

# Sex-Specific Longitudinal Modeling of Youth Peak Oxygen Uptake

Neil Armstrong and Jo Welsman

University of Exeter

**Purpose:** To investigate peak oxygen uptake ( $\dot{V}O_2$ ) in relation to sex, age, body mass, fat-free mass (FFM), maturity, and overweight status. **Methods:** Multiplicative, allometric models of 10- to 18-year-olds were founded on 1057 determinations of peak  $\dot{V}O_2$  supported by anthropometry and estimates of maturity status. **Results:** Baseline models with body mass controlled for showed age to exert a positive effect on peak  $\dot{V}O_2$ , with negative estimates for age<sup>2</sup>, sex, and a sex-by-age interaction. Sex-specific models showed maturity status to have a positive effect on peak  $\dot{V}O_2$  in addition to the effects of age and body mass. Introducing skinfold thicknesses to provide, with body mass, a surrogate for FFM explained maturity effects and yielded a significantly ( $P < .05$ ) better statistical fit in all models compared with those based on FFM estimated from youth-specific skinfold equations. With girls only, the introduction of overweight, defined by body mass index, resulted in a small but significant ( $P < .05$ ) negative effect, with an age-by-overweight status interaction. **Conclusions:** FFM has a powerful influence on peak  $\dot{V}O_2$  in both sexes. Interpretation of the development of youth aerobic fitness and its application to health should reflect the sex- and maturity-associated variation in FFM.

**Keywords:** aerobic fitness, body mass, children, fat-free mass, maturity status, overweight

Aerobic fitness defines the ability to deliver oxygen to the muscles and to use it to generate energy to support muscle activity during exercise. Peak oxygen uptake ( $\dot{V}O_2$ ), the highest  $\dot{V}O_2$  achieved during an incremental exercise test to exhaustion, is internationally recognized as the single best indicator of youth aerobic fitness, but its interpretation in relation to sex, age, body mass, fat-free mass (FFM), and maturity status is controversial (4).

The vast majority of published data are cross-sectional and, on balance, show that boys' absolute peak  $\dot{V}O_2$  (ie, in  $L \cdot \text{min}^{-1}$ ) increases in a near-linear manner with age and girls' data show a similar trend, although some studies report a leveling-off in girls' peak  $\dot{V}O_2$  from the mid- to late teens. Peak  $\dot{V}O_2$  is, however, correlated with body mass; to control for body mass in youth, most researchers have opted to simply divide peak  $\dot{V}O_2$  by body mass and express it in  $\text{mL} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$  (ie, with a body mass exponent of 1). This is traditionally referred to as *ratio scaling* (40). Cross-sectional data indicate that boys' ratio-scaled peak  $\dot{V}O_2$  remains essentially unchanged from 10 to 18 years of age, whereas girls' values progressively decline with age, particularly from ~13 to 14 years of age (6).

Ratio scaling is "convenient and traditional" (13), but it does not have a rigorous scientific rationale, is often not statistically justified, favors lighter individuals, and penalizes heavier (overweight or more mature) youth (45,46,50). Allometric scaling has challenged the "convenient and traditional" interpretation of youth aerobic fitness and indicated that, with body mass appropriately controlled for, there is a progressive increase in peak  $\dot{V}O_2$  with age in both sexes (48). Peak  $\dot{V}O_2$  in  $\text{mL} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$  has been reported to be unrelated to maturity status (11), but with body mass controlled for using allometry, it has been demonstrated that there

are significant, sex-specific, positive effects of maturity status on peak  $\dot{V}O_2$  independent of age (9).

Fat tissue does not influence peak  $\dot{V}O_2$  (20), but it is common for overfat or overweight youth to be classified as "unfit" on the basis of peak  $\dot{V}O_2$  ratio scaled with body mass (35). "Clinical red flags" have been raised against 8- to 18-year-olds with "cut points" of peak  $\dot{V}O_2$  below 42 and 35  $\text{mL} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$  for boys and girls, respectively (24). The difference between the peak  $\dot{V}O_2$  of overfat and lean children has, however, been shown to be markedly reduced with body mass controlled for using allometry compared with values in  $\text{mL} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$  (47). Moreover, as compellingly argued by Tanner (40) 70 years ago, inappropriate ratio scaling leads to spurious correlations between peak  $\dot{V}O_2$  and other health-related variables. More recently, a comprehensive review concluded that the association of ratio-scaled peak  $\dot{V}O_2$  with cardiovascular risk factors in overweight youth may reflect overweight status to a greater extent than aerobic fitness (25).

Longitudinal data provide a more informative interpretation of the development of youth aerobic fitness, but rigorous longitudinal studies are remarkably sparse. Some studies include a small number of participants and terminate in the early teens. A few investigations only include boys or follow girls for a shorter period of time. Data interpretation often consists of an evaluation of a series of annual cross-sectional analyses of absolute and ratio-scaled peak  $\dot{V}O_2$ . Studies are difficult to interpret through the use of different methodology, equipment, and exercise testing modes, but trends are generally consistent. Taken together, longitudinal studies indicate that absolute peak  $\dot{V}O_2$  is higher in boys than girls from at least 10 years of age; boys' values increase with age through the teen years, whereas girls' peak  $\dot{V}O_2$  appears to level off from ~13 years of age. Boys' ratio-scaled peak  $\dot{V}O_2$  is higher than that of girls and remains stable with age, although a gradual decline from 12 years of age to 13 years of age has also been noted. Girls' ratio-scaled peak  $\dot{V}O_2$  decreases with age. Age at peak height velocity (PHV) has been used as a maturational benchmark in several studies; over the period 2 years prior to 2 years following PHV,

Armstrong and Welsman are with Children's Health and Exercise Research Centre, University of Exeter, Exeter, United Kingdom. Armstrong ([N.Armstrong@exeter.ac.uk](mailto:N.Armstrong@exeter.ac.uk)) is corresponding author.

both boys' and girls' mean absolute peak  $\dot{V}O_2$  has been reported to increase. By contrast, ratio-scaled peak  $\dot{V}O_2$  has been noted to decline in both sexes over the same time period, with a greater decline observed in girls. The influence of concurrent changes in body mass, FFM, and maturity status on peak  $\dot{V}O_2$  has not been addressed longitudinally, and overweight status in relation to the development of aerobic fitness has not been explored (16,18,22,23,29,34,36,38).

