

## Improving reference equations for cardiorespiratory fitness using multiplicative allometric rather than additive linear models: Data from the Fitness Registry and the Importance of Exercise National Database Registry

Item Type	Journal article
Authors	Nevill, Alan M;Myers, Jonathan;Kaminsky, Leonard A;Arena, Ross
Citation	Nevill, A. M., Myers, J., Kaminsky, L. A. and Arena, R. (2019) Improving reference equations for cardiorespiratory fitness using multiplicative allometric rather than additive linear models: Data from the Fitness Registry and the Importance of Exercise National Database Registry, Progress in Cardiovascular Diseases, <a href="https://doi.org/10.1016/j.pcad.2019.11.011">https://doi.org/10.1016/j.pcad.2019.11.011</a>
DOI	<a href="https://doi.org/10.1016/j.pcad.2019.11.011">10.1016/j.pcad.2019.11.011</a>
Publisher	Elsevier
Journal	Progress in Cardiovascular Diseases
Download date	2026-03-16 18:25:03
License	<a href="https://creativecommons.org/licenses/by-nc-nd/4.0/">https://creativecommons.org/licenses/by-nc-nd/4.0/</a>
Link to Item	<a href="http://hdl.handle.net/2436/622940">http://hdl.handle.net/2436/622940</a>

1 **Improving reference equations for cardiorespiratory fitness using**  
2 **multiplicative allometric rather than additive linear models: Data from the**  
3 **Fitness Registry and the Importance of Exercise National Database Registry**

4

5 **Running Title:** An allometric reference equation for  $VO_{2max}$

6 Authors: Alan M. Nevill<sup>1</sup>, Jonathan Myers<sup>2,3</sup>, Leonard A. Kaminsky<sup>4</sup>, Ross Arena<sup>5</sup>

7 <sup>1</sup>Faculty of Education, Health and Wellbeing, University of Wolverhampton, Walsall  
8 Campus, Walsall, WS1 3BD, UK

9 <sup>2</sup>Division of Cardiology, VA Palo Alto Healthcare System, Palo Alto, CA

10 <sup>3</sup>Stanford University, Stanford, CA

11 <sup>4</sup>Fisher Institute of Health and Well-Being and Clinical Exercise Physiology  
12 Laboratory, Ball State University, Muncie, IN

13 <sup>5</sup>Department of Physical Therapy, College of Applied Sciences, University of Illinois  
14 at Chicago, Chicago, IL

15

16 **Address for correspondence:**

17 Professor Alan M. Nevill, Ph.D.

18 University of Wolverhampton,

19 Faculty of Education, Health and Wellbeing,

20 Walsall Campus

21 Gorway Road

22 Walsall, WS1 3BD

23 Tel: +44 (0)1902 322838

24 Fax: +44 (0)1902 322894

25 Email: [a.m.nevill@wlv.ac.uk](mailto:a.m.nevill@wlv.ac.uk)

26

27

28 **Abstract**

29 New improved reference equations for cardiorespiratory fitness have recently been  
30 published, using Data from the Fitness Registry and the Importance of Exercise  
31 National Database (FRIEND Registry). The new linear equation for  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ )  
32 was additive, derived using multiple-linear regression. An alternative multiplicative  
33 allometric model has also been published recently, thought to improve further the quality  
34 of fit. The purpose of the current study was to compare the accuracy and  
35 quality/goodness-of-fit of the linear, additive model with the multiplicative allometric  
36 model using the FRIEND database. The results identified that the allometric model out  
37 performs the linear model based on all model-comparison criteria. The allometric model  
38 demonstrates; 1) greater explained variance ( $R^2=0.645$ ;  $R=0.803$ ) vs. ( $R^2=0.62$ ;  
39  $R=0.79$ ), 2) residuals that were more normally distributed, 3) residuals that yielded less  
40 evidence of curvature, 4) superior goodness-of-fit statistics i.e., greater maximum log-  
41 likelihood (MLL) and smaller Akaike Information Criterion (AIC) statistics, 5) less  
42 systematic bias together with smaller unexplained standard error of estimates. The  
43 Bland and Altman plots also confirmed little or no evidence of curvature with the  
44 allometric model, but systematic curvature (lack-of-fit) in the linear model. The  
45 multiplicative allometric model to predict  $VO_{2max}$  was;

46 
$$VO_{2max} (ml.kg^{-1}.min^{-1}) = M^{0.854} \cdot H^{1.44} \cdot \exp(0.424 - .346 \cdot (sex) - 0.011 \cdot age),$$

47 where M=body mass and H=height ( $R^2=0.645$ ;  $R=0.803$ ) and sex is entered as a  
48 [0,1] indicator variable (male=0 and female=1). Another new insight obtained from  
49 the allometric model (providing construct validity) is that the height-to-body-mass  
50 ratio is similar to inverse body mass index or the lean body mass index, both  
51 associated with leanness when predicting  $VO_{2max}$ . In conclusion adopting allometric

52 models will provide more accurate predictions of  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ ) using more  
53 plausible, biologically sound and interpretable models.

54

55

56 **Keywords:** Curvature, power-function models, quality of fit, residuals.

57 **Abbreviations and acronyms**

58 BMI = body mass index

59 CPX = cardiopulmonary exercise test

60 CRF = cardiorespiratory fitness

61 CV = cardiovascular

62 CVD = cardiovascular disease

63 H = height

64 iBMI = inverse body mass index

65 KS = Kolmogorov-Smirnov

66 LBMI = lean body mass index

67 Ln =  $\log_e$

68 M = body mass

69 RER = respiratory exchange ratio

70  $VO_{2max}$  = maximal oxygen uptake

71

72 **Conflicts of interest statement**

73 The authors have no conflicts of interest

74

75 **Introduction**

76 Cardiorespiratory fitness (CRF), which can be accurately assessed by direct  
77 measurement of maximal oxygen consumption ( $VO_{2max}$ ), is a well-established and  
78 robust indicator of cardiovascular (CV) health, as well as a valuable predictor of all-  
79 cause mortality.<sup>1,2,3,4</sup> However, despite its importance, direct measurements of  
80  $VO_{2max}$  in epidemiological or population studies are rare primarily due to feasibility  
81 issues related to the time requirement to administer the test and the lack of having  
82 this measure routinely measured in clinical practice. As such, estimating or  
83 predicting CRF (i.e.,  $VO_{2max}$ ) has emerged as an attractive alternative; research to  
84 refine/improve  $VO_{2max}$  prediction models is ongoing.

