

## Early View

Original article

# Comparing individual and population differences in VE/VCO<sub>2</sub> slopes using centile growth curves and log-linear allometry

Alan M. Nevill, Jonathan Myers, Leonard A. Kaminsky, Ross Arena, Tony D. Myers

Please cite this article as: Nevill AM, Myers J, Kaminsky LA, *et al.* Comparing individual and population differences in VE/VCO<sub>2</sub> slopes using centile growth curves and log-linear allometry. *ERJ Open Res* 2021; in press (<https://doi.org/10.1183/23120541.00088-2021>).

This manuscript has recently been accepted for publication in the *ERJ Open Research*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJOR online.

Copyright ©The authors 2021. This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact [permissions@ersnet.org](mailto:permissions@ersnet.org)

**Comparing individual and population differences in VE/VCO<sub>2</sub> slopes using centile growth curves and log-linear allometry.**

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

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

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

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

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

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

<sup>6</sup>Sport and Health, Newman University, Birmingham, UK

**Address for correspondence:**

Professor Alan M. Nevill, Ph.D.  
University of Wolverhampton,  
Faculty of Education, Health and Wellbeing,  
Walsall Campus  
Gorway Road  
Walsall, WS1 3BD  
Tel: +44 (0)1902 322838  
Fax: +44 (0)1902 322894  
Email: [a.m.nevill@wlv.ac.uk](mailto:a.m.nevill@wlv.ac.uk)

## Abstract

Identifying vulnerable groups and/or individuals' cardiorespiratory fitness (CRF) is an important challenge for clinicians/researchers alike. To quantify CRF accurately, the assessment of several variables is now standard practice including maximal oxygen uptake ( $\text{VO}_2$ ) and ventilatory efficiency, the latter assessed using the minute ventilation/carbon dioxide production ( $\text{VE}/\text{VCO}_2$ ) slope. Recently, reference values (centiles) for  $\text{VE}/\text{VCO}_2$  slopes for men and women aged 20 to 80 have been published, using cardiopulmonary exercise testing (CPX) data (treadmill protocol) from the Fitness Registry and the Importance of Exercise National Database (FRIEND Registry).

In the current observational study we provide centile *curves* for the FRIEND Registry  $\text{VE}/\text{VCO}_2$  slopes, fitted using the generalised additive model for location, scale and shape (GAMLSS), to provide individuals with a more precise estimate of where their  $\text{VE}/\text{VCO}_2$  slopes fall within the population. We also confirm that by adopting allometric models (incorporating a log-transformation), the resulting ANCOVAs provided more normal and homoscedastic residuals, with superior goodness-of-fit using the Akaike information criterion  $\text{AIC}=14671$  (compared with traditional ANCOVA's  $\text{AIC}=15008$ ) that confirms allometric models are vastly superior to traditional ANCOVA models.

In conclusion, providing sex-by-age centile curves rather than referring to reference tables for ventilatory efficiency ( $\text{VE}/\text{VCO}_2$  slopes) will provide more accurate estimates of where an individual's particular  $\text{VE}/\text{VCO}_2$  slope falls within the population. Also, by adopting allometric models researchers are more likely to identify real and valid inferences when analysing population/group differences in  $\text{VE}/\text{VCO}_2$  slopes.

Key words: Ventilatory efficiency, AIC goodness of fit, population differences, cardiorespiratory fitness (CRF), exercise testing.

## Introduction

Numerous studies and subsequent position statements have confirmed the importance of cardiopulmonary exercise testing (CPX) as the gold-standard approach for the assessment of cardiorespiratory fitness [1–3]. The term cardiorespiratory fitness (CRF) refers to the ability of the circulatory and respiratory systems to supply oxygen to skeletal muscles during sustained physical activity. The clinical and research focus of CPX is usually focused on the measurement/quantification of peak or maximal oxygen uptake ( $\text{VO}_2$ ). Accordingly, considerable effort has been made to establish reference values for peak/maximal  $\text{VO}_2$  across the lifespan in both men and women [4–6].

Although peak/maximal  $\text{VO}_2$  is the gold-standard measurement for exercise capacity specifically, it does not capture all of the characteristics associated with CRF, or all the CPX responses that are useful in stratifying risk [3]. To quantify CPX responses more comprehensively, the assessment of other parallel variables is necessary. One such response is ventilatory efficiency, most commonly assessed as the minute ventilation/carbon dioxide production ( $\text{VE}/\text{VCO}_2$ ) slope, a measure that dramatically improves diagnostic and prognostic resolution as well as more accurately quantifies the degree of therapeutic efficacy for surgical, pharmacologic and lifestyle interventions [1, 3].

In their recent study, Arena et al., [7] provide important reference values as discrete point estimates for  $\text{VE}/\text{VCO}_2$  slopes across the lifespan in both men and women (reporting mean values at 10 percent intervals per decile). Although a helpful guide, these discrete point estimates require a certain degree of “linear interpolation” to estimate precisely where an individual’s  $\text{VE}/\text{VCO}_2$  slope falls within the table (e.g., somewhere between the 80<sup>th</sup> and 90<sup>th</sup> centile for a 68-year-old). An alternative **continuous** “growth” curve-fitting methodology for centile reference values (fitted using the Lamda-mu-sigma (LMS) proposed by Cole [8] will provide a more accurate estimate.