In the first study of youth aerobic fitness to include both boys and girls, Åstrand (12) argued that ideally, peak  $\dot{V}O_2$  should be interpreted in relation to active muscle mass rather than total body mass. Given the complications of measuring muscle mass, however, peak  $\dot{V}O_2$  should at least be scaled to FFM. It has been reported that peak  $\dot{V}O_2$  is more strongly correlated with FFM, estimated using dual-energy X-ray absorptiometry, than with body mass (21) and that the peak  $\dot{V}O_2$  of overweight youth is similar to that of lean youth when expressed in relation to FFM (20). In contrast, over the period 11 to 18 years of age, changes in FFM determined using densitometry have been reported to be of secondary importance in the development of peak  $\dot{V}O_2$  compared with changes in body mass, at least in boys (38). Rigorous analysis of the influence of FFM on aerobic fitness in youth is, however, challenging (14), and FFM measures of the same young people have been shown to vary widely across established laboratory techniques (17).

In several longitudinal studies, %body fat has been predicted from skinfold thicknesses (eg, 16,22,23,34,36), but all combinations of skinfold thicknesses and prediction equations have major limitations (14). The youth-specific equations to predict %body fat from skinfolds developed by Slaughter et al (37) are the most commonly used (eg, 22,34), but validation studies of the Slaughter equations have revealed wide limits of agreement and a tendency to underpredict fat in girls and overpredict fat in boys (32). It has been recommended that researchers should consider simply using the sum of skinfold thicknesses in conjunction with body mass as a surrogate of FFM rather than predicting %body fat from skinfold-based equations and subsequently estimating FFM (32). This proposal has not been investigated longitudinally.

Both cross-sectional and longitudinal studies have contributed to current knowledge, but variations and limitations in experimental designs and analyses have confused physiologic understanding of the development of youth aerobic fitness. The emergence (1) and refinement (19) of multilevel regression modeling has opened new analytical approaches to developmental exercise physiology. Multilevel regression modeling enables the effects of variables such as sex, age, body mass, FFM, and maturity status on peak  $\dot{V}O_2$  to be partitioned concurrently within an allometric framework to provide a flexible and sensitive interpretation of youth aerobic fitness.

In an innovative reanalysis of data from a study of elite youth athletes, Nevill et al (30) introduced multiplicative, allometric modeling to pediatric sport science and demonstrated age and—in boys but not girls—maturity status effects on peak  $\dot{V}O_2$  that were additional to those due to changes in body size. Unfortunately, issues related to selection and retention in elite youth sports and participation in different high-intensity, sport-specific training programs both before and during the study preclude extrapolation of the data to an untrained population of young people. The sex-specific development of aerobic fitness in healthy, untrained boys and girls using a longitudinal design with a multiplicative, allometric modeling approach has not been examined. The

longitudinal influence of FFM on youth aerobic fitness has rarely been explored. In addition, a multilevel modeling approach enables the efficacy of body mass and the sum of skinfold thicknesses as a surrogate of FFM to be compared with FFM estimated from body mass and %body fat predicted from youth-specific skinfold equations.

The purpose of the present study is, therefore, to enhance understanding of the development of peak  $\dot{V}O_2$  from 10 to 18 years of age in relation to sex, age, maturity status, body mass, and FFM. Subsidiary objectives are (1) to compare sex-specific models of the development of peak  $\dot{V}O_2$  founded on body mass and the sum of skinfold thicknesses as a surrogate of FFM with those based on FFM estimated from body mass and %body fat predicted from skinfold equations and (2) to explore sex-specific effects of overweight status on peak  $\dot{V}O_2$ .

## Methods

### Subjects

Initially, 234 (115 girls and 119 boys) students from year 6 (10- to 11-year-olds) in local state schools volunteered to take part in a longitudinal project involving 3 annual testing occasions. The project was subsequently extended to cover the age range 10–18 years, with an additional 71 (35 girls and 36 boys) students from year 7 (11- to 12-year-olds) and 67 (31 girls and 36 boys) students from year 9 (13- to 14-year-olds) volunteering to join the project. The study received ethical approval from the District Health Authority Ethical Committee, and all participants provided written informed consent signed by themselves and their parent or guardian. Participant recruitment and representativeness have been described in interim publications, where descriptive data and some of the initial longitudinal data have been reported as the project progressed (7,10). The complete data set has not previously been brought together for analysis; sex-specific models have not been developed, and neither FFM nor overweight status has been explored.

### Experimental Procedures

Participants visited the research center annually and were well habituated to the laboratory environment, laboratory personnel, and the specific experimental procedures. Age was computed from date of birth to date of examination. Anthropometric measures were taken as recommended by the International Biological Program (44), and apparatus was calibrated according to the manufacturers' instructions. Stature was measured by using a Holtain stadiometer (Holtain, Crymych, UK); body mass was determined by using Avery balance scales (Avery, Birmingham, UK); body mass index (BMI) was calculated as body mass  $\cdot$  stature<sup>-2</sup> (ie, kg  $\cdot$  m<sup>-2</sup>), and participants coded into the categories "overweight" and "not overweight" based on the international BMI criteria recommended by Cole et al (15). Skinfold thicknesses over the triceps and subscapular regions were measured by using Holtain skinfold calipers. Percentage body fat was predicted from skinfolds by using the equations developed by Slaughter et al (37), and FFM was subsequently estimated from body mass and predicted %body fat. On 972 occasions, maturity status was visually assessed by the Research Centre nurse using the indices for pubic hair (PH) development described by Tanner (41).

Following a standardized warm-up, peak  $\dot{V}O_2$  was determined during a progressive running test to voluntary exhaustion on a

motorized treadmill (Woodway; Cranlea Medical, Birmingham, UK). Throughout each test, heart rate and expired respiratory gases were monitored continuously using an electrocardiograph and an Oxycon Sigma online gas analysis system (Cranlea Medical), which was calibrated prior to each test by using gases of verified concentration. Depending on age, the test began at a treadmill belt speed of either  $1.94 \text{ m}\cdot\text{s}^{-1}$  ( $7 \text{ km}\cdot\text{h}^{-1}$ ) or  $2.22 \text{ m}\cdot\text{s}^{-1}$  ( $8 \text{ km}\cdot\text{h}^{-1}$ ) and increased by  $0.28 \text{ m}\cdot\text{s}^{-1}$  ( $1 \text{ km}\cdot\text{h}^{-1}$ ) for each 2- or 3-minute stage until a speed of  $2.78 \text{ m}\cdot\text{s}^{-1}$  ( $10 \text{ km}\cdot\text{h}^{-1}$ ) was reached. Subsequently, belt speed was held constant, and the gradient was incrementally increased by 2.5% until the participants were exhausted. The highest 30-second  $\dot{V}O_2$  attained was accepted as a maximal index if the subject demonstrated clear signs of intense exertion (eg, hyperpnea, facial flushing, unsteady gait, profuse sweating) supported by a heart rate that was leveling off over the final stages of the test at a value within 5% of the mean maximum heart rate we have previously reported for youth of these ages (11). All participants included herein met these criteria on each test occasion, totaling 1057 determinations of peak  $\dot{V}O_2$ .

