BMJ Open Sport & Exercise Medicine

Should we use fat-free mass or body mass and percentage body fat as separate predictors to predict maximum oxygen uptake?

Alan Nevill ,[1] Matthew Harber,[2] Matthew Wyon ,[3,4] Jonathan Myers ,[5] Ross Arena,[6] Tony Myers,[7] Leonard A Kaminsky[8]

To cite: Nevill A, Harber M, Wyon M, *et al*. Should we use fat-free mass or body mass and percentage body fat as separate predictors to predict maximum oxygen uptake? *BMJ Open Sport & Exercise Medicine* 2026;**12**:e003029. doi:10.1136/bmjsem-2025-003029

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjsem-2025-003029).

Accepted 15 February 2026

For numbered affiliations see end of article.

Correspondence to
Professor Alan Nevill;
a.m.nevill@wlv.ac.uk

ABSTRACT

Background/objective To assess whether fat-free mass (FFM) or body mass (M) is the more appropriate body size variable to predict maximum oxygen uptake (VO_{2max}, $L{\cdot}min^{-1}$).

Methods Data (3930 cardiopulmonary exercise tests) were provided from the FRIEND registry. Our prediction equations adopted the well-known allometric/power function model VO_{2max} ($L{\cdot}min^{-1}$)=$a{\cdot}X^b$, using either FFM or M as the predictor variable (X). These models can be linearised with a log-transformation, and analysis of covariance (ANCOVA) can then be used to estimate the unknown parameters.

Results Initially, when predicting Ln(VO_{2max}) using only Ln(FFM) adjusted for age and sex, the explained variance was R^2=0.718 (Akaike information criterion (AIC)=−1882.5), with the FFM exponent b=0.658. However, when predicting Ln(VO_{2max}) using M AND bodyfat% separately, the explained variance increased to R^2=0.733 (AIC=−2077.4), with the M exponent b=0.636. The difference in R^2 and AICs confirmed the benefit of predicting VO_{2max} using separate M and bodyfat% terms. The analysis identified an enlarged negative bodyfat% term that improved the prediction of VO_{2max}, explained latterly by central adiposity (waist circumference). These final, more inclusive M and FFM exponents were estimated to be b=0.67, suggesting that VO_{2max} should be normalised using VO_{2max} ($mL{\cdot}FFM^{-2/3}{\cdot}min^{-1}$) or *preferably* VO_{2max} ($mL{\cdot}M^{-2/3}{\cdot}min^{-1}$) rather than VO_{2max} ($mL{\cdot}FFM^{-1}{\cdot}min^{-1}$). We also found that linear prediction models systematically *under-estimate* the VO_{2max} of overweight and underweight individuals, but *over-estimate* the VO_{2max} of average-weight individuals.

Conclusion Incorporating FFM into equations to predict VO_{2max} fails to explain the negative effect of central adiposity. However, by incorporating M and percentage body fat (BF%) separately into the allometric models, a greater/enlarged negative BF% term explains this apparent omission/absence.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Cardiorespiratory/aerobic fitness (maximum oxygen uptake (VO_{2max})) is an important indicator of health and longevity. Predictive equations have been fundamental in large population studies where the direct measurement of VO_{2max} is not feasible.

WHAT THIS STUDY ADDS

⇒ Our results demonstrate the importance of including both body mass and body fat (%) separately rather than combining both and then using fat-free mass alone to predict VO_{2max}. An enlarged negative body fat (%) term, obtained when fitted separately, was subsequently explained by excess central adiposity (waist circumference).

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ With the global increase in obesity, especially in Western countries, the inclusion of body mass and body fat (%) separately when predicting VO_{2max} provides a more flexible/robust prediction model.