85 Various prediction models using additive linear equations have been  
86 published recently, some include estimates of physical activity (PA)<sup>5</sup> while others do  
87 not.<sup>6</sup> However, Nevill and Cook<sup>7</sup> highlighted a number of concerns with these linear,  
88 additive models. Firstly, the models suggest linear associations with all key  
89 predictors such as age and body mass (M). However, there is strong evidence,  
90 certainly from the findings reported by Myers et al.<sup>6</sup> (see Figures 1 and 2), that  
91 curvature exists suggesting that at least one of these associations is likely to be non-  
92 linear.

93 The second concern is the absence of a body weight/mass term in the linear  
94 model proposed by Nes et al.,<sup>5</sup> and the absence of a height term in the model  
95 reported by Myers et al.<sup>6</sup> Astrand and Rodahl<sup>8</sup> in their Figure 9-4 on page 400 and  
96 Nevill et al.,<sup>9</sup> reported a strong negative association between  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ )  
97 and body mass. This is because absolute  $VO_{2max}$  ( $l.min^{-1}$ ) scales to, or is associated  
98 with, body mass ( $M^{0.67}$ ). Thus, when researchers calculate  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ )

99 relative to body mass by dividing  $VO_{2max}$  ( $l \cdot min^{-1}$ ) by body mass (M), the resulting  
100 ratio over-scales, leaving  $VO_{2max}$  ( $ml \cdot kg^{-1} \cdot min^{-1}$ ) proportional to  $M^{0.33}$ . This **non-linear**  
101 association with body mass was not considered by Nes et al.<sup>5</sup> or Myers et al.<sup>6</sup> By  
102 incorporating a power-function body-mass term as a predictor variable in their  
103 models, the result is likely to explain some of the curvature described above and, at  
104 the same time, improve the accuracy of the model. Furthermore, neither of the linear  
105 models proposed by Nes et al.<sup>5</sup> nor Myers et al.<sup>6</sup> incorporated height as a predictor  
106 variable. When height was included as a predictor variable in the allometric model to  
107 predict  $VO_{2max}$  ( $ml \cdot kg^{-1} \cdot min^{-1}$ ) proposed by Nevill and Cook,<sup>7</sup> a strong association  
108 was detected ( $P < 0.001$ ).

109 Another concern with these fitted linear additive models is the fact that the  
110 residuals from both models are unlikely to be: a) normally distributed; and b)  
111 homoscedastic (errors remain constant throughout the range of measurements; see  
112 **Figure 2** in Myers et al.<sup>6</sup>). If the residuals demonstrate a lack of normality and  
113 heteroscedasticity, then the validity of the models (i.e., the statistical significance of  
114 the estimated parameters) will be questionable. The alternative approach proposed  
115 by Nevill and Cook,<sup>7</sup> incorporating proportional, multiplicative allometric models, was  
116 found to overcome or at least reduce many of these problems; in particular,  
117 improving the normality and heteroscedasticity observed in the residuals. For a brief  
118 and concise history of allometric modeling, see Winter and Nevill.<sup>10</sup>

119 The purpose of the current study was to fit the same linear, additive model  
120 adopted by Myers et al.<sup>6</sup> to directly measure  $VO_{2max}$  ( $ml \cdot kg^{-1} \cdot min^{-1}$ ) data from the  
121 "Fitness Registry and the Importance of Exercise: A National Data Base" (FRIEND)  
122 Registry, to compare the original linear additive model with an alternative,  
123 proportional allometric model.<sup>7</sup> The comparison assessed whether the latter

124 provides: 1) a superior quality of fit (using  $R^2$ , maximum log-likelihood and AIC  
125 criterion); 2) more normally distributed residuals; 3) less explainable bias and smaller  
126 unexplained standard deviation of differences (standard error of the estimate); and 4)  
127 a more plausible, biologically sound and interpretable model.

## 128 **Methods**

129         The procedures used for acquiring and managing the data for the FRIEND  
130 registry have been previously reported.<sup>11</sup> In brief, laboratories determined by the  
131 advisory board to use valid and reliable calibration and cardiopulmonary exercise  
132 test (CPX) procedures administered by experienced personnel were invited to be  
133 considered for inclusion in the FRIEND Registry. Although there were some  
134 variations in laboratory equipment, protocols, and procedures defining  $VO_{2max}$ , the  
135 characteristics of all participating CPX laboratories are consistent with  
136 recommendations outlined in recently published guidelines.<sup>12,13</sup> Local institutional  
137 review board approval for participation in the FRIEND Registry was obtained by each  
138 participating CPX laboratory to submit de-identified, coded data to the data  
139 coordinating center at Ball State University, which then forwarded these data to the  
140 core CPX laboratory housed at the University of Illinois at Chicago. Institutional  
141 review board approval for the core CPX laboratory was also obtained at the  
142 University of Illinois at Chicago. Data from each CPX laboratory were reviewed for  
143 uniformity and to ensure data were within expected normal ranges by both the  
144 coordinating center and the core laboratory prior to merging into the FRIEND  
145 database.

## 146 **Study sample**

147 The study sample included 7759 subjects (4601 men and 3158 women, mean  
148  $46 \pm 13$  years) from ten participating CPX laboratories with geographical  
149 representation from Indiana, Louisiana, North Carolina, Oregon, Pennsylvania,  
150 Tennessee, and Texas. For inclusion, subjects were required to meet the following  
151 criteria: 1) age  $\geq 20$  years; 2) a maximal exercise test performed on a treadmill; and  
152 3) a peak respiratory exchange ratio (RER)  $\geq 1.00$ . The indications for the exercise  
153 tests were determination of fitness level before entry into an exercise program or for  
154 a research study. Laboratories provided data on individuals who at the time of the  
155 test were without known CV disease (CVD; i.e., coronary artery disease, peripheral  
156 vascular disease, or heart failure) or chronic obstructive pulmonary disease. Any  
157 subject identified as having a pre-existing diagnosis of CVD or pulmonary disease  
158 was excluded.

### 159 **Statistical methods**

160 The additive, linear model proposed by Myers et al.<sup>6</sup> is given by

$$161 \quad \text{VO}_{2\text{max}} (\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}) = a + b_1\cdot\text{age} + b_2\cdot M + b_3\cdot H + \varepsilon, \quad (\text{Eq. 1})$$

162 where M=mass, H=height (note that H=height has been added for comparative  
163 purposes) and  $\varepsilon$  is an additive error term that is assumed to be both normally  
164 distributed and homoscedastic (remains constant throughout the range of  
165 observations). The intercept “a” was allowed to vary with sex.

166 An alternative multiplicative model with allometric body size components  
167 originally proposed by Nevill and Holder<sup>14</sup> and subsequently reported by Nevill and  
168 Cooke<sup>7</sup> is given by

$$169 \quad \text{VO}_{2\text{max}} (\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}) = M^{k_1} \cdot H^{k_2} \cdot \exp(a + b_1\cdot\text{age}) \cdot \varepsilon, \quad (\text{Eq. 2})$$

170 where M=mass, H=height and 'ε' is a multiplicative, error ratio that assumes the error  
171 will be in proportion to  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ ).