## Data Analyses

Data were analyzed using SPSS v24 software (IBM Corp, Armonk, NY). To explore age, body mass, and body composition relationships with peak  $\dot{V}O_2$  across the data set, Pearson correlation coefficients were determined by sex between peak  $\dot{V}O_2$  and age, body mass, and estimated FFM, respectively. For descriptive purposes, peak  $\dot{V}O_2$  data were plotted against age and against body mass, respectively.

Factors associated with the longitudinal growth of peak  $\dot{V}O_2$  were analyzed using multilevel regression modeling (MLwiN version 3.02; Centre for Multilevel Modelling, University of Bristol, Bristol, UK), adopting the multiplicative, allometric approach described by Nevill et al (30). Initial (baseline) models sought to investigate sex differences across the age range 10–18 years of age, with age and body mass (or estimated FFM) controlled for within an allometric framework. Subsequently, multilevel models were computed for boys and girls separately.

The log-linearized baseline model (Equation 1) formed the starting point for analyses:

$$\log_e y = k_1 \cdot \log_e \text{mass} + a_j + b \cdot \text{age} + c \cdot \text{age}^2 + \log_e (\epsilon_{ij}). \quad (1)$$

All investigated parameters were fixed with the exception of the constant (intercept term), which was allowed to vary randomly at level 2 (between individuals), and the multiplicative error term  $\epsilon$ , which also varied randomly at level 1 (within individuals). The subscripts  $i$  and  $j$  in Equation 1 denote this random variation at levels 1 and 2, respectively. Age was centered on the group mean. From the baseline model, including age and body mass, additional explanatory variables were explored. In the initial modeling of the combined groups, sex differences were investigated using the indicator variable boys = 0, girls = 1, plus the interaction term age-by-sex. In subsequent analyses, the additional effect of the sum of triceps and subscapular skinfold thicknesses and additional indicator variables for maturity status (ie, PH stages 2, 3, 4, and 5 compared with PH stage 1) and overweight (not overweight = 0, overweight = 1) were explored. Finally, maturity status was introduced to baseline models of age and estimated FFM.

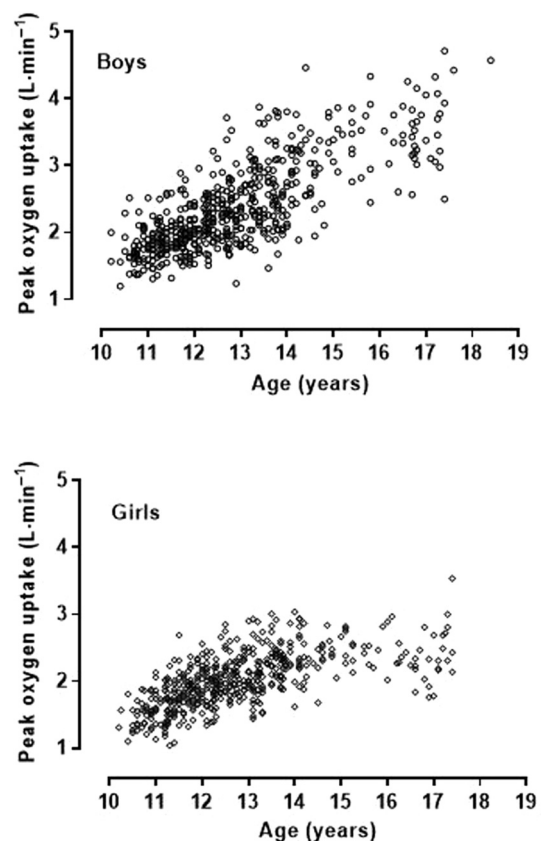
Parameter estimates were considered significant ( $P < .05$ ) where their value exceeded  $2 \times$  the standard error. The change in deviance statistic ( $-2 \times \log$  likelihood) was used to assess the goodness of fit of the models.

## Results

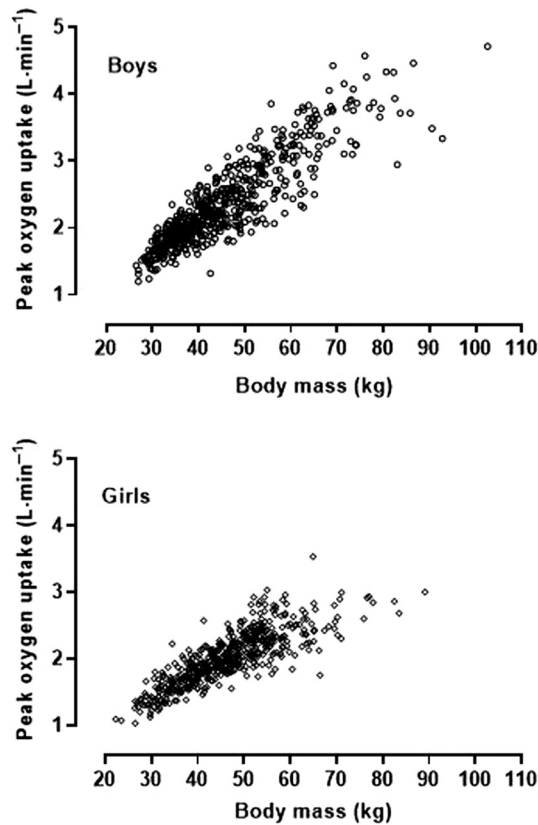
In contrast to traditional methods that require a complete longitudinal data set, both the number of observations per individual and the temporal spacing of the observations can vary within a multilevel analysis. Moreover, in the present study, there were no significant differences ( $P > .05$ ) between those who were unable to attend a subsequent test occasion and the rest of their sex-specific group in body mass, skinfold thicknesses, or peak  $\dot{V}O_2$ . The analyses were founded on 1057 (556 boys and 501 girls) measurements of age, body mass, skinfold thicknesses, and peak  $\dot{V}O_2$  supported by 972 (516 boys and 456 girls) assessments of maturity status.