Comparing population differences in VE/VCO<sub>2</sub> slopes are usually performed using ANOVA or ANCOVA. These analyses assume these data are normally distributed and homoscedastic. As far as the ratio variables such as peak/maximal VO<sub>2</sub> (mL.kg<sup>-1</sup>.min<sup>-1</sup>) and VE/VCO<sub>2</sub> slopes are concerned, these assumptions are unlikely to be satisfied [6] with evidence of the heteroscedastic spread of the VE/VCO<sub>2</sub> slope data with age in the centile plots below in **Figures 1a and 1b**. Ratio data such as peak/maximal VO<sub>2</sub> (mL.kg<sup>-1</sup>.min<sup>-1</sup>) and the VE/VCO<sub>2</sub> slope will invariably benefit from the log transformation [9] . Hence the purpose of the current study is as follows: 1) to provide “individual” comparisons using the centile *curves* for VE/VCO<sub>2</sub> slope values rather than point estimates (to help overcome the linear interpolation issues described above), curves estimated using the generalised additive model for location, scale and shape (GAMLSS), fitted using the GAMLSS package [10] in R [11]; and 2) to compare population differences in the VE/VCO<sub>2</sub> slope using a more valid analysis (ANOVA/ANCOVA) with the help of an allometric model (incorporating a log-transformation) that are likely to provide more normally and homoscedastic residuals, thus making any inference concerning population differences more trustworthy.

## Methods

### The FRIEND Registry

The FRIEND Registry was established with the primary charge of establishing reference CRF values across the adult lifespan [4, 12]. The CPX laboratories currently contributing data to the FRIEND Registry function consistent with recommendations provided in recently published guidelines including the use of valid and reliable calibration and testing procedures as well as employing experienced personnel qualified to conduct exercise tests [13, 14]. Participating CPX laboratories were responsible for obtaining local Institutional review board approval, or exemption from review, for inclusion in the FRIEND Registry, providing documentation that they were authorized to submit de-identified, coded data to the core CPX laboratory housed at the University of Illinois at Chicago (UIC). Institutional review board approval for the core CPX laboratory was also obtained at UIC.

## Cohort Characteristics, Inclusion/Exclusion Criteria and Study Data Points

The current analysis included 2,511 tests from the following 10 CPX laboratories, see Table 1. All CPX tests undertaken by the laboratories adopted a treadmill protocol. The mean ( $\pm$ sd) time to completion of the exercise tests was 11.23 (3.06) mins.

Inclusion criteria for the current analysis included CPX data on men and women: 1) aged > 20 years; and 2) with a maximal CPX performed on a treadmill following established guidelines [1, 2]. Note that only treadmill protocols were performed in all 10 laboratories. Any subject identified as having a pre-existing medical condition was excluded. Thus, the cohort included apparently healthy subjects (i.e., subjects without a pre-existing medical condition) who achieved maximal exertion by established peak RER criteria [1]. Descriptive data for all subjects by sex and age groups are provided in Table 2.

VO<sub>2</sub>, peak RER, peak heart rate (HR) and the VE/VCO<sub>2</sub> slope were the CPX variables reported in the current study. Each participating CPX laboratory input VE and VCO<sub>2</sub> data from the initiation of exercise to peak and calculated the VE/VCO<sub>2</sub> slope via least squares linear regression ( $y = mx + b$ ,  $m$ =slope).

### **Statistical methods (to assess individual differences using centile curves)**

The centile curves for the VE/VCO<sub>2</sub> slope by age were produced using the Lambda-mu-sigma (LMS) method. This method assumes that using a power transformation, data can be appropriately normalised, by stretching one tail of the distribution and shrinking the other and so removing skewness [8]. The power transformation applied in the present models used the Box–Cox  $t$  (BCT [ $\mu$ ,  $\sigma$ ,  $v$ ,  $\tau$ ]). This distribution is defined by  $Y^v$  having a shifted truncated  $t$  distribution with  $\tau$  degrees of freedom,  $\mu$  is the median of the distribution,  $\sigma \cdot (\tau / (\tau - 2))^{0.5}$  is approximate the coefficient of variation,  $v$  controls the skewness and  $\tau$  the kurtosis of the distribution [15]. The generalised additive model for location, scale and shape (GAMLSS) allows each of the parameters of the distribution to be modelled as linear and or non-linear parametric, and or smooth non-parametric functions of explanatory variables. A Fisher scoring algorithm was used to fit the model by maximising a penalised likelihood [10].

The analysis and centile curves were fitted and plotted using the GAMLSS package [16] in R [11]. Post- estimation diagnostics for these models included standard QQ-plots but also detrended normal QQ-plots (worm plots) to check the age-conditional normality of the transformed data [17].

### **Statistical methods (to assess population differences)**

The additive, linear model for VE/VCO<sub>2</sub> slope proposed by [7] is given by

$$\text{VE/VCO}_2 \text{ slope} = a + b_1 \cdot \text{age} + b_2 \cdot \text{BMI} + \varepsilon, \quad (\text{Eq. 1})$$

where  $\varepsilon$  is an additive error term that is assumed to be both normally distributed and homoscedastic (remains constant throughout the range of observations). The intercept “a” was allowed to vary with sex.

An alternative multiplicative model for the VE/VCO<sub>2</sub> slope with allometric body size components originally proposed for VO<sub>2max</sub> by Nevill and Holder [18] and subsequently adapted by Nevill and Cooke [19] to include an age<sup>2</sup> term, is given by

$$\text{VE/VCO}_2 \text{ slope} = M^{k_1} \cdot H^{k_2} \cdot \exp (a + b_1 \cdot \text{age} + b_2 \cdot \text{age}^2) \cdot \varepsilon, \quad (\text{Eq. 2})$$

where M=mass, H=height and ‘ $\varepsilon$ ’ is a multiplicative, error ratio that assumes the error will be in proportion to VE/VCO<sub>2</sub> slope.