172 The model (Eq. 2) can be linearized with a log transformation (using  $Ln=log_e$ ). A  
173 linear regression analysis on  $Ln(VO_{2max})$  can then be used to estimate the unknown  
174 parameters in the log transformed model i.e., the transformed model (Eq3) is now  
175 additive that conforms with the assumptions associated with ordinary least squares  
176 and ANOVA:

$$177 \quad Ln(VO_{2max})= k_1 \cdot Ln(M)+k_2 \cdot Ln(H) + a + b_1.age + Ln(\epsilon), \quad (Eq. 3)$$

178 where the residual errors  $Ln(\epsilon)$  are assumed to be normally distributed and  
179 homoscedastic and the intercept "a" is allowed to vary with sex.

180 Normality was assessed using the Ryan-Joiner and Kolmogorov-Smirnov  
181 tests. The Ryan-Joiner statistic measures how well the data follow a normal  
182 distribution by correlating the association between the measured data and the  
183 calculated normal scores. If the correlation coefficient is near 1, the population is  
184 likely to be normal. Larger values for the Kolmogorov-Smirnov statistic (KS) indicate  
185 that the data do not follow the normal distribution.

## 186 **Cross-validation**

187 To assess the validity or success of the two competing models, we adopted  
188 the same cross-validation proposed by Nes et al.,<sup>5</sup> by splitting the FRIEND data into  
189 two independent groups using a random split (80:20). We used the 80% sample to  
190 predict the competing two models and the 20% sample to test/validate the models.  
191 This was achieved by predicting the  $VO_{2max}$  of the 20% sample using the model  
192 derived from the 80% prediction models. The success of cross-validation process

193 was then assessed by comparing the measured  $VO_{2max}$  of 20% validation sample  
194 with the predicted  $VO_{2max}$  scores using correlations and Bland and Altman's limits of  
195 agreement<sup>15</sup> extended to assess systematic bias due to sex and age groups (using a  
196 two-way ANOVA), as well as the usual standard deviation of differences (The  
197 standard error of the estimate based on the ANOVA's residual mean-square errors).  
198 Bland and Altman plots are also reported and compared.

## 199 **Results**

### 200 *Additive linear models*

201 Fitting a similar linear model to that reported by Myers et al.,<sup>6</sup> plus an  
202 additional term for height (H), improves the fit as follows,

$$203 \quad VO_{2max} = 41.38 - 10.88 \cdot (\text{sex}) - 0.378 \cdot \text{age} - 0.310 \cdot M + 0.227 \cdot H, \quad (\text{Eq. 4})$$

204 where sex is entered as a [0,1] indicator variable (male=0 and female=1). The  $R^2$   
205 was 0.632 (R=0.795) with the standard error of estimate being 6.96 (ml.kg<sup>-1</sup>.min<sup>-1</sup>).

206 Note the original fit reported by Myers et al.<sup>6</sup> (excluding height) was  $R^2 = 0.62$   
207 (R=0.79) and the standard error of estimate was 7.2 (ml.kg<sup>-1</sup>.min<sup>-1</sup>).

208 When the residuals were saved and plotted against the predicted values (fits)  
209 as part of the usual model assessment diagnostics, the additive linear model  
210 demonstrates a clear lack of fit (evidence of curvature and heteroscedasticity) as  
211 seen in **Figure 1a**. The Ryan-Joiner statistic was r=0.996 and the Kolmogorov-  
212 Smirnov statistic KS=0.032, indicating that the residuals were not normally  
213 distributed, P<0.01.

214 ■ Figure 1a and 1b about here--

215 In an attempt to explain the curvature seen in **Figure 1a**, we examined the  
216 associations between  $VO_{2max}$  and the three predictor variables - body mass, height  
217 and age. Results suggest that the curvature was predominately due to the  
218 association between  $VO_{2max}$  and body mass (see **Figure 2**).

219 ■ Figure 2 about here--

### 220 *Multiplicative allometric models*

221 Fitting the multiplicative allometric model to the FRIEND's data using Eq. 3, we  
222 obtained,

$$223 \quad \ln(VO_{2max}) = -.854 \cdot \ln(M) + 1.44 \cdot \ln(H) + .424 - .346 \cdot (\text{sex}) - .011 \cdot \text{age}, \quad (\text{Eq. 5})$$

224 where sex is entered as a [0,1] indicator variable (male=0 and female=1). The  $R^2$   
225 was 0.678 ( $R=0.824$ ). (The  $R^2$  for the linear additive model Eq. 4 was 0.632  
226 ( $R=0.795$ )).

227 In contrast to the additive linear model, when the residuals were saved and plotted  
228 against the predicted values, the log-transformed multiplicative allometric model  
229 demonstrates an acceptable fit (little or no evidence of curvature nor  
230 heteroscedasticity), see **Figure 1b**. The Ryan-Joiner statistic was  $r=0.997$  and the  
231 Kolmogorov-Smirnov statistic  $KS=0.028$ ,  $P<0.01$ . Despite the fact that both statistics  
232 indicate that the residuals saved from the allometric model were not normally  
233 distribution they were closer to normality than the residuals saved using the linear  
234 additive model (Eq. 4).

235 Taking anti-logs of (Eq. 5), we obtain the multiplicative allometric model to predict  
236  $VO_{2max}$  ( $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ ) as follows;

237 
$$VO_{2max} = M^{0.854} \cdot H^{1.44} \cdot \exp(0.424 - .346 \cdot (\text{sex}) - 0.011 \cdot \text{age}), \quad (\text{Eq. 6})$$

238 where sex is entered as a [0,1] indicator variable (male=0 and female=1). The R<sup>2</sup>  
239 (measured vs. the predicted VO<sub>2max</sub> using Eq. 6) was 0.645 (R=0.803).

#### 240 *Comparing the goodness-of-fit*

241 Since the models are not nested or hierarchical, a direct comparison between  
242 two competing model forms (linear vs allometric) is not possible using traditional  
243 criteria such as the residual sum-of-squares, the standard error and the coefficient of  
244 determination (R<sup>2</sup>). However, Nevill and Holder<sup>14</sup> and Nevill et al.<sup>16</sup> chose the  
245 maximum likelihood criterion and the Akaike Information Criteria (AIC) as their  
246 standard criterion of model assessment (quality of fit) that does not require the  
247 competing models to be either nested or hierarchical.

248 A simple modification of the maximum log likelihood criterion is able to  
249 produce the Akaike Information Criteria (AIC= -2×(maximum log-likelihood) +  
250 2×(number of parameters fitted)), see goodness-of-fit data from both the linear and  
251 allometric models (**Table 1**).