The relationships across the whole data set between peak  $\dot{V}O_2$  and age and peak  $\dot{V}O_2$  and body mass by sex are illustrated in Figures 1 and 2. Age was significantly ( $P < .001$ ) correlated with peak  $\dot{V}O_2$  in both boys and girls, with coefficients of  $r = .78$  and  $r = .64$ , respectively. Body mass was significantly ( $P < .001$ ) correlated with peak  $\dot{V}O_2$  in both boys ( $r = .89$ ) and girls ( $r = .83$ ). Peak  $\dot{V}O_2$  was significantly ( $P < .001$ ) correlated with estimated FFM in boys ( $r = .94$ ) and girls ( $r = .87$ ).

To investigate sex differences in the development of aerobic fitness and whether or not ratio scaling is appropriate, the male and female data were combined in Table 1, where each model is founded on 1057 determinations of peak  $\dot{V}O_2$ . Model 1.1 confirmed body mass as a significant anthropometric covariate, yielding an allometric exponent (standard error) of 0.69 (0.02) (ie, with



**Figure 1** — Peak oxygen uptake and age in 10- to 18-year-old boys and girls.



**Figure 2** — Peak oxygen uptake and body mass in 10- to 18-year-old boys and girls.

95% confidence intervals (CIs) encompassing the range 0.65–0.73). With body mass controlled for, a positive effect of age on peak  $\dot{V}O_2$  was demonstrated, with significant negative estimates obtained for the age<sup>2</sup>, sex, and age-by-sex interaction terms. In model 1.2, the sum of subscapular and triceps skinfold thicknesses was entered to provide, with body mass, a surrogate for FFM. The introduction of skinfold thicknesses generated a negative exponent, resulting in a large increase in the body mass exponent and reductions in the positive age and negative sex, age-by-sex, and age<sup>2</sup> terms. Model 1.3 confirmed estimated FFM as a significant covariate. With this covariate controlled for, a positive effect of age on peak  $\dot{V}O_2$  remained, although smaller than that noted in model 1.1. Similar to models 1.1 and 1.2, significant negative estimates were obtained for the age<sup>2</sup>, sex, and age-by-sex interaction terms. A comparison of the goodness of fit of the different models is obtained from the  $-2 \times \log$ -likelihood, with reference to the number of fitted parameters. In a comparison of 2 models with the same number of fitted parameters, the model with the smallest  $-2 \times \log$ -likelihood reflects that with the best fit. Additional parameters contribute to improved fit from the change in  $-2 \times \log$  likelihood according to a chi-square statistic for additional degrees of freedom. Thus, in Table 1, the model with the best fit is clearly model 1.2 (with body mass and skinfolds as covariates).

Table 2 contains the models describing the boys' data. Model 2.1, the baseline model, shows that with body mass controlled for, age had a significant positive effect on peak  $\dot{V}O_2$ , with a small but significant negative effect of age<sup>2</sup>. In model 2.2, maturity status (PH development) was entered into the baseline model, and PH stages 2 to 5 each made a significant positive contribution in addition to age and body mass in explaining peak  $\dot{V}O_2$ . With the introduction of the sum of triceps and subscapular skinfolds, the

**Table 1** Multilevel Regression Models to Examine Longitudinal Sex Differences in Peak Oxygen Uptake

	Model 1.1		Model 1.2		Model 1.3	
	<u>Log<sub>e</sub> peak oxygen uptake</u>	<u>SE</u>	<u>Log<sub>e</sub> peak oxygen uptake</u>	<u>SE</u>	<u>Log<sub>e</sub> peak oxygen uptake</u>	<u>SE</u>
Fixed part						
constant	-1.757	0.085	-2.123	0.076	-2.251	0.088
log <sub>e</sub> body mass	0.686	0.022	0.910	0.024	–	
age <sup>a</sup>	0.096	0.006	0.055	0.006	0.044	0.006
sex	-0.162	0.009	-0.129	0.008	-0.099	0.008
age-by-sex	-0.037	0.003	-0.023	0.003	-0.015	0.003
age squared	-0.007	0.001	-0.004	0.001	-0.004	0.001
log <sub>e</sub> skinfolds	–		-0.167	0.011	–	
log <sub>e</sub> FFM <sup>b</sup>	–		–		0.861	0.024
Random part						
level 2						
var (cons)	0.006	0.001	0.004	0.000	0.004	0.000
level 1						
var (cons)	0.004	0.000	0.004	0.000	0.005	0.000
$-2 \times \log$ -likelihood	-2126.339		-2330.576		-2200.463	

Abbreviations: FFM, fat-free mass; SE, standard error; var, variance; cons, constant; –, not entered into model.

<sup>a</sup>Centered on mean age 12.8 years.

<sup>b</sup>FFM computed from the equations of Slaughter et al. (37).

**Table 2** Multilevel Regression Models for Peak Oxygen Uptake in Boys

	Model 2.1		Model 2.2		Model 2.3		Model 2.4	
	Log <sub>e</sub> peak oxygen uptake		Log <sub>e</sub> peak oxygen uptake		Log <sub>e</sub> peak oxygen uptake		Log <sub>e</sub> peak oxygen uptake	
	SE		SE		SE		SE	
Fixed part								
constant	-1.861	0.121	-1.694	0.123	-2.273	0.099	-2.300	0.114
log <sub>e</sub> body mass	0.713	0.032	0.655	0.033	0.964	0.031	-	
age <sup>a</sup>	0.051	0.005	0.031	0.005	0.023	0.004	0.025	0.005
age squared	-0.004	0.001	ns		-0.003	0.001	-0.003	0.001
pubic hair								
stage 2	-		0.030	0.011	ns		ns	
stage 3	-		0.063	0.013	ns		ns	
stage 4	-		0.091	0.015	ns		ns	
stage 5	-		0.091	0.023	ns		ns	
log <sub>e</sub> skinfolds	-		-		-0.185	0.013	-	
log <sub>e</sub> FFM <sup>b</sup>	-		-		-		0.874	0.032
overweight	-		-		ns		-	
overweight by age	-		-		ns		-	
Random part								
level 2								
var (cons)	0.007	0.001	0.006	0.001	0.003	0.000	0.004	0.001
level 1								
var (cons)	0.005	0.000	0.004	0.001	0.004	0.000	0.005	0.000
-2 × log-likelihood	-1088.073		-1028.937		-1238.092		-1154.915	

Abbreviations: FFM, fat-free mass; ns, not significant ( $P > .05$ ); SE, standard error; var, variance; cons, constant; -, not entered into model.

<sup>a</sup>Centered on mean age 12.9 years.