INTRODUCTION

Over 20 years ago, Nevill *et al* explained why an allometric or power function model ($Y=a{\cdot}X^b$) was theoretically, physiologically and empirically superior to scale physiological variables known to vary with body size, compared with either simple ratio standards or linear regression adjustment methods.[1] Based on the fitted allometric/power function models, the best method of scaling maximum oxygen uptake (VO_{2max}), peak and mean power output (W) to be *independent* of body size, required these variables to be divided by body mass (M), recorded in the units $kg^{2/3}$. Hence, the power function ratio standards ($mL.kg^{-2/3}.min^{-1}$) and ($W.kg^{-2/3}$) were best able to describe a wide range of participants in terms of their ability to use oxygen or record power maximally, *independent* of body size. Theoretical support for this model comes from Åstrand and Rodahl[2] who provided a plausible explanation as to why VO_{2max} ($L.min^{-1}$) of 'trained' athletes should scale to body mass $M^{2/3}$ or L^2 (L being a linear dimension of body size).

However, there has been considerable debate in the literature as to whether M or fat-free mass (FFM=M×(1−BF%/100), BF%=percentage body fat) should be used as the body size normalising variable. Goran *et al*[3] argued that because the major influence of M on VO_{2max} was explained by FFM while fat mass (FM) has no effect on VO_{2max}, VO_{2max} should be expressed per FFM (kg) as a more appropriate normalising methodology. Recently, Imboden *et al* also argued that VO_{2max} should be normalised to FFM using VO_{2max} (mL·kg FFM^{-1}·min^{-1}) to better predict all-cause mortality.[4] Indeed, Lolli *et al* published a meta-analysis of allometric studies that concluded "against the normalization of VO_{2max} using whole body mass (M) and underlined the validity of normalising using fat-free mass (FFM)".[5] They argued that their large sample of study cohorts identified a robust estimate of the expected range of true FFM exponents 'b', the point estimate given by b=0.90 (95% prediction interval: 0.68–1.12).

However, estimating FFM has its limitations. FFM is a heterogeneous component, and internal organs with high resting metabolic rates (eg, liver, kidneys, brain) do not scale with or directly contribute to the oxygen consumption of locomotor muscle during whole-body exercise.[6] Therefore, VO_{2max} is more closely related to the mass of active skeletal muscle rather than to total FFM.[7] The accuracy of FFM measurement also depends on the method (DEXA (dual-energy X-ray absorptiometry), bioelectrical impedance analysis (BIA), skinfolds (SF)), each with its own error;[8] DEXA, considered the gold standard in FFM measurement, has limitations based on software algorithms and athlete hydration effects.[9] This adds noise to estimated FFM.

In a letter to the editor, Nevill and Wyon[10] identified an oversight with the study by Lolli and coworkers.[5] Nevill and Wyon[8 10] asked the following research question—"what is the physiological/ biological explanation behind the fat-free mass exponent estimated to be 0.90"? As far as we are aware, there is no physiological/biological explanation. However, Nevill and Wyon[10] offered a simple alternative justification for Lolli *et al*'s proposed exponent of 0.90.[5] They stated that the allometric models used to predict VO_{2max} will be explicable only if the all-important confounding variables have been explained and removed. They recognise that there are at least four well-known confounders associated with VO_{2max}, but at least one less well-known confounder, the latter of which appears to have been overlooked in all the studies reported by Lolli *et al* The four well-known confounding variables associated with predicting VO_{2max} are M, age, sex and BF% or FM.[5] However, the less well-known confounding variable strongly associated with VO_{2max} was lung function, which was shown to be a stronger predictor of VO_{2max} than BF%.[11 12]

Hence, in the current study, the main objective was to assess whether either FFM or by combining M and BF% as separate predictors might be the more appropriate body size variables to predict VO_{2max}. Secondary objectives were to explore the effect that other confounding variables might be able to help optimise the prediction of VO_{2max} providing new insights into the mechanisms associated with VO_{2max} and also to assess whether these multiplicative allometric models are more successful than previously published linear additive models when predicting the VO_{2max} ($mL.kg^{-1}.min^{-1}$) of individuals that vary considerably in weight from a non-athletic population.