The model (Eq. 2) can be linearized with a log transformation (using Ln=log<sub>e</sub>). A linear regression or ANCOVA analysis on Ln(VE/VCO<sub>2</sub> slope) can then be used to estimate the unknown parameters in the log transformed model i.e., the transformed model (Eq3) is now additive that conforms with the assumptions associated with ordinary least squares and ANOVA:

$$\text{Ln(VE/VCO}_2 \text{ slope)} = k_1 \cdot \text{Ln(M)} + k_2 \cdot \text{Ln(H)} + a + b_1 \cdot \text{age} + b_2 \cdot \text{age}^2 + \text{Ln}(\varepsilon), \quad (\text{Eq. 3})$$

where the residual errors Ln( $\varepsilon$ ) are assumed to be normally distributed and homoscedastic.

For researchers unfamiliar with analyzing these allometric models (Eq.3), the dependent variable is defined as the log-transformed VE/VCO<sub>2</sub> slopes, Ln(VE/VCO<sub>2</sub> slope), and Ln(M),

$\ln(H)$ , age and age<sup>2</sup> are incorporated as co-variables. Population differences in VE/VCO<sub>2</sub> slope between groups, for example sex, can be assessed by allowing the intercept “a” to vary for each fixed factor in the ANOVA/ANCOVA analyses.

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

Model comparison (goodness-of-fit) between the linear and allometric models was assessed using the Akaike information criterion (AIC). Traditionally, R<sup>2</sup> is frequently used to measure goodness of fit. However, higher R<sup>2</sup> values do not always indicate a better fit. Higher R<sup>2</sup> can indicate overfitting and adding noise variables will also inflate R<sup>2</sup>. So R<sup>2</sup> while useful, is not necessarily the best method of comparing competing models. An alternative method of model comparison is to use the Akaike information criterion. AIC can be conceptualised as a “distance” or error between the data and a model, lower values meaning a better model. Unlike R<sup>2</sup>, that rewards added noise variables with a higher (better) value, adding noise variables will increase AIC which penalises its values (lower being better). A difference between two AIC values less than 2, is really irrelevant; differences > 2 ≤ 6, evidence for the lower model is somewhat positive > 6 ≤ 10; the evidence for the model with the lower AIC value is strong, > 10; the evidence for the model with the lower AIC is very strong. Given differences between AIC values can be challenging to comprehend - beyond appreciating that one model is better than another - evidence ratios (see 19] will also be calculated to present the magnitude of the difference in terms of how many times better the best model is in terms of its predictive accuracy [21, 22].

## **Results (individual comparisons)**



The centile curves for the VE/VCO<sub>2</sub> slope by age are given for males and females separately in **Figures 1a and 1b**, and comparative purposes, the 50<sup>th</sup> percentile values for the male and female subjects by age in **Figure 1c**.

These curves enable the reader to estimate an individual's VE/VCO<sub>2</sub> slope value within a population of healthy individuals (N=2511) for comparative purposes. Point-estimate centile tables by age for male and female individuals are also given in **Table 3**.

## **Results (population comparisons)**

### *Linear, additive models*

Comparing population differences in VE/VCO<sub>2</sub> are usually performed using ANOVA or ANCOVA. To assess differences in sex using equation 1, we obtain,

$$\text{VE/VCO}_2 \text{ slope} = 26.4 + 0.09(\text{age}) - 0.08(\text{BMI in kg/m}^2) + 0.83(\text{sex}),$$

where the variable sex is incorporated as a [0, 1] indicator variable, given by male = 0 and female = 1. However, the total R value for this equation was 0.33 with an adjusted R-square of 0.10 and a SEE of 4.8. Parameter estimates and 95% confidence intervals are given in Table 4a.

When the residuals were plotted against the “fits” or predicted (VE/VCO<sub>2</sub> slope) values, the plot (Figure 2a) illustrates a lack of fit (the “fits” quadratic term was B=0.26, t=8.14; P<0.001) and heteroscedasticity (confirmed when the absolute residuals were correlated against the predicted values (N=2511), r=0.089, P<0.001).

The Shapiro-Wilk statistic was r=0.967 and the Kolmogorov-Smirnov statistic KS=0.05, indicating that the residuals were not normally distributed, P<0.001.

### *Multiplicative allometric model*

Fitting the multiplicative allometric model to the FRIEND's data using Eq. 3, we obtained the fitted parameters given in **Table 4b**.

In order to compare the quality of fit (AIC) values using the same parameters as linear model Eq.1, but adopting the allometric model structure, we obtained AIC= 14671.04. Using the allometric model with age (but not age<sup>2</sup>), height, weight and gender as predictors, AIC= 14672.33. Note that the AIC for the linear model Eq.1 was 15007.99 (**Table 4a**), a difference in AIC=336.95. The evidence ratio for this difference in AIC values indicates the allometric model is  $3.921438 \times 10^{71}$  times better in terms of model parsimony than the additive linear model. Parsimony is important because it helps discriminate signal from the noise, facilitating generalisation and increasing predictive ability for new data.

When the residuals were plotted against the “fits” or predicted Ln(VE/VCO<sub>2</sub> slope) values, the plot (**Figure 2b**) illustrates acceptable fit (the “fits” quadratic term was B=-0.9, t=1.72; P=0.086) and homoscedasticity (confirmed when the absolute residuals were correlated against the predicted values (N=2511), r=0.000, P=0.998).

The Shapiro-Wilk statistic was r=0.992 (compared with r=0.967 using linear Eq.1) and the Kolmogorov-Smirnov statistic KS=0.031 (compared with KS=0.05 using Eq.1), indicating that the residuals were still not normally distributed, P<0.001 but behave considerably better (more normal) than the residuals obtained from the linear additive model (Eq. 1).

To illustrate the possible inferential pitfalls that might occur when investigating population differences using linear models, we re-examined the VE/VCO<sub>2</sub> slopes from the FRIEND dataset using allometric models but incorporating a sex-by-age interaction term as well as a sex main effect and allowing subjects to be nested within sites. The re-analysis revealed a significant age-by-sex interaction (P<0.001) that can be clearly seen in **Figure 1c**, a finding that appears to have been overlooked by [7] using linear additive model (Eq. 1).