<sup>b</sup>FFM computed from the equations of Slaughter et al. (37).

effects of maturity status became nonsignificant. Model 2.3, founded on the whole-male data set, shows the effects of body mass, age, age<sup>2</sup>, and skinfold thicknesses on peak  $\dot{V}O_2$ .

Model 2.4 is similar to model 2.3, showing that with estimated FFM controlled for, maturity status was not significant and age and age<sup>2</sup> had significant effects on peak  $\dot{V}O_2$ . Comparing models 2.3 and 2.4 for goodness of fit based on the  $-2 \times \log$  likelihood, model 2.3, with body mass and skinfolds as covariates, provided a better description of developmental changes in boys' peak  $\dot{V}O_2$  than model 2.4 with estimated FFM as covariate. The addition of overweight as a predictor variable yielded a nonsignificant parameter estimate in model 2.3.

The models describing the girls' data are presented in Table 3. Model 3.1, the baseline model, shows that with body mass controlled for, age has a significant positive effect on peak  $\dot{V}O_2$ , with a significant negative age<sup>2</sup> term. When maturity status was entered into the baseline model as model 3.2, PH stages 2 to 5 each made a significant contribution to explaining peak  $\dot{V}O_2$  in addition to body mass, age, and age<sup>2</sup>. The entry of the sum of skinfolds negated the effects of maturity status. Model 3.3, which was founded on the whole-female data set, shows the effects of body mass and skinfold thicknesses on peak  $\dot{V}O_2$ . As observed for both the combined data (Table 1) and boys' data (Table 2), the model with the best statistical fit for girls' peak  $\dot{V}O_2$  is that with body mass and skinfolds as covariates (model 3.3) rather than with estimated FFM as a covariate (model 3.5). The introduction of overweight in model 3.4 yielded a significant

negative estimate, with a significant negative age-by-overweight interaction.

## Discussion

The raw data presented in Figures 1 and 2 support indications from previous cross-sectional and longitudinal studies (5). In boys, there is a near-linear increase in peak  $\dot{V}O_2$  with age, indicating that between 16 and 18 years of age, the age effect is reduced as the rate of change in growth decreases. In girls, a near-linear increase is observed until ~13 to 14 years of age, followed by a leveling-off in values. Plotted against body mass, the size dependency of peak  $\dot{V}O_2$  is evident, as are sex differences throughout the range of values observed.

The baseline models clearly illustrate the fallacy of the ratio-scaled peak  $\dot{V}O_2$  interpretation of age-related aerobic fitness. Allometric exponents for body mass are 0.69 (model 1.1), 0.71 (model 2.1), and 0.63 (model 3.1), with an exponent of 1.0 (as assumed in ratio scaling) falling outside the 95% CI on each occasion. The positive age, negative sex, and negative age-by-sex terms in Table 1 show that with either body mass (with or without skinfolds) or estimated FFM controlled for, peak  $\dot{V}O_2$  is higher in boys than in girls and increases with age in both sexes, although the age effect is smaller in girls. The negative age<sup>2</sup> term indicates that the magnitude of the age effect is reduced as the rate of change in growth decreases. Model 1.3 is a better statistical fit

**Table 3** Multilevel Regression Models for Peak Oxygen Uptake in Girls

Response	Model 3.1		Model 3.2		Model 3.3		Model 3.4		Model 3.5	
	Log <sub>e</sub> peak oxygen uptake		Log <sub>e</sub> peak oxygen uptake		Log <sub>e</sub> peak oxygen uptake		Log <sub>e</sub> peak oxygen uptake		Log <sub>e</sub> peak oxygen uptake	
	SE		SE		SE		SE		SE	
Fixed part										
constant	-1.701	0.119	-1.657	0.127	-2.004	0.117	-2.204	0.128	-2.215	0.142
log <sub>e</sub> body mass	0.631	0.031	0.609	0.034	0.815	0.038	0.857	0.039	-	
age <sup>a</sup>	0.035	0.004	0.024	0.006	0.020	0.005	0.017	0.005	0.022	0.005
age squared	-0.010	0.001	-0.008	0.001	-0.007	0.001	-0.007	0.001	-0.006	0.001
pubic hair										
stage 2	-		0.038	0.013	ns		-		ns	
stage 3	-		0.046	0.015	ns		-		ns	
stage 4	-		0.052	0.018	ns		-		ns	
stage 5	-		0.055	0.023	ns		-		ns	
log <sub>e</sub> skinfolds	-		-		-0.129	0.018	-0.114	0.018	-	
log <sub>e</sub> FFM <sup>b</sup>	-		-		-		-		0.824	0.040
overweight	-		-		-		-0.054	0.015	-	
overweight by age	-		-		-		-0.015	0.006	-	
Random part										
level 2										
var (cons)	0.006	0.001	0.006	0.001	0.004	0.001	0.004	0.001	0.005	0.001
level 1										
var (cons)	0.004	0.000	0.004	0.000	0.004	0.000	0.004	0.000	0.004	0.000
-2 × log-likelihood	-1060.443		-951.197		-1107.768		-1129.861		-1052.430	

Abbreviations: FFM, fat-free mass; ns, not significant ( $P > .05$ ); SE, standard error; var, variance; cons, constant; -, not entered into model.

<sup>a</sup>Centered on mean age 12.8 years.

<sup>b</sup>FFM computed from the equations of Slaughter et al. (37).

than model 1.1, but model 1.2, founded on body mass and skinfold thicknesses (as a surrogate for FFM), is a significantly ( $P < .05$ ) better fit than either model 1.1 or model 1.3. This trend is consistent through the sex-specific models in Tables 2 and 3, with baseline models founded on estimated FFM significantly ( $P < .05$ ) superior to those based on body mass. The models with the best fit ( $P < .05$ ), however, are those founded on body mass and the sum of skinfold thicknesses. Therefore, the discussion that follows will focus on model 2.3 and model 3.3 when interpreting the influence of FFM on the development of aerobic fitness.

It is well documented that age-related increases in youth peak  $\dot{V}O_2$  are the result of increases in maximal stroke volume ( $SV_{max}$ ) and maximal arteriovenous oxygen difference (a- $vO_2$  diff max). The reasons for sex differences in the development of aerobic fitness are less clear but generally attributed to differences in physical activity,  $SV_{max}$ , or a- $vO_2$  diff max (5).