METHODS

Cohort

The present analysis includes 3930 cardiopulmonary exercise tests (CPX) (3262 treadmill, 668 cycle ergometer) that were performed between 1 January 1971 and 31 December 2021. Inclusion criteria for the present dataset were as follows: (1) men and women aged >18 years; (2) no known pre-existing diagnosis of cardiovascular disease (coronary artery disease, myocardial infarction, heart failure, peripheral arterial disease or stroke); (3) no known pre-existing diagnosis of chronic obstructive pulmonary disease; (4) not taking beta blocker medications; (5) maximal CPX performed on a treadmill or cycle ergometer with a peak respiratory exchange ratio >1.0; (6) measurement of BF% was assessed using either SF, Bod Pod (BP) or DEXA (DXA); (7) measurement of pulmonary function (forced vital capacity (FVC)). Physical activity status was binary (active or inactive) with physical inactivity defined as performing <30 min of moderate-intensity activity on 5 days per week or <20 min of vigorous-intensity activity on 3 days per week. Sex-specific descriptive characteristics of the cohort are presented in table 1.

Statistical methods

We adopted the well-known allometric or power function model ($Y=a\cdot X^b$) used by Nevill *et al* to predict VO_{2max} but using FFM rather than M as the predictor variable,[1] as follows:

$$VO_{2max}\ (L\cdot min^{-1}) = a\cdot FFM^{b}\cdot \varepsilon,\ (Eq.\ 1)$$

where '*b*' is the FFM scaling exponent, the parameter '*a*' is the scaling constant and 'ε' is a multiplicative error ratio, assumed to increase in proportion to body size (see figure 1a–d in Results), a characteristic in data known as heteroscedasticity, that can be controlled by taking logarithms (ln=$\log_e$), as described below. The model can be linearised with a log-transformation, and linear regression (or analysis of covariance (ANCOVA)) can then be used to estimate the unknown parameters 'Ln(*a*)', and '*b*', as follows:

$$Ln(VO_{2max}) = \ln(a) + b\cdot Ln(FFM) + Ln(\varepsilon).\ (Eq.2)$$

Any categorical differences, such as sex and age (groups), can easily be incorporated into the model by allowing the parameter 'Ln(*a*)' to vary as fixed factors within the ANCOVA.

To compare the quality of fit of FFM with M adjusted for BF% when predicting VO_{2max}, we adopted an alternative multiplicative allometric model that has been used in

Table 1 Sex-specific descriptive characteristics (N, means and SD) of the cohort

	Male			Female		
	N	Mean	SD	N	Mean	SD
Age (years)	2167	46.4	14.5	1763	45.2	15.7
Height (m)	2167	1.78	0.07	1763	1.64	0.07
Body mass (kg)	2167	84.9	15.68	1763	70.4	16.47
Waist (cm)	1904	93.3	12.18	1643	81.9	13.35
BF%	2167	22.8	8.78	1763	33.4	9.48
FVC (L)	2167	4.73	0.90	1763	3.45	0.72
VO_{2max} ($L.min^{-1}$)	2167	3.20	0.78	1763	1.98	0.53
VO_{2max} ($mL.kg^{-1}.min^{-1}$)	2167	38.5	10.57	1763	29.1	8.97

BF%, percentage body fat; FVC, forced vital capacity; VO_{2max}, maximum oxygen uptake.

the past (see Nevill *et al* to predict VO_{2max} ($L \cdot min^{-1}$),[13] by adjusting M for differences in BF%. This involves incorporating BF% into the multiplicative allometric model within an exponential function as follows:,

$$VO_{2max}\ (L \cdot min^{-1}) = a \cdot M^{b} \cdot \exp\ (c \cdot BF\%) \cdot \varepsilon, \text{ (Eq.3)}$$

As in Eq. 1, the model (Eq. 3) can be linearised with a log-transformation, and linear regression (or ANCOVA) can be used to estimate the unknown parameters '$\ln(a)$', 'b' and 'c',