## Discussion

The first aim/purpose of the current study was to provide centile “growth” curves (by age) for ventilatory efficiency as recorded by the VE/VCO<sub>2</sub> slope values. Arena et al. [7] provided discrete point-estimate centile reference values for the VE/VCO<sub>2</sub> slope by age (using deciles) and sex, when all treadmill-testing exercise data were used to calculate the slopes. In the current study we report **continuous** “growth” curves for the centile reference values (fitted using the Lamda-mu-sigma (LMS) method proposed by Cole [8]) that provides the reader the opportunity to more accurately **interpolate** an individual’s reference values between the point estimates (deciles) reported by Arena et al., [7]. In addition to this inherent error (or lack of precision) associated from using point estimates provided by reference tables, the reference tables would only allow a CPX administrator to say how close to ‘normal’ the patient’s VE/VCO<sub>2</sub> slope value was. In comparison, the centile curves provide a valuable added context associated with the additional level of precision. The interpretation of centiles is straightforward. For example, in the case of an individual’s VE/VCO<sub>2</sub> slope and age, if their estimate is on the 25th centile, it means that for every 100 individuals of that age, 25 would have a lower VE/VCO<sub>2</sub> slope and 75 a higher VE/VCO<sub>2</sub> slope. These results may prove useful in enhancing the interpretation of CPX results. We recognise that similar centile curves have been published elsewhere recently by Wagner et al. [23] but using a different CPX protocol (bicycle ergometry) and with a reduced sample size [24].

Comparing the centile curves reported by Wagner et al [23] using bicycle ergometry with those reported in the current study using a treadmill protocol, we detected large differences. The 50<sup>th</sup> centiles for the male VE/VCO<sub>2</sub> slopes reported by Wagner et al. [22] ranged from 34.2 to 37.2 (for 30 years to 60-year-old males respectively). In contrast, 50<sup>th</sup> centiles for the male VE/VCO<sub>2</sub> slopes reported in the current study (treadmill protocol) ranged from 26.31 to 29.46 for 30 years to 60-year-old males respectively. A similar contrast is observed with female VE/VCO<sub>2</sub> slopes, The 50<sup>th</sup> centiles for the female VE/VCO<sub>2</sub> slope scores reported by Wagner et al.[23] ranged from 34.0 to 36.5 (30 years to 60 year old

females respectively). The 50<sup>th</sup> centiles for the female VE/VCO<sub>2</sub> slope scores reported in the current study (treadmill) ranged from 27.76 to 29.97 for 30 years to 60-year-old females respectively. These differences suggest that the protocol used to obtain the VE/VCO<sub>2</sub> slopes is crucial and outlines the necessity to refer to the appropriate centile curves depending of the CPX protocol used.

The second aim/purpose of the current study was to recommend how researchers should compare population differences in the VE/VCO<sub>2</sub> slopes. The majority of researchers exploring population differences in VE/VCO<sub>2</sub> slopes have used traditional ANOVA, ANCOVA or regression methods (such as Eq. 1) to explore differences between categorical variables such as age, sex, smoking habits, BMI categories and exercise habits [5, 7, 25, 26]. However, the range of ratio variables (such as VO<sub>2</sub>max [mL.kg<sup>-1</sup>.min<sup>-1</sup>] and the VE/VCO<sub>2</sub> slope) that are also associated with body size are invariably bounded by zero to the left (cannot be negative) but unbounded to the right. This leads its frequency distribution to be positively skewed and not normally distributed.

The solution to overcome such problems would appear to adopt an alternative multiplicative model with allometric body size components (Eq. 2). The model can be linearized with a log transformation (using  $\text{Ln}=\log_e$ ) and linear regression or ANCOVA analysis on  $\text{Ln}(\text{VE/VCO}_2 \text{ slope})$  can then be used to estimate the unknown parameters (see Eq. 3). The log-transformation will naturally overcome both the positive skewness and heteroscedasticity, the latter seen clearly in the centile curves Figures 1a and 1b, where the more extreme curves (e.g., 90% and 10% centile) diverge with increasing age in both the male and female figures.

The allometric model (Eq. 2) performs better than the linear model (Eq.1) based on ALL model-comparison criteria: 1) The explained variance was greater using the allometric model ( $R^2=0.151$ ) compared with the linear models ( $R^2 = 0.105$ ); 2) Both the Shapiro-Wilk statistic and Kolmogorov-Smirnov tests indicate that the saved residuals from the linear and allometric models were not normally distributed. However, the residuals from the allometric model were closer to a normal distribution than the residuals saved from the linear model,

with the Q-Q plot correlations for the linear and allometric models being 0.967 and 0.992 respectively; 3) The residuals vs the predicted values (fits) plots provided evidence of curvature (lack-of-fit) as seen with the linear model in **Figure 2a**. No such evidence was apparent with the allometric model in **Figure 2b**; and 4) The goodness of fit assessed using the Akaike Information Criterion (AIC) was much lower (AIC=14679.3) for the allometric model compared with the linear additive model (AIC=15008.0). The 328.7 difference between models, points to a major improvement, some 71 orders of magnitude greater terms of parsimony and therefore predictive accuracy.

Clearly, these results suggest that allometric models (incorporating a log-transformation) are likely to provide a better fit to the  $VE/VCO_2$  slope values with more normally and homoscedastic residuals, thus making any inference concerning population differences more valid and trustworthy. This was reinforced when the allometric model identified the sex-by-age interaction (see Figure 1c) that appears to have been missed using linear additive models (Arena et al. 2020). Some authors [25] have also reported an age-by-sex interaction (in their Table 4), but as yet no mechanism has been provided to explain why ventilatory efficiency ( $VE/VCO_2$  slope) is higher in females than males in younger age groups but converges (little or no difference) in the older age groups. This can be considered a limitation of the study.