Physical activity tends to decrease through the teen years, and boys are generally more physically active than girls, but there is no compelling evidence to suggest that current levels of daily physical activity contribute to age or sex differences in peak  $\dot{V}O_2$  (3). To test whether there was any relationship between daily physical activity and aerobic fitness in the present study, the physical activity patterns of ~200 of the participants were monitored annually from 11 to 13 years. The data showing that no significant ( $P > .05$ ) relationship between physical activity and peak  $\dot{V}O_2$  exists has been published elsewhere (8), but in summary, with age and body mass controlled for, peak  $\dot{V}O_2$  was introduced to a

multilevel regression model of physical activity as an additional variable and a nonsignificant ( $P > .05$ ) parameter estimate generated. Physical activity of moderate intensity and above significantly ( $P < .05$ ) decreased with age, whereas peak  $\dot{V}O_2$  in both  $L \cdot min^{-1}$  and with body mass controlled for using allometry significantly ( $P < .05$ ) increased over the same time period in both girls and boys.

The introduction of noninvasive technologies to developmental exercise physiology has promoted research into comparisons of the components of peak  $\dot{V}O_2$  in girls and boys, but data are sparse and conflicting. Two Doppler echocardiography studies have reported sex differences to be the result of a higher  $SV_{max}$  in boys, but whether this is because of differences in cardiac size or function is contentious (33,43). By contrast, a study using thoracic bioelectrical impedance reported that a 19% sex difference observed in mean peak  $\dot{V}O_2$  was caused by a 17% difference in a- $vO_2$  diff max with no significant sex difference in  $SV_{max}$ . The same study explored resting heart size variables by using magnetic resonance imaging and reported no significant sex differences (49). A poorer matching of muscle oxygen delivery to oxygen utilization in girls was reported in a study that used near-infrared spectroscopy to measure microcirculatory changes in deoxygenated hemoglobin and myoglobin. The authors concluded that this phenomenon may contribute to girls' lower peak  $\dot{V}O_2$  (28). However, although the physiologic mechanisms underlying sex differences in youth aerobic fitness are emerging, they have yet to be fully elucidated (5).

Girls typically enter puberty before similarly aged boys; for example, a 12-year-old girl in PH stage 3 is not equivalent to a 14-year-old boy at the same pubertal stage (26). It is therefore appropriate to analyze data in sex-specific models (Tables 2 and 3), which provide more sensitive explorations of the development of girls' and boys' peak  $\dot{V}O_2$  than the combined data in Table 1.

The sex-specific baseline models are in direct conflict with the interpretation of the development of aerobic fitness based on ratio-scaled data. Both model 2.1 (boys) and model 3.1 (girls) demonstrate that with body mass controlled for, age exerts a significant additional effect on peak  $\dot{V}O_2$ . Models 2.2 (boys) and 3.2 (girls) show significant additional effects of maturity status on peak  $\dot{V}O_2$  with age and body mass controlled for. However, when skinfold thicknesses were entered (model 2.3 and model 3.3 for boys and girls, respectively), a negative estimate was obtained, the value of the body mass exponent markedly increased, and the effects of maturity status on peak  $\dot{V}O_2$  were negated. The increase in the body mass exponent with the introduction of skinfold thicknesses has been previously observed and attributed to the effect that excess fat mass has on increasing body mass without an increase in peak  $\dot{V}O_2$  (42). Model 2.3 and model 3.3 illustrate this effect in the complete sex-specific data sets. These findings clearly demonstrate the importance of maturational changes in FFM on youth peak  $\dot{V}O_2$  in both sexes.

Increases in peak  $\dot{V}O_2$  are driven by gains in  $SV_{\max}$ ,  $a-vO_2$  diff max, or both, but growth in muscle mass, reflected by increases in FFM, not only enhances total muscle  $\dot{V}O_2$  during exercise but, through the peripheral muscle pump, also augments venous return to the heart and boosts  $SV_{\max}$ . FFM increases by ~90% and ~40% in boys and girls, respectively, from 11 to 16 years of age. The influence of sex-specific maturity status is evidenced by percentage changes in FFM being at their peak around the time of PHV. Boys' FFM increases by ~83% over the period 2 years pre-PHV to 2 years post-PHV. The greatest increase in girls' FFM (~31%) occurs over a shorter 2-year period centered on PHV, and then levels off in accordance with the development of peak  $\dot{V}O_2$  (2,27).

The introduction of overweight as an indicator variable yielded a nonsignificant parameter estimate in boys (model 2.3) confirming the ability of body mass and skinfolds, as surrogates for FFM, to account for any differences in peak  $\dot{V}O_2$  between boys classified by BMI for age as overweight (15) vs those of healthy mass. Interestingly, in girls (model 3.4), this was not the case, and girls defined as overweight by BMI for age presented a peak  $\dot{V}O_2$ , adjusted for body mass and skinfold thicknesses that was ~5% smaller and statistically significantly different from girls of healthy body mass. A significant negative age-by-overweight interaction indicated an additional small (~2%) decrement with age. We suspect that these findings result not from a true difference in peak  $\dot{V}O_2$  between the groups but rather from the use of BMI to define assumed levels of excess adipose tissue. BMI was designed as a population metric but is currently used with relatively small samples—at an individual level, an application that is inherently flawed (39). Data from adults have cast serious doubts on the validity of BMI to represent adiposity compared with surface anthropometry such as skinfolds (31). Specifically, a ratio of mass to stature<sup>2</sup> cannot reflect maturational changes in fat accumulation and patterning in girls. We suggest that classifying young people as overweight by using a direct measure of adiposity may better enable more valid comparisons of peak  $\dot{V}O_2$  in relation to overweight status.

Taken collectively, the data show that the development of the peak  $\dot{V}O_2$  of overweight boys and girls is closely related to their

FFM, as in healthy weight youth. However, their elevated fat mass and body mass mean that they are highly likely to be graded as “unfit” or to raise “clinical red flags” in current classification systems based on ratio scaling with body mass (24,35).

In conclusion, a multiplicative allometric approach, applied to longitudinal data, has demonstrated that in both sexes with body mass controlled for, peak  $\dot{V}O_2$  increases with age, but the effect is smaller in girls than in boys. In both boys and girls, maturity status exerts a positive independent effect on peak  $\dot{V}O_2$  in addition to that of age and body mass, but it is the changes in FFM, as indicated herein by body mass and skinfolds, associated with changes in maturity status that are the principal influence on the increase of peak  $\dot{V}O_2$  in youth. The present data are in direct conflict with the ratio-scaled interpretation of youth peak  $\dot{V}O_2$  and expose the fallacy of using single peak  $\dot{V}O_2$  data points in  $\text{mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$  as “cut points” for acceptable levels of aerobic fitness or for raising “clinical red flags” for youth falling between 10 and 18 years of age. Moreover, the interpretation of youth aerobic fitness and its rigorous application to health and well-being should reflect the maturity status- and sex-associated variation in FFM, which we have demonstrated herein to be a powerful influence on the development of peak  $\dot{V}O_2$  in both sexes and in both overweight and healthy-weight youth. The use of body mass and skinfold thicknesses as a surrogate for FFM has been revealed to provide a model with a superior statistical fit than the application of FFM estimated from body mass and %body fat predicted from youth-specific skinfold equations, and this technique can be recommended for future research.