252 ■ Table 1 about here--

#### 253 *Cross-validation assessment*

254 The FRIEND's dataset was randomly split into two independent samples: the  
255 prediction model sample =6214 (80.1%) and the validation/test sample = 1545  
256 (19.9%).

257 The results from the cross-validation assessment (using the 20% validation sample)  
258 found the correlations between the measured VO<sub>2max</sub> and the predicted VO<sub>2max</sub> were  
259 r=0.796 (linear) and r=0.809 (allometric). The differences between the measured and

260 predicted  $VO_{2max}$  (bias) were assessed using Bland and Altman's Limits of  
261 Agreements, extended to incorporate "explainable" bias due to sex and age groups.  
262 The bias (measured – predicted  $VO_{2max}$ ) from linear and allometric prediction models  
263 were assessed using a two-way ANOVA. The mean bias by age and sex can be  
264 seen in **Figures 3a and 3b**, respectively. Note that the "explainable" variance in bias  
265 from the linear additive model using the two-way ANOVA was  $R^2=1.9\%$ . The  
266 equivalent "explainable" variance in bias from the allometric-model using the two-  
267 way ANOVA was  $R^2=1.2\%$ .

268 ■ Figure 3a and 3b about here--

269 Having explained these "systematic" biases due to age and sex (using ANOVA), the  
270 remaining unexplained standard deviation of differences (standard error of estimate)  
271 were  $\pm 6.87$  ( $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ) (linear model) and  $\pm 6.71$  ( $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ) (allometric model)  
272 (based on the ANOVA's mean-square errors). The Bland and Altman plots  
273 (differences vs means) for the linear and allometric prediction models are given in  
274 **Figures 4a and 4b**, respectively.

275 ■ Figure 4a and 4b about here--

## 276 Discussion

277 The answer to the question "can we improve the reference equation for normal  
278 standards for  $VO_{2max}$ , using multiplicative allometric rather than additive linear  
279 models?" would appear to be categorically yes. The allometric model (Eq. 2)  
280 performs better than the linear model (Eq.1), originally proposed by Myers et al.<sup>6</sup>  
281 based on ALL model-comparison criteria: 1) The explained variance. The explained  
282 variance was greater using the allometric models, using either the log-transformed  
283 model ( $R^2=0.678$ ;  $R=0.824$ ) or the multiplicative allometric model ( $R^2=0.645$ ;

284 R=0.803) compared with the linear models (excluding height) ( $R^2 = 0.62$ ; R=0.79) or  
285 (including height) ( $R^2=0.632$ ; R=0.795); 2) Tests of normality: Both the Ryan-Joiner  
286 and Kolmogorov-Smirnov tests indicate that the residuals saved from the linear and  
287 allometric models were not normally distributed. However, the residuals from the  
288 allometric model were closer to a normal distribution than the residuals saved from  
289 the linear model. However, the Ryan-Joiner statistics (Q-Q plot correlations) for the  
290 linear and allometric models were 0.996 and 997 respectively, resulting in arguably  
291 acceptable linearity; 3) The residuals vs the predicted values (fits) plots. Evidence of  
292 curvature (lack-of-fit) was observed with the linear model in **Figure 1a**. No such  
293 evidence was apparent with the allometric model in **Figure 1b**; 4) The goodness of  
294 fit assessed using the maximum log-likelihood (MLL) and the Akaike Information  
295 Criterion (AIC). The MLL was greater, and the AIC was smaller with the allometric  
296 model compared with the linear additive models (**see Table 1**).; 5) The cross-  
297 validation assessment. Based on the validation sample, the correlation between the  
298 measured  $VO_{2max}$  and the predicted  $VO_{2max}$  was  $r=0.796$  (linear) and  $r=0.809$   
299 (allometric). The bias was less evident from the allometric model ( $R^2=1.2\%$   
300 compared with  $R^2=1.9\%$  with the linear model) see **Figure 3b** compared with the  
301 linear model **Figure 3a**. The Bland and Altman plots confirmed a lack-of-fit  
302 (curvature) in the linear model but little evidence of curvature with the allometric  
303 model.

304         Perhaps the most persuasive argument for choosing the allometric model  
305 comes from the lack-of-fit/curvature observed in the three figures based on linear  
306 models: **Figure 1a**, **Figure 2** and **Figure 4a**. The same curvature was also observed  
307 in **Figures 1** and **2** reported by Myers et al.,<sup>6</sup> In effect, the linear model  
308 systematically underestimates  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ ) among both lighter and heavier

309 participants but overestimates  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ ) among the more “average” body  
310 mass participants. This might explain the negative bias of 70 year-old male  
311 participants (see **Figure 3a**) who are relatively light (“average” mass) having  
312 “survived” to the age of 70, and also the negative bias of a large, homogeneous  
313 group of 30 year-old female participants (see **Figure 3a**) who have maintained an  
314 “average” body mass especially compared to their male counterparts.

315 Two other helpful new insights were obtained from the fitted multiplicative  
316 allometric model (Eq. 6). The first comes from observing that the fitted mass (M) and  
317 height (H) exponents have opposite signs. The resulting product  $H^{1.44} \cdot M^{-0.854}$  can be  
318 expressed as a ratio ( $H^{1.44}/M^{0.854}$ ) or ( $H^{1.69}/M$ )<sup>0.854</sup> not dissimilar to the inverse BMI  
319 ( $iBMI=H^2/M$ ). This index has a sound biological interpretation in terms of its  
320 association with cardiorespiratory fitness [i.e.,  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ )]. The inverse  
321  $BMI^{17}$  also known as the Lean Body Mass Index (LBMI)<sup>18</sup>, is a measure of leanness.  
322 Clearly having a greater lean body mass (LBMI) should be strongly and positively  
323 associated with  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ ), thus providing further biological support (and  
324 construct validity) for the allometric model (Eq. 6).

325 The second helpful new insight obtained from the allometric model (Eq. 6)  
326 comes from the negative decline in  $VO_{2max}$  with age. The linear model incorporates  
327 the age decline as a negative linear term that would theoretically predict a negative  
328  $VO_{2max}$  for very old participants. The allometric model incorporates the negative age  
329 decline within an exponential term, thus ensuring the prediction of  $VO_{2max}$  remains  
330 positive irrespective of the participants’ age.

331 We acknowledge that the current study is not without limitations. Although we  
332 have been able to demonstrate the benefits of the allometric model for  $VO_{2max}$ ,

333 obtaining similar results to those reported by Nevill and Cooke,<sup>7</sup> further work is  
334 required. For example, future research should attempt to explain the systematic bias  
335 observed in **Figure 3b**, in particular the negative bias observed in the 70 year-old  
336 age groups.