In summary, providing centile curves (by age and sex) rather than having to refer to mean reference tables for ventilatory efficiency ( $VE/VCO_2$  slopes) should provide individuals with a more accurate estimate of where their particular  $VE/VCO_2$  slope falls within the general population. Also, by adopting the allometric model (Eq. 2), fitted using the log-transformed model/equation (Eq. 3), researchers are more likely to identify real and/or valid inferences, leading to more trustworthy interpretations of population differences (and interactions) when analysing  $VE/VCO_2$  slopes.

## References

1. Balady GJ, Arena R, Sietsema K, Myers J, Coke L, Fletcher GF, Forman D, Franklin B, Guazzi M, Gulati M. Clinician's guide to cardiopulmonary exercise testing in adults: a scientific statement from the American Heart Association. *Circulation* 2010; 122: 191–225.
2. Guazzi M, Adams V, Conraads V, Halle M, Mezzani A, Vanhees L, Arena R, Fletcher GF, Forman DE, Kitzman DW. EACPR/AHA Scientific Statement. Clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. *Circulation* 2012; 126: 2261–2274.
3. Ross R, Blair SN, Arena R, Church TS, Després J-P, Franklin BA, Haskell WL, Kaminsky LA, Levine BD, Lavie CJ. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation* 2016; 134: e653–e699.
4. Kaminsky LA, Arena R, Myers J. Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing: data from the Fitness Registry and the Importance of Exercise National Database. *Mayo Clin. Proc.* Elsevier; 2015. p. 1515–1523.
5. Loe H, Steinshamn S, Wisløff U. Cardio-respiratory reference data in 4631 healthy men and women 20-90 years: the HUNT 3 fitness study. *PLoS One* 2014; 9: e113884.
6. Ingle L, Rigby A, Brodie D, Sandercock G. Normative reference values for estimated cardiorespiratory fitness in apparently healthy British men and women. *PLoS One* 2020; 15: e0240099.
7. Arena R, Myers J, Harper M, Phillips SA, Severin R, Ozemek C, Peterman JE, Kaminsky LA. The VE/VCO<sub>2</sub> Slope During Maximal Treadmill Cardiopulmonary Exercise Testing: Reference Standards from FRIEND (Fitness Registry and the Importance of Exercise): An International Database. 2020; .
8. Cole TJ. The LMS method for constructing normalized growth standards. *Eur. J. Clin. Nutr.* 1990; 44: 45–60.
9. Nevill AM. Why the analysis of performance variables recorded on the ratio scale will invariably benefit from a log transformation. *J. Sports Sci.* 1997; 15, 457–458.
10. Rigby RA, Stasinopoulos DM. Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Stat. Model.* 2006; 6: 209–229.
11. R Core Team. R: A language and environment for statistical computing. 2020; .
12. Kaminsky LA, Arena R, Beckie TM, Brubaker PH, Church TS, Forman DE, Franklin BA, Gulati M, Lavie CJ, Myers J. The importance of cardiorespiratory fitness in the United States: the need for a national registry: a policy statement from the American Heart Association. *Circulation* 2013; 127: 652–662.
13. Myers J, Arena R, Franklin B, Pina I, Kraus WE, McInnis K, Balady GJ. Recommendations for clinical exercise laboratories: a scientific statement from the American Heart Association. *Circulation* 2009; 119: 3144–3161.

14. Myers J, Forman DE, Balady GJ, Franklin BA, Nelson-Worel J, Martin B-J, Herbert WG, Guazzi M, Arena R. Supervision of exercise testing by nonphysicians: a scientific statement from the American Heart Association. *Circulation* 2014; 130: 1014–1027.
15. Stasinopoulos MD, Rigby RA, Bastiani FD. GAMLSS: a distributional regression approach. *Stat. Model.* 2018; 18: 248–273.
16. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *J. R. Stat. Soc. Ser. C Appl. Stat.* 2005; 54: 507–554.
17. Buuren S van, Fredriks M. Worm plot: a simple diagnostic device for modelling growth reference curves. *Stat. Med.* 2001; 20: 1259–1277.
18. Nevill AM, Holder RL. Modelling Maximum Oxygen Uptake—a Case- Study in Nonlinear Regression Model Formulation and Comparison. *J. R. Stat. Soc. Ser. C Appl. Stat.* 1994; 43: 653–666.
19. Nevill AM, Cooke CB. The Dangers of Estimating  $\dot{V}O_{2\max}$  Using Linear, Nonexercise Prediction Models. *Med. Sci. Sports Exerc.* [Internet] 2017; 49 Available from: [https://journals.lww.com/acsm-msse/Fulltext/2017/05000/The\\_Dangers\\_of\\_Estimating\\_V\\_O2max\\_Using\\_Linear,.22.aspx](https://journals.lww.com/acsm-msse/Fulltext/2017/05000/The_Dangers_of_Estimating_V_O2max_Using_Linear,.22.aspx).
20. Anderson D, Burnham K. Model selection and multi-model inference. *Second. Springer-Verl.* 2004; 63: 10.
21. Akaike H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* 1974; 19: 716–723.
22. Spiegelhalter D, Best NG, Carlin BP, Linde A van der. Bayesian measures of model complexity and fit. *Qual. Control Appl. Stat.* 2003; 48: 431–432.
23. Wagner J, Knaier R, Infanger D, Königstein K, Klenk C, Carrard J, Hanssen H, Hinrichs T, Seals D, Schmidt-Trucksäss A. Novel CPET Reference Values in Healthy Adults: Associations with Physical Activity. *Med. Sci. Sports Exerc.* United States; 2021; 53: 26–37.
24. Cole TJ. Sample size and sample composition for constructing growth reference centiles. *Stat. Methods Med. Res.* 2021; 30: 488–507.
25. Koch B, Schäper C, Ittermann T, Spielhagen T, Dörr M, Völzke H, Opitz CF, Ewert R, Gläser S. Reference values for cardiopulmonary exercise testing in healthy volunteers: the SHIP study. *Eur. Respir. J.* 2009; 33: 389–397.
26. Salvioni E, Corrà U, Piepoli M, Rovai S, Correale M, Paolillo S, Pasquali M, Magrì D, Vitale G, Fusini L. Gender and age normalization and ventilation efficiency during exercise in heart failure with reduced ejection fraction. *ESC Heart Fail.* 2020; 7: 368–377.