## Acknowledgments

The authors gratefully acknowledge the commitment and loyalty of the participants, the logistic support of all schools involved, and the technical support of the Children's Health and Exercise Research Centre team. The work was funded by the British Heart Foundation, the Community Fund, and the Darlington Trust.

## References

1. Aitkin M, Anderson D, Hinde J. Statistical modelling of data on teaching styles. *J Roy Stat Soc A*. 1981;144:419–61. doi:10.2307/2981826
2. Armstrong N. *Development of the Youth Athlete*. Oxford, UK: Routledge;2018:5–26.
3. Armstrong N. Paediatric physical activity and aerobic fitness. In: Draper N, Stratton G, eds. *Physical Activity: A Multi-Disciplinary Approach*. Oxford, UK: Routledge; 2018:186–204.
4. Armstrong N. Top 10 research questions related to youth aerobic fitness. *Res Quart Exerc Sport*. 2017;88:130–48. PubMed ID: 28402178 doi:10.1080/02701367.2017.1303298
5. Armstrong N, McManus AM. Aerobic fitness. In: Armstrong N, van Mechelen W, eds. *Oxford Textbook of Children's Sport and Exercise Medicine*. 3rd ed. Oxford, UK: Oxford University Press; 2017:161–80.
6. Armstrong N, Welsman JR. Assessment and interpretation of aerobic fitness in children and adolescents. *Exerc Sport Sci Rev*. 1994;22:435–76. PubMed ID: 7925551 doi:10.1249/00003677-199401000-00016
7. Armstrong N, Welsman JR. Peak oxygen uptake in relation to growth and maturation in 11-17-year-old humans. *Eur J Appl Physiol*. 2001;85:546–51. PubMed ID: 11718283 doi:10.1007/s004210100485