337 In summary, the quality of fit associated with predicting  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ )  
338 using the allometric model was superior to the linear additive models, based on all  
339 model-comparison criteria. As seen in **Figure 2**, fitting a linear additive model will  
340 systematically over-estimate the  $VO_{2max}$  for participants who are either very light or  
341 very heavy, but under-estimate  $VO_{2max}$  for participants who have “average” body  
342 mass. The allometric model also identified a stature-to-body-mass ratio, very similar  
343 to LBMI or inverse BMI, known to be associated with leanness, a new insight that  
344 leads to a more plausible, biologically sound and interpretable model when  
345 predicting  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ ).

346 **Acknowledgments** FRIEND Consortium Contributors: Ball State University  
347 (Leonard Kaminsky), Brooke Army Medical Center (Kenneth Leclerc), Cone Health  
348 (Paul Chase), Johns Hopkins University (Kerry Stewart), Pennington Biomedical  
349 Research Center (Timothy Church), Southern Connecticut State University (Robert  
350 Axtell), Taylor University (Erik Hayes), University of Illinois–Chicago (Jacob Haus),  
351 University of Tennessee, Knoxville (David Bassett).

352

## 353 **References**

354 1. Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory  
355 fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific

- 356 statement from the American Heart Association. *Circulation*  
357 2016;134(24):e653-e699.
- 358  
359 2. Ozemek C, Laddu DR, Lavie CJ, et al. An update on the role of  
360 cardiorespiratory fitness, structured exercise and lifestyle physical activity in  
361 preventing cardiovascular disease and health risk. *Prog Cardiovasc Dis*  
362 2018;61:484-490.
- 363 3. Kaminsky LA, Arena R, Ellingsen Ø, et al. Cardiorespiratory fitness and  
364 cardiovascular disease: the past, present, and future. *Prog Cardiovasc Dis*  
365 2019;62:86-93.
- 366 4. Imboden MT, Harber MP, Whaley MH, et al. The association between the  
367 change in directly measured cardiorespiratory fitness across time and  
368 mortality risk. *Prog Cardiovasc Dis* 2019;62:157-162.
- 369 5. Nes BM, Janszky I, Vatten LJ, et al. Estimating VO<sub>2</sub>peak from a Non-exercise  
370 Prediction Model: The HUNT Study, Norway. *Med. Sci. Sports Exerc.*  
371 2011;43:2024–30.
- 372 6. Myers, J., Kaminsky, L. A., Lima, R., et al. A reference equation for normal  
373 standards for VO<sub>2</sub> max: analysis from the Fitness Registry and the  
374 Importance of Exercise National Database (FRIEND Registry). *Prog*  
375 *Cardiovasc Dis* 2017;60:21-29.
- 376 7. Nevill, A.M. and Cooke, C.B. The Dangers of Estimating VO<sub>2</sub>max Using  
377 Linear, Non-exercise Prediction Models. *Med. Sci. Sports Exerc.*  
378 2017;49:1036–1042 <[https://journals.lww.com/acsm-  
379 msse/Fulltext/2017/05000/The\\_Dangers\\_of\\_Estimating\\_V\\_O2max\\_Using\\_Lin  
380 ear,.22.aspx](https://journals.lww.com/acsm-msse/Fulltext/2017/05000/The_Dangers_of_Estimating_V_O2max_Using_Linear,.22.aspx)>.

- 381 8. Astrand P-O and Rodahl K: Textbook of work physiology, 3rd ed. New York:  
382 McGraw-Hill; 1986.
- 383 9. Nevill AM, Ramsbottom R, and Williams C. Scaling physiological  
384 measurements for individuals of different body size. *Eur J Appl*  
385 *Physiol.*1992;65: 110-7.
- 386 10. Winter E.M. and Nevill A.M. Scaling: Adjusting for differences in body size. In  
387 "Kinanthropometry and Exercise Physiology Laboratory Manual for Tests,  
388 Procedures and Data, 3rd ed." (eds) R. G. Eston and T. Reilly, Abingdon,  
389 Oxon, Routledge; 2009;300-314 p.
- 390 11. Kaminsky LA, Arena R, Beckie TM, et al. The importance of cardiorespiratory  
391 fitness in the United States: the need for a national registry. A policy  
392 statement from the American Heart Association. *Circulation.* 2013;127(5):652-  
393 662.
- 394 12. Balady GJ, Arena R, Seitsema K, et al. A clinician's guide to cardiopulmonary  
395 exercise testing. A scientific statement from the American Heart Association.  
396 *Circulation.* 2010;122: 191-225.
- 397 13. Myers J, Arena R, Franklin B, et al. Recommendations for clinical exercise  
398 laboratories: a scientific statement from the American Heart Association.  
399 *Circulation.* 2009;119(24):3144-3161.
- 400 14. Nevill AM and Holder RL. Modelling maximum oxygen uptake - A case study  
401 in non-linear regression model formulation and comparison. *Journal of the*  
402 *Royal Statistical Society, Series C.* 1994; 43: 653-66.
- 403 15. Bland, JM and Altman, D. Statistical methods for assessing agreement  
404 between two methods of clinical measurement. *The lancet,* 1986;327(8476):  
405 307-310.

- 406 16. Nevill AM, Holder RL, Baxter-Jones A, et al. Modeling developmental changes  
407 in strength and aerobic power in children. *J Appl Physiol.* 1998;84: 963-70.
- 408 17. Nevill AM, Stavropoulos-Kalinoglou A., Metsios G.S., et al. Inverted BMI  
409 rather than BMI is a better proxy for percentage of body fat. *Ann Hum Biol.*  
410 2011; 38: 681-4.
- 411 18. Nevill AM and Holder RL. Body Mass Index; a measure of fatness or  
412 leanness? *The British Journal of Nutrition*, 1995;73: 507-516.
- 413

414 **Legends to Tables**

415 **Table 1. The maximum log-likelihood (MLL) and Akaike Information Criterion**  
416 **(AIC) together with the number of fitted parameters for the competing models to**  
417 **predict  $VO_{2max}$**

418

419

420 **Legends to figures**

421 **Figure 1. The association between the residuals and the predicted values (fits)**  
422 **saved from fitting 1a) the additive linear model (Eq. 1), and 1b) the log-**  
423 **transformed multiplicative allometric model (Eq. 5)**

424 **Figure 2. The association between  $VO_{2max}$  ( $ml.kg^{-1}.min^{-1}$ ) and body mass (kg)**  
425 **for female participants (linear  $R^2=0.316$ ; power function  $R^2=0.374$ )**

426 **Figure 3. The mean bias (measured – predicted  $VO_{2max}$ ) from 3a) the linear**  
427 **prediction model, and 3b) the allometric prediction model, assessed using a**  
428 **two-way ANOVA (sex and age group)**

