(2):211–277. doi:10.1164/rccm.167.2.211
- Åstrand I (1960) Aerobic work capacity in men and women with special reference to age. *Acta Physiol Scand Suppl* 49(169):1–92
- Åstrand PO, Rodahl K (1970) Textbook of work physiology. McGraw-Hill, New York
- Åstrand PO, Ryhming I (1954) A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during sub-maximal work. *J Appl Physiol* 7(2):218–221
- Bassett DR Jr, Howley ET (2000) Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med Sci Sports Exerc* 32(1):70–84
- Beekley MD, Brechue WF, deHoyos DV, Garzarella L, Werber-Zion G, Pollock ML (2004) Cross-validation of the YMCA submaximal cycle ergometer test to predict  $\text{VO}_2\text{max}$ . *Res Q Exerc Sport* 75(3):337–342. doi:10.1080/02701367.2004.10609165
- Bennett H, Parfitt G, Davison K, Eston R (2015) Validity of submaximal step tests to estimate maximal oxygen uptake in healthy adults. *Sports Med*. doi:10.1007/s40279-015-0445-1
- Blair SN, Kohl HW 3rd, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW (1989) Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA* 262(17):2395–2401
- Blomqvist CG, Saltin B (1983) Cardiovascular adaptations to physical training. *Annu Rev Physiol* 45:169–189. doi:10.1146/annurev.ph.45.030183.001125
- Borg G (1970) Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med* 2(2):92–98
- Brouha L, Graybiel A, Heath CW (1943) The step test: a simple method of measuring maximal physical fitness for hard muscular work in man. *Rev Canad Biol* 2:86–92
- Chatterjee S, Chatterjee P, Mukherjee PS, Bandyopadhyay A (2004) Validity of Queen's College step test for use with young Indian men. *Br J Sports Med* 38(3):289–291
- Cink RE, Thomas TR (1981) Validity of the Åstrand–Ryhmings nomogram for predicting maximal oxygen intake. *Br J Sports Med* 15(3):182–185
- Ekblom B, Engstrom LM, Ekblom O (2007) Secular trends of physical fitness in Swedish adults. *Scand J Med Sci Sports* 17(3):267–273

- Ekblom-Bak E, Bjorkman F, Hellenius ML, Ekblom B (2014) A new submaximal cycle ergometer test for prediction of  $\text{VO}_2\text{max}$ . *Scand J Med Sci Sports* 24(2):319–326. doi:[10.1111/sms.12014](https://doi.org/10.1111/sms.12014)
- Engels HJ, Zhu W, Moffatt RJ (1998) An empirical evaluation of the prediction of maximal heart rate. *Res Q Exerc Sport* 69(1):94–98. doi:[10.1080/02701367.1998.10607673](https://doi.org/10.1080/02701367.1998.10607673)
- Eriksen L, Gronbaek M, Helge JW, Tolstrup JS (2015) Cardiorespiratory fitness in 16 025 adults aged 18–91 years and associations with physical activity and sitting time. *Scand J Med Sci Sports*. doi:[10.1111/sms.12608](https://doi.org/10.1111/sms.12608)
- Francis K, Culpepper M (1989) Height-adjusted, rate-specific, single-stage step test for predicting maximal oxygen consumption. *South Med J* 82(5):602–606
- Golding L, Myers CR, Sinning WE (1989) Y's way to physical fitness, 3rd edn. Human Kinetics, Champaign
- Hartung GH, Krock LP, Crandall CG, Bisson RU, Myhre LG (1993) Prediction of maximal oxygen uptake from submaximal exercise testing in aerobically fit and nonfit men. *Aviat Space Environ Med* 64(8):735–740
- Jessup GT, Riggs CE, Lambert J, Miller WD (1977) The effect of pedalling speed on the validity of the Astrand–Ryhming aerobic work capacity test. *J Sports Med Phys Fit* 17(4):367–371
- Kline GM, Porcari JP, Hintermeister R, Freedson PS, Ward A, McCarron RF, Ross J, Rippe JM (1987) Estimation of  $\text{VO}_2\text{max}$  from a one-mile track walk, gender, age, and body weight. *Med Sci Sports Exerc* 19(3):253–259
- Knight E, Stuckey MI, Petrella RJ (2014) Validation of the step test and exercise prescription tool for adults. *Can J Diabetes* 38(3):164–171. doi:[10.1016/j.jcjd.2014.03.007](https://doi.org/10.1016/j.jcjd.2014.03.007)
- Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, Sugawara A, Totsuka K, Shimano H, Ohashi Y, Yamada N, Sone H (2009) Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *JAMA* 301(19):2024–2035. doi:[10.1001/jama.2009.681](https://doi.org/10.1001/jama.2009.681)
- Legge BJ, Bannister EW (1986) The Astrand–Ryhming nomogram revisited. *J Appl Physiol* 61:1203–1209
- Loe H, Rognmo O, Saltin B, Wisloff U (2013) Aerobic capacity reference data in 3816 healthy men and women 20–90 years. *PLoS One* 8(5):e64319. doi:[10.1371/journal.pone.0064319](https://doi.org/10.1371/journal.pone.0064319)
- Majerczak J, Korostynski M, Nieckarz Z, Szkutnik Z, Duda K, Zoladz JA (2012) Endurance training decreases the non-linearity in the oxygen uptake-power output relationship in humans. *Exp Physiol* 97(3):386–399. doi:[10.1113/expphysiol.2011.062992](https://doi.org/10.1113/expphysiol.2011.062992)
- McArdle WD, Katch FI, Pechar GS, Jacobson L, Ruck S (1972) Reliability and interrelationships between maximal oxygen intake, physical work capacity and step-test scores in college women. *Med Sci Sports* 4(4):182–186
- Perroni F, Cortis C, Minganti C (2013) Maximal oxygen uptake of Italian firefighters: laboratory vs. field evaluations. *Sport Sci Health* 9(2):31–35. doi:[10.1007/s11332-013-0142-0](https://doi.org/10.1007/s11332-013-0142-0)
- Pollock ML, Bohannon RL, Cooper KH, Ayres JJ, Ward A, White SR, Linnerud AC (1976) A comparative analysis of four protocols for maximal treadmill stress testing. *Am Heart J* 92(1):39–46
- Santo AS, Golding LA (2003) Predicting maximum oxygen uptake from a modified 3-minute step test. *Res Q Exerc Sport* 74(1):110–115. doi:[10.1080/02701367.2003.10609070](https://doi.org/10.1080/02701367.2003.10609070)
- Siconolfi SF, Cullinane EM, Carleton RA, Thompson PD (1982) Assessing  $\text{VO}_2\text{max}$  in epidemiologic studies: modification of the Astrand–Ryhming test. *Med Sci Sports Exerc* 14(5):335–338
- Solway S, Brooks D, Lacasse Y, Thomas S (2001) A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. *Chest* 119(1):256–270
- Swank AM, Serapiglia L, Funk D, Adams KJ, Durham M, Berning JM (2001) Development of a branching submaximal treadmill test for predicting  $\text{VO}_2\text{max}$ . *J Strength Cond Res/Natl Strength Cond Assoc* 15(3):302–308
- Zoladz JA, Rademaker AC, Sargeant AJ (1995) Non-linear relationship between  $\text{O}_2$  uptake and power output at high intensities of exercise in humans. *J Physiol* 488(Pt 1):211–217
- Zoladz JA, Szkutnik Z, Majerczak J, Duda K (1998) Detection of the change point in oxygen uptake during an incremental exercise test using recursive residuals: relationship to the plasma lactate accumulation and blood acid base balance. *Eur J Appl Physiol* 78(4):369–377