8. Armstrong N, Welsman JR, Kirby BJ. Longitudinal changes in 11-13-year-olds' physical activity. *Acta Paediatr.* 2000;89:775–80. PubMed ID: [10943956](#) doi:[10.1111/j.1651-2227.2000.tb00384.x](#)
9. Armstrong N, Welsman JR, Kirby BJ. Peak oxygen uptake and maturation in 12 year-olds. *Med Sci Sports Exerc.* 1998;30:165–9. PubMed ID: [9475659](#) doi:[10.1097/00005768-199801000-00023](#)
10. Armstrong N, Welsman JR, Nevill AM, Kirby BJ. Modeling growth and maturation changes in peak oxygen uptake in 11-13 yr olds. *J Appl Physiol.* 1999;87:2230–6. PubMed ID: [10601172](#) doi:[10.1152/jappl.1999.87.6.2230](#)
11. Armstrong N, Williams J, Balding J, Gentle P, Kirby B. Peak oxygen uptake of British children with reference to age, sex and sexual maturity. *Eur J Appl Physiol.* 1991;62:369–75. PubMed ID: [1874245](#) doi:[10.1007/BF00634975](#)
12. Åstrand PO. *Experimental Studies of Physical Working Capacity in Relation to Sex and Age.* Copenhagen, Denmark: Munksgaard; 1952.
13. Bar-Or O, Rowland TW. *Pediatric Exercise Medicine.* Champaign, IL: Human Kinetics; 2004.
14. Claessens AL, Beunen G, Malina RM. Anthropometry, physique, body composition, and maturity. In: Armstrong N, van Mechelen W, eds. *Paediatric Exercise Science and Medicine.* 2nd ed. Oxford, UK: Oxford University Press; 2008:24–36.
15. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *Br Med J.* 2000;320:1240–3. PubMed ID: [10797032](#) doi:[10.1136/bmj.320.7244.1240](#)
16. Cunningham DA, Paterson DH, Blimkie CJR, Donner AP. Development of cardiorespiratory function in circumpubertal boys: a longitudinal study. *J Appl Physiol.* 1984;56:302–7. PubMed ID: [6706741](#) doi:[10.1152/jappl.1984.56.2.302](#)
17. Ferri-Morales A, Nascimento-Ferreira MV, Vlachopoulos D, et al. Agreement between standard body composition methods to estimate percentage of body fat in young male athletes. *Pediatr Exerc Sci.* 2018;30:402–10. PubMed ID: [29543127](#) doi:[10.1123/pes.2017-0171](#)
18. Geithner CA, Thomis MA, Vanden Eynde B, et al. Growth in peak aerobic power during adolescence. *Med Sci Sports Exerc.* 2004;36:1616–24. PubMed ID: [15354046](#) doi:[10.1249/01.MSS.0000139807.72229.11](#)
19. Goldstein H, Rasbash J, Plewis I, et al. *A User's Guide to MLwiN.* London, UK: University of London Institute of Education; 1998.
20. Goran M, Fields DA, Hunter GR, Herd SL, Weinsten RL. Total body fat does not influence maximal aerobic capacity. *Int J Obes Relat Metab Disord.* 2000;24:841–8. PubMed ID: [10918530](#) doi:[10.1038/sj.ijo.0801241](#)
21. Graves LEF, Nevill AM, Fowweather L, et al. Scaling of peak oxygen uptake in children. A comparison of body size index models. *Med Sci Sports Exerc.* 2013;45:2341–5. PubMed ID: [23698249](#) doi:[10.1249/MSS.0b013e31829bfa79](#)
22. Janz KF, Burns TL, Witt JD, Mahoney LT. Longitudinal analysis of scaling for differences in body size during puberty: the Muscatine study. *Med Sci Sports Exerc.* 1998;30:1436–44. PubMed ID: [9741614](#)
23. Kemper HCG, Verschuur R. Maximal aerobic power. *Med Sport Sci.* 1985;20:107–26. doi:[10.1159/issn.0076-6070](#)
24. Lang JJ, Tomkinson GR, Janssen I, et al. Making a case for cardiorespiratory fitness surveillance among children and youth. *Exerc Sport Sci Rev.* 2018;46:66–75. PubMed ID: [29346159](#) doi:[10.1249/JES.0000000000000138](#)
25. Loftin M, Sothorn M, Abe T, Bonis M. Expression of VO<sub>2</sub> peak in children and youth with special reference to allometric scaling. *Sports Med.* 2016;46:1451–60. PubMed ID: [27139725](#) doi:[10.1007/s40279-016-0536-7](#)
26. Malina RM. Assessment of maturation. In: Armstrong N, van Mechelen W, eds. *Oxford Textbook of Children's Sport and Exercise Medicine.* 3rd ed. Oxford, UK: Oxford University Press; 2017:3–10.
27. Malina RM, Bouchard C, Bar-Or O. *Growth, Maturation, and Physical Activity.* 2nd ed. Champaign, IL: Human Kinetics; 2004: 113–6.
28. McNarry MA, Farr C, Middlebrooke A, et al. Aerobic function and muscle deoxygenation dynamics during ramp exercise in children. *Med Sci Sports Exerc.* 2015;47:1877–84. PubMed ID: [25551403](#) doi:[10.1249/MSS.0000000000000609](#)
29. Mirwald RL, Bailey DA. *Maximal Aerobic Power.* London, Ontario: Sports Dynamics; 1986:1–80.
30. Nevill AM, Holder RL, Baxter-Jones A, Round JM, Jones DA. Modelling developmental changes in strength and aerobic power in children. *J Appl Physiol.* 1998;84:963–70. PubMed ID: [9480958](#) doi:[10.1152/jappl.1998.84.3.963](#)
31. Nevill AM, Stewart AD, Olds T, Holder RL. Relationship between adiposity and body size reveals limitations of BMI. *Am J Phys Anthropol.* 2006;129:151–6. PubMed ID: [16270304](#) doi:[10.1002/\(ISSN\)1096-8644](#)
32. Roemmich JN, Clark PA, Weltman A, Rogol AD. Alterations in growth and body composition during puberty. I. Comparing multicompartiment body composition models. *J Appl Physiol.* 1997;83:927–35. PubMed ID: [9292482](#) doi:[10.1152/jappl.1997.83.3.927](#)
33. Rowland T, Goff D, Martel L, Ferrone L. Influence of cardiac functional capacity on gender differences in maximal oxygen uptake in children. *Chest.* 2000;117:629–35. PubMed ID: [10712984](#) doi:[10.1378/chest.117.3.629](#)
34. Rowland T, Vanderburgh P, Cunningham L. Body size and the growth of maximal aerobic power in children: a longitudinal analysis. *Pediatr Exerc Sci.* 1997;9:262–74. doi:[10.1123/pes.9.3.262](#)
35. Ruiz JR, Cavero-Redondo I, Ortega FB, Welk GJ, Andersen LB, Martinez-Vizcaino V. Cardiorespiratory fitness cut points to avoid cardiovascular disease risk in children and adolescents; what level of fitness should raise a red flag? A systematic review and meta-analysis. *Br J Sports Med.* 2016;50:1451–8. PubMed ID: [27670254](#) doi:[10.1136/bjsports-2015-095903](#)
36. Rutenfranz J, Andersen KL, Seliger V, Klimmer F, Berndt I, Ruppel M. Maximum aerobic power and body composition during the puberty growth period: similarities and differences between children of two European countries. *Eur J Pediatr.* 1981;136:122–33. PubMed ID: [7227388](#)
37. Slaughter MH, Loham TG, Boileau RA, et al. Skinfold equations for estimation of body fatness in children and youth. *Hum Biol.* 1998;60:709–23. PubMed ID: [3224965](#)
38. Sprynarova S, Parizkova J, Bunc S. Relationships between body dimensions and resting and working oxygen consumption in boys aged 11 to 18 years. *Eur J Appl Physiol Occup Physiol.* 1987;56:725–36. PubMed ID: [3678228](#) doi:[10.1007/BF00424817](#)
39. Sweeting H. Measurement and definitions of obesity in childhood and adolescence: a field guide for the uninitiated. *Nutr J.* 2007; 6:32. PubMed ID: [17963490](#) doi:[10.1186/1475-2891-6-32](#)
40. Tanner JM. Fallacy of per-weight and per-surface area standards and their relation to spurious correlation. *J Appl Physiol.* 1949;2:1–15. PubMed ID: [18133122](#) doi:[10.1152/jappl.1949.2.1.1](#)
41. Tanner JM. *Growth at Adolescence.* 2nd ed. Oxford, UK: Blackwell; 1962:28–39.
42. Vanderburgh PM, Mahar MT, Chou CH. Allometric scaling of grip strength by body mass in college-age men and women. *Res Q Exerc Sport.* 1995;66:80–4. PubMed ID: [7777699](#) doi:[10.1080/02701367.1995.10607658](#)



43. Vinet A, Mandigout S, Nottin S, et al. Influence of body composition, hemoglobin concentration, and cardiac size and function on gender differences in maximal oxygen uptake in prepubertal children. *Chest*. 2003;124:1494–9. PubMed ID: [14555585](#) doi:[10.1378/chest.124.4.1494](#)
44. Weiner JS, Lourie JA. *Practical Human Biology*. London, UK: Academic Press; 1981:33–51.
45. Welsman J, Armstrong N. Interpreting aerobic fitness in youth: the fallacy of ratio scaling. *Pediatr Exerc Sci*. doi:[10.1123/pes.2018-0141](#).
46. Welsman JR, Armstrong N. Interpreting exercise performance data in relation to body size. In: Armstrong N, van Mechelen W, eds. *Paediatric Exercise Science and Medicine*. 2nd ed. Oxford, UK: Oxford University Press; 2008:13–22.
47. Welsman JR, Armstrong N, Kirby BJ. Scaling the aerobic fitness of overfat children (Abstract). *J Sport Sci*. 1997;15:34.
48. Welsman JR, Armstrong N, Kirby BJ, Nevill AM, Winter EM. Scaling peak  $\dot{V}O_2$  for differences in body size. *Med Sci Sports Exerc*. 1996;28:259–65. PubMed ID: [8775163](#) doi:[10.1097/00005768-199602000-00016](#)
49. Winsley RJ, Fulford J, Roberts AC, Welsman JR, Armstrong N. Sex difference in peak oxygen uptake in prepubertal children. *J Sci Med Sport*. 2009;12:647–51. PubMed ID: [18768359](#) doi:[10.1016/j.jsams.2008.05.006](#)
50. Winter EM. Importance and principles of scaling for size differences. In: Bar-Or O, ed. *The Child and Adolescent Athlete*. Oxford, UK: Blackwell; 1996:673–9.