429 **Figure 4. The Bland and Altman plot (differences vs means) using the**  
430 **validation sample, data from 4a) the linear model and 4b) the allometric model**  
431 **to predict  $VO_{2max}$**

432

433

434

435 **Table 1.**

436	<b>Competing models</b>	<b>MLL (N par)</b>	<b>AIC</b>
437	Additive linear model	-25919.01 (4)	51846
438	Log-transformed allometric model (Eq. 3)	-25116.57 (4)	50241

439 N par =number of fitted parameters

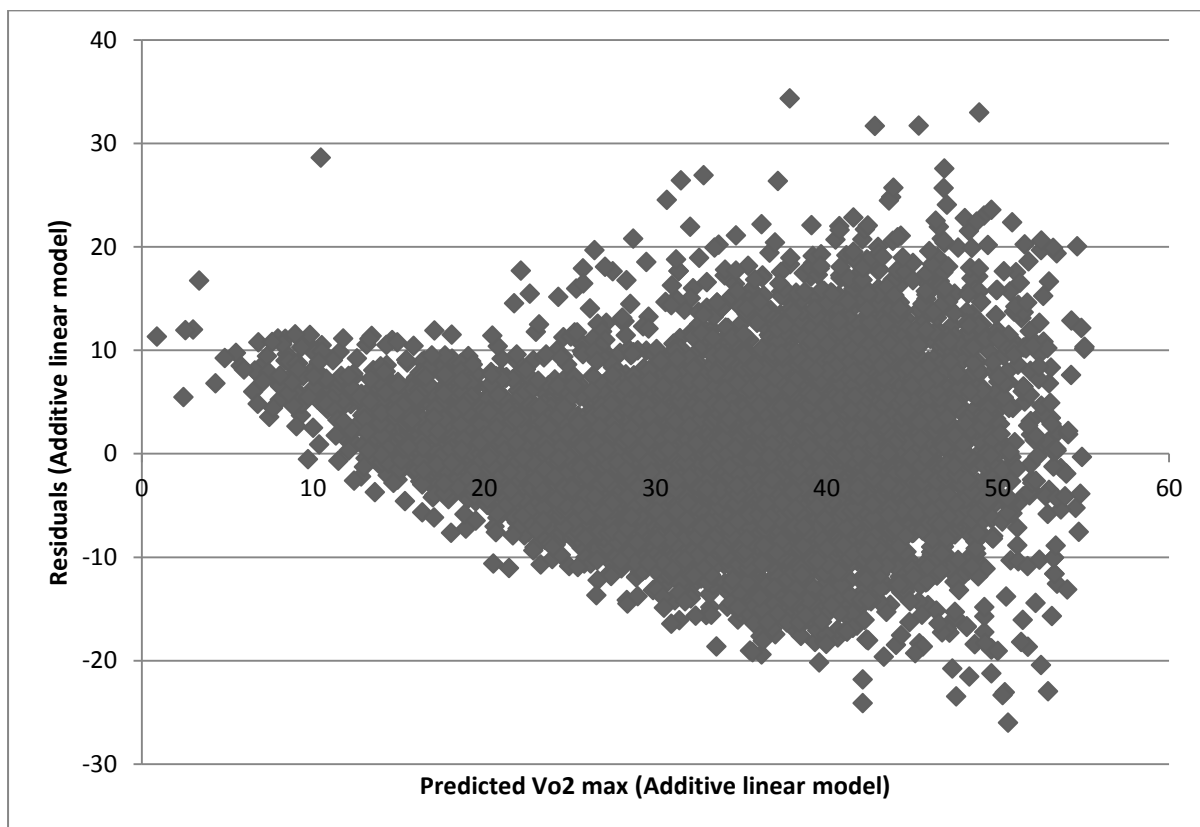
440 Note that the best model is the one with the greatest MLL (i.e., least negative)