Table 1. Participating CPX laboratories

- 1) Ball State University, Muncie, Indiana;
- 2) Brigham and Women's Hospital, Boston, Massachusetts;
- 3) Brooke Army Medical Center, Houston, Texas;
- 4) Goshen Heart and Vascular Center, Goshen, Indiana;
- 5) Massachusetts General Hospital, Boston, Massachusetts;
- 6) Saint Francis University, Loretto, Pennsylvania;
- 7) San Francisco State University, San Francisco, California;
- 8) University of Kansas Medical Center, Kansas City, Kansas;
- 9) Veteran's Hospital – Palo Alto, Palo Alto, California;
- 10) Yale University, New Haven, Connecticut.



Table 2. Descriptive data for all subjects by sex and age group in years

Sex	age group	Age			Weight		Height		BMI	
		N	(yrs)	sd	(kg)	sd	(cm)	sd	(kg.m <sup>-2</sup> )	sd
Female	20.00	252	23.91	2.283	64.97	2.983	146.57	34.319	24.40	5.364
	30.00	156	34.83	2.970	65.25	2.375	171.91	48.196	28.37	7.735
	40.00	206	44.79	2.871	64.44	2.614	163.41	43.749	27.66	7.128
	50.00	227	53.78	2.728	64.38	2.845	165.01	40.204	28.02	6.666
	60.00	209	64.17	2.787	64.46	2.537	160.40	33.313	27.19	5.650
	70.00	140	74.19	2.810	63.69	2.543	154.72	27.747	26.85	4.491
	80.00	43	82.37	2.000	62.30	2.306	139.92	25.157	25.37	4.249
	total	1233	48.85	17.848	64.48	2.740	159.02	38.996	26.90	6.344
Male	20.00	242	24.34	2.532	70.52	3.030	180.40	32.833	25.51	4.482
	30.00	221	34.39	2.939	70.68	2.924	204.40	51.681	28.72	6.738
	40.00	252	44.50	2.941	70.72	2.870	208.93	37.885	29.35	4.927
	50.00	220	54.04	2.769	70.32	2.848	203.79	39.039	29.00	5.458
	60.00	181	64.60	2.941	70.00	2.876	194.83	36.489	27.95	4.825
	70.00	111	73.99	2.862	69.44	3.118	188.40	31.797	27.52	4.380
	80.00	51	82.78	2.248	69.33	3.042	170.77	22.598	24.99	2.720
	total	1278	47.51	17.336	70.34	2.961	196.56	40.423	27.92	5.375

sd=standard deviation

**Table 3. Centile table for VE/VCO<sub>2</sub> slope for healthy 20-85-year-olds (centiles 1-99)**

<b>Males</b>		<b>Centiles</b>							
<b>age (ys)</b>	<b>C1</b>	<b>C3</b>	<b>C10</b>	<b>C25</b>	<b>C50</b>	<b>C75</b>	<b>C90</b>	<b>C97</b>	<b>C99</b>
<b>20</b>	19.34	20.88	22.92	25.02	27.48	30.24	33.18	36.79	40.18
<b>25</b>	19.01	20.54	22.56	24.63	27.05	29.77	32.68	36.26	39.65
<b>30</b>	18.69	20.21	22.20	24.24	26.62	29.31	32.18	35.75	39.15
<b>35</b>	18.45	19.96	21.94	23.96	26.31	28.97	31.83	35.40	38.83
<b>40</b>	18.35	19.87	21.85	23.86	26.21	28.86	31.73	35.34	38.84
<b>45</b>	18.44	19.99	22.01	24.04	26.40	29.08	31.99	35.68	39.31
<b>50</b>	18.71	20.31	22.37	24.44	26.85	29.58	32.56	36.38	40.18
<b>55</b>	19.13	20.79	22.92	25.05	27.52	30.32	33.40	37.39	41.42
<b>60</b>	19.71	21.45	23.67	25.88	28.44	31.34	34.55	38.76	43.08
<b>65</b>	20.34	22.18	24.51	26.81	29.46	32.47	35.83	40.30	44.97
<b>70</b>	20.93	22.87	25.30	27.69	30.43	33.55	37.06	41.81	46.84
<b>75</b>	21.60	23.66	26.22	28.70	31.55	34.79	38.48	43.54	49.04
<b>80</b>	22.41	24.61	27.32	29.93	32.91	36.31	40.21	45.66	51.72
<b>85</b>	23.10	25.45	28.31	31.03	34.12	37.67	41.77	47.63	54.31
<b>Females</b>									
<b>20</b>	18.90	21.10	23.73	26.15	28.77	31.50	34.27	37.57	40.60
<b>25</b>	18.61	20.79	23.40	25.83	28.46	31.23	34.07	37.46	40.62
<b>30</b>	18.28	20.43	23.03	25.46	28.11	30.92	33.81	37.32	40.62
<b>35</b>	17.97	20.09	22.67	25.10	27.76	30.61	33.57	37.20	40.66
<b>40</b>	17.77	19.86	22.42	24.84	27.51	30.40	33.44	37.21	40.85
<b>45</b>	17.76	19.82	22.35	24.77	27.45	30.37	33.49	37.41	41.24
<b>50</b>	17.94	19.97	22.48	24.88	27.57	30.52	33.71	37.78	41.82
<b>55</b>	18.36	20.37	22.87	25.26	27.97	30.96	34.23	38.47	42.76
<b>60</b>	19.11	21.14	23.65	26.07	28.82	31.89	35.28	39.74	44.34
<b>65</b>	20.12	22.17	24.71	27.17	29.97	33.13	36.66	41.38	46.35
<b>70</b>	21.15	23.21	25.77	28.25	31.09	34.32	37.97	42.93	48.26
<b>75</b>	22.12	24.18	26.75	29.23	32.09	35.36	39.10	44.28	49.96
<b>80</b>	23.26	25.34	27.91	30.41	33.28	36.60	40.44	45.84	51.90
<b>85</b>	24.73	26.84	29.44	31.97	34.90	38.29	42.25	47.93	54.45

1.

1.

**3.**

**3.**

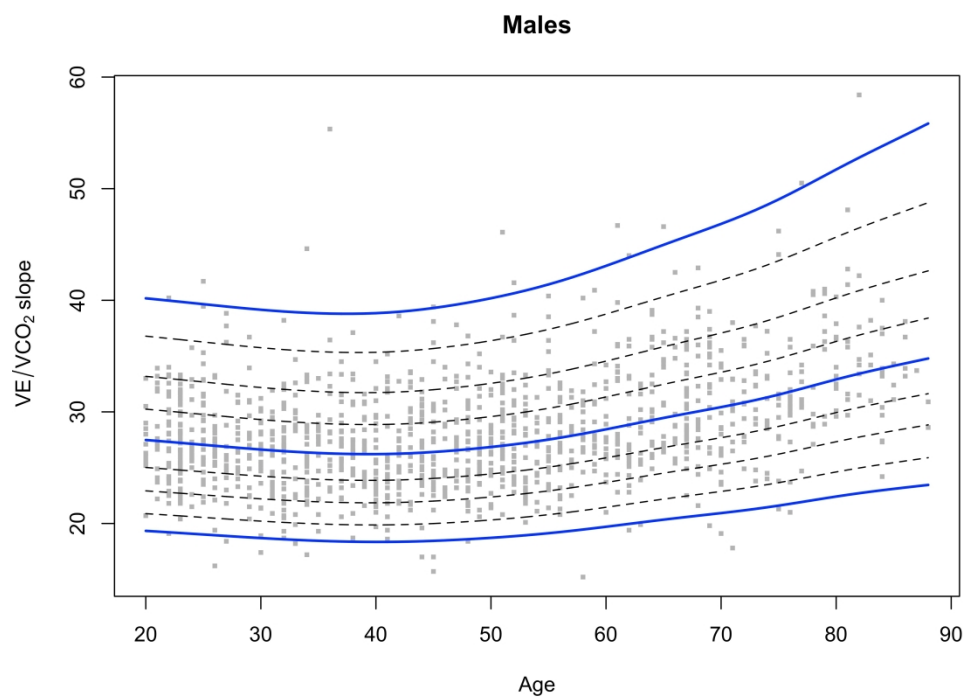


Figure 1a. The centile curves for VE/VCO<sub>2</sub> slope by age for males

1693x1270mm (72 x 72 DPI)

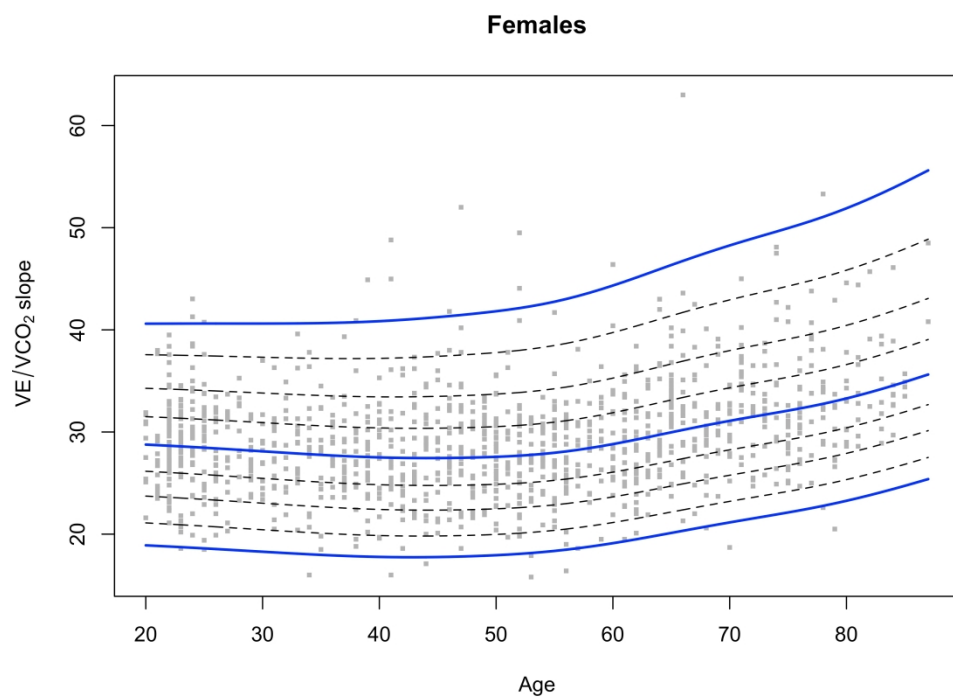


Figure 1b. The centile curves for VE/VCO<sub>2</sub> slope by age for females

1693x1270mm (72 x 72 DPI)