441 and/or the smallest AIC

442

443

444 **Figure 1a.**

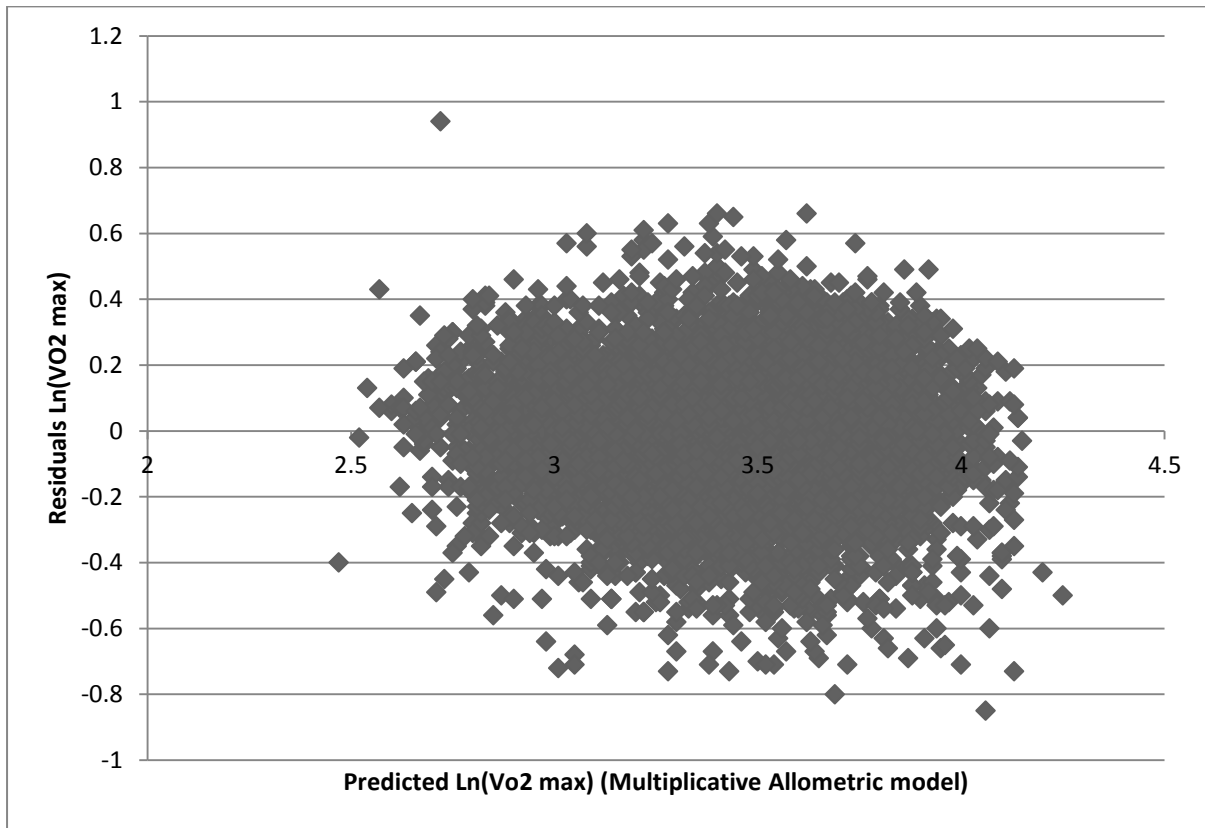


445

446

447

448 **Figure 1b**

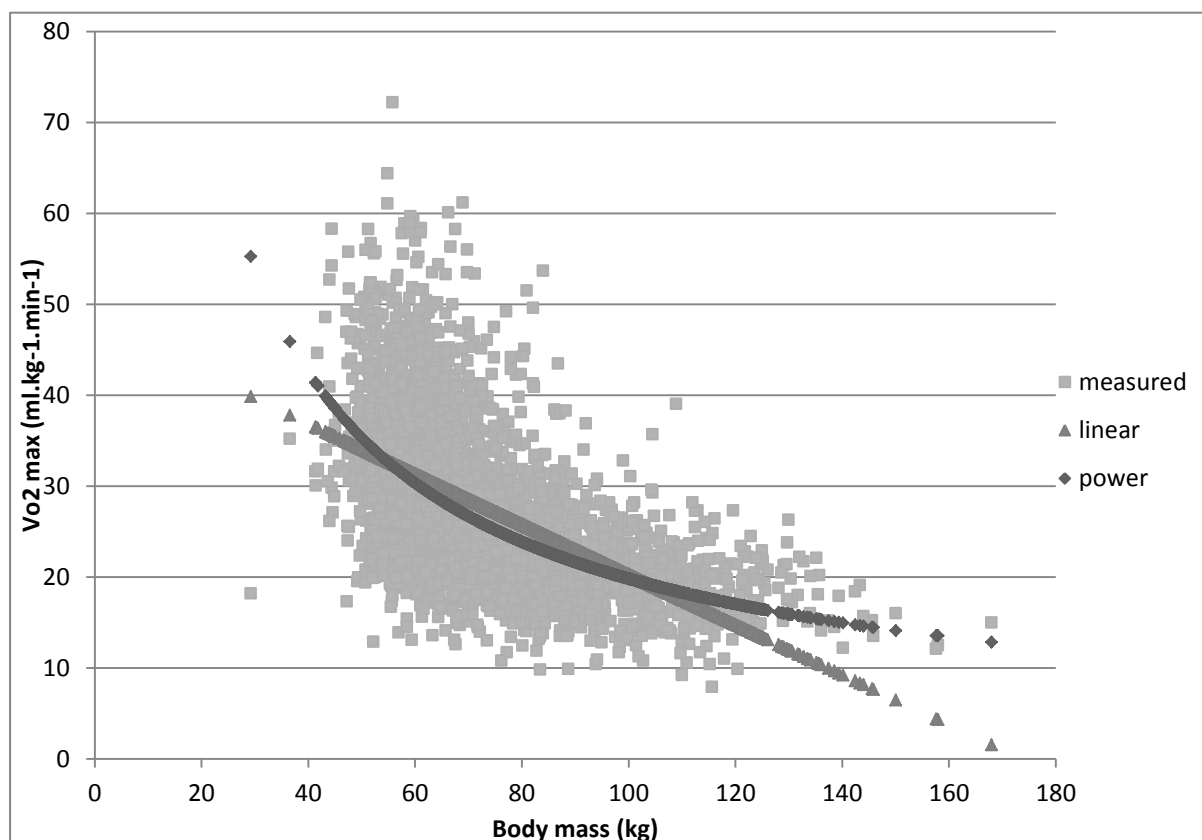


449

450

451

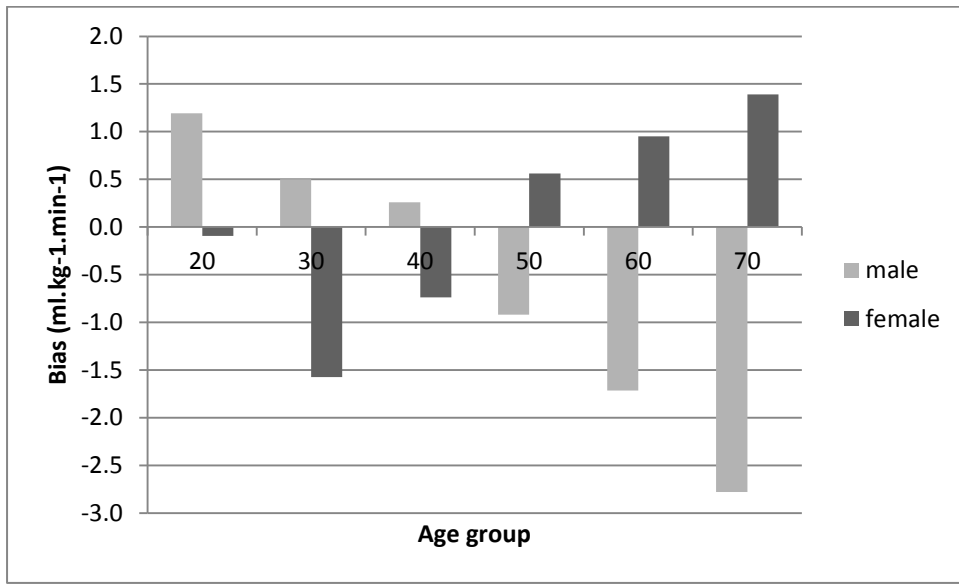
452 **Figure 2.**