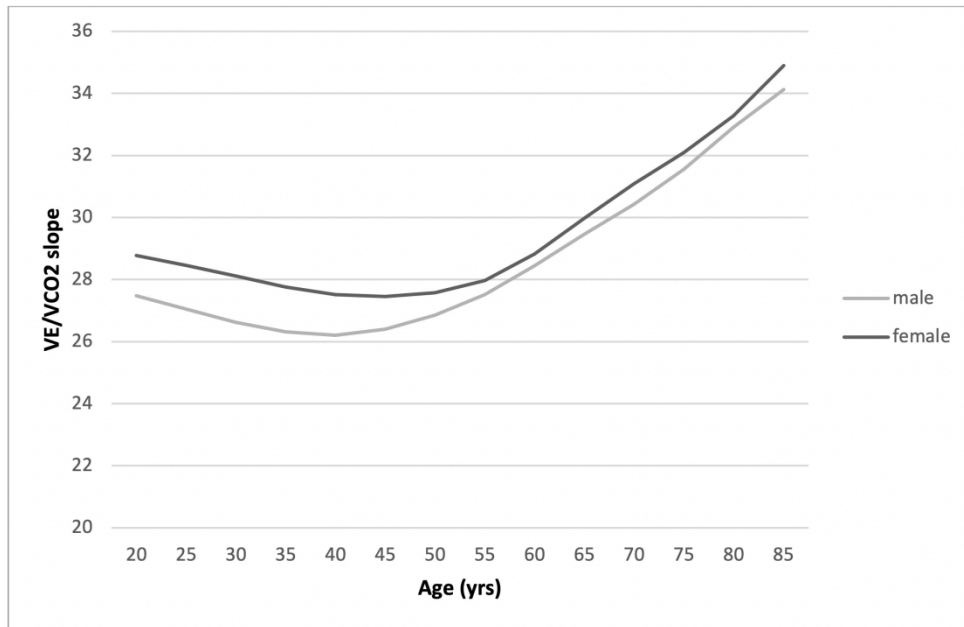


Figure 1c. The 50th centile curves for VE/VCO<sub>2</sub> slope for males and females by age.

464x302mm (144 x 144 DPI)

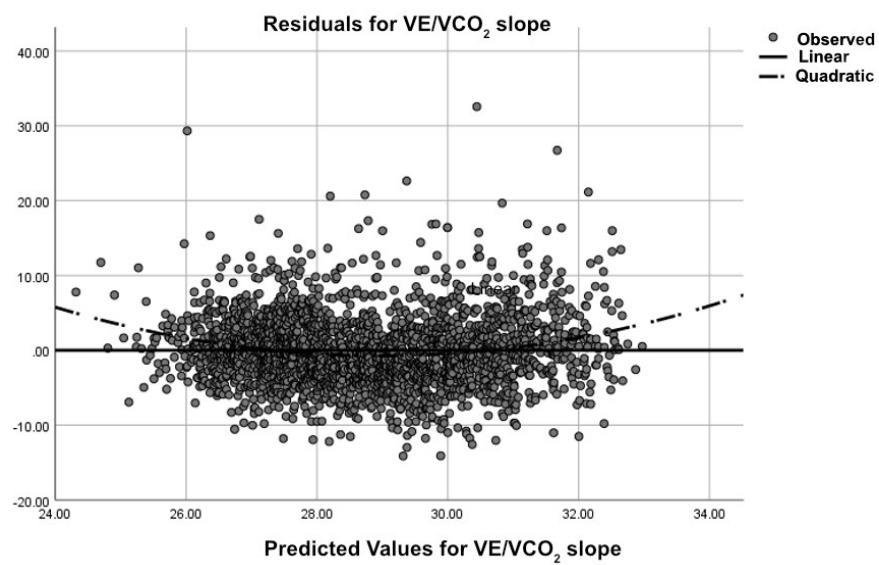


Figure 2a. The residuals versus the "fits" or predicted (VE/VC<sub>02</sub> slope) values using Eq. 1.

166x97mm (150 x 150 DPI)



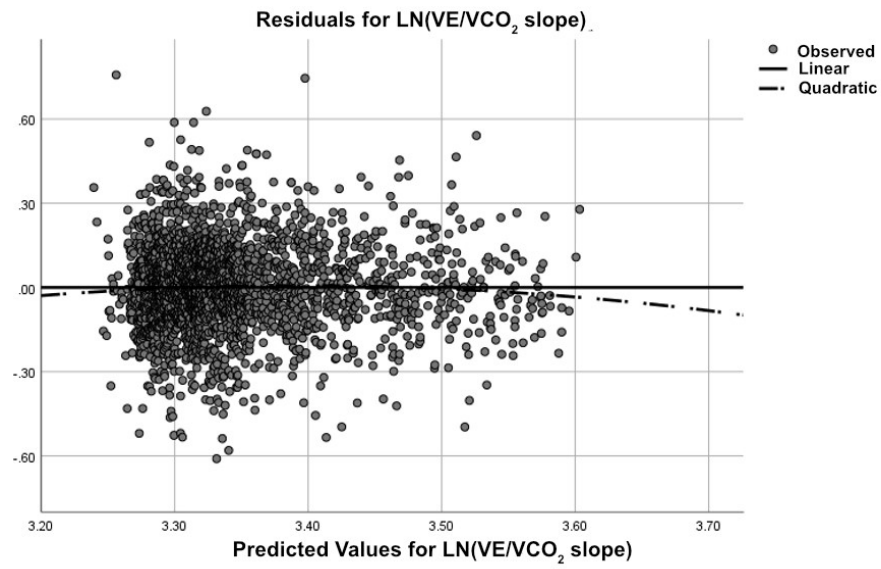


Figure 2b. The residuals versus the "fits" or predicted (VE/VCO<sub>2</sub> slope) values using Eq. 3.

166x97mm (150 x 150 DPI)