453

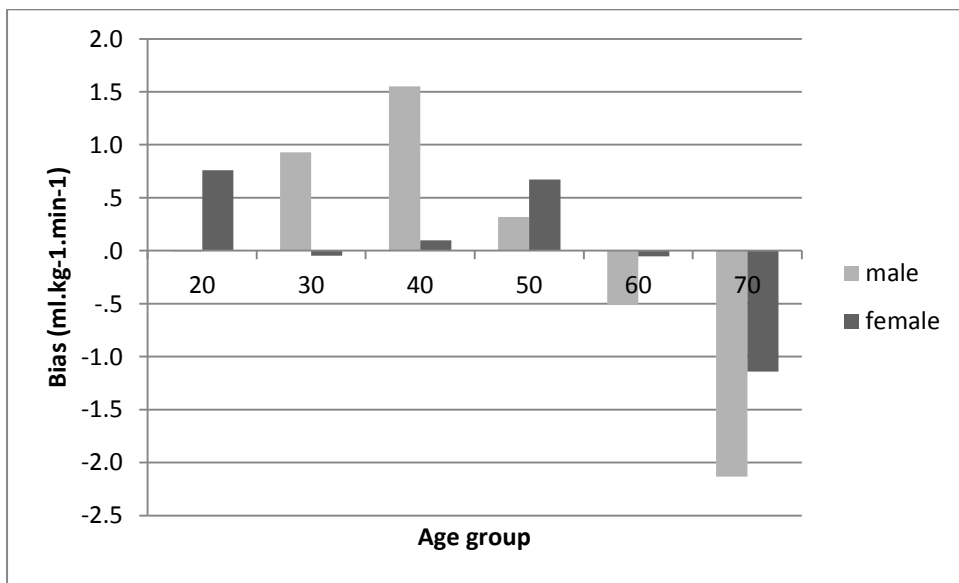
454

455 **Figure 3a**



456

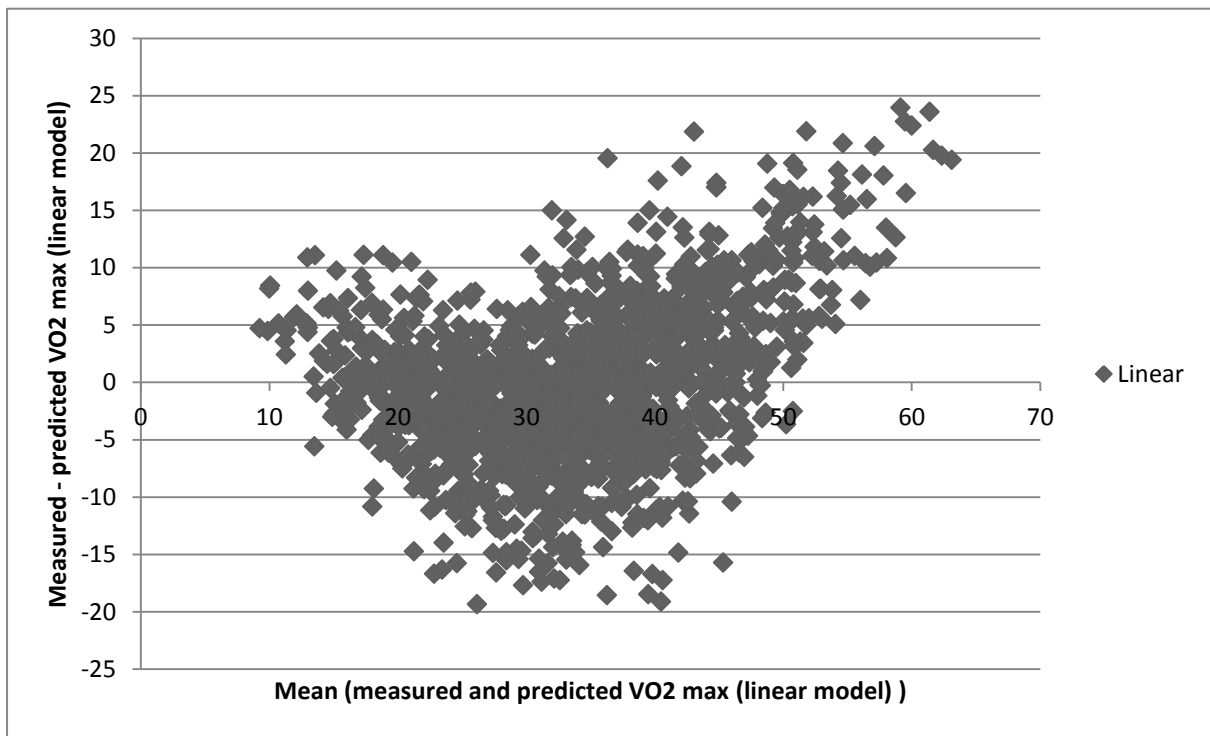
457 **Figure 3b.**



458

459

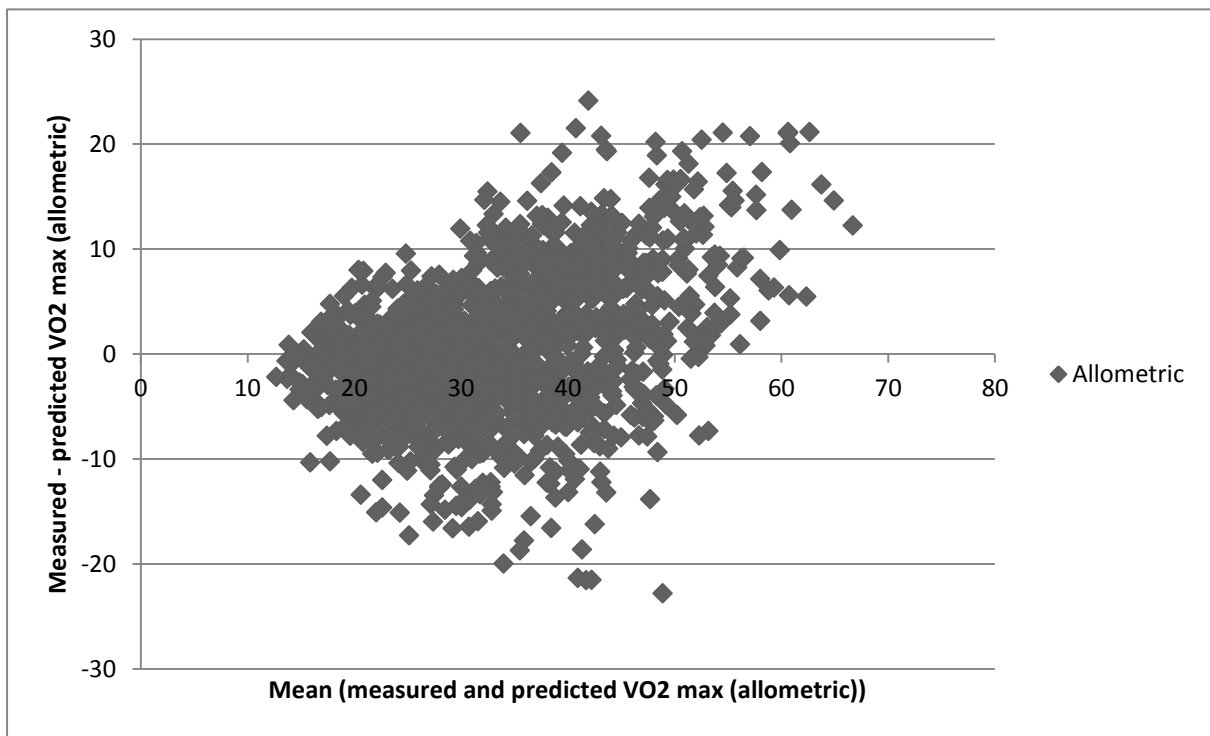
460 **Figure 4a.**



461

462

463 **Figure 4b.**



464