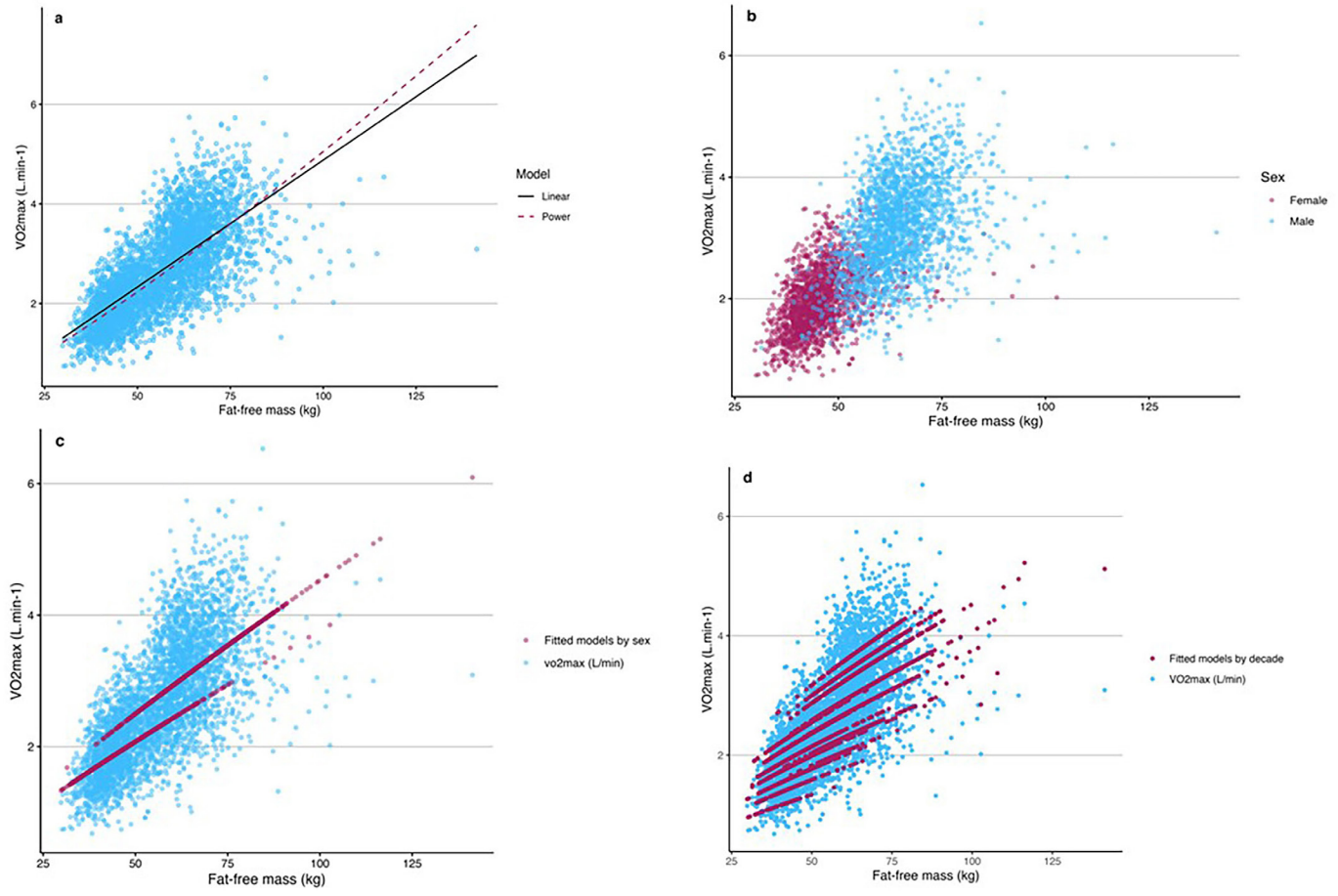

Figure 1 (a) The association between maximum oxygen uptake (VO_{2max}; $L.min^{-1}$) and fat-free mass (FFM; kg), r=0.704. (b) The association between VO_{2max} and FFM by sex. (c) The association between VO_{2max} and FFM with the fitted allometric models for male and females. (d) The association between VO_{2max} and FFM allowing separate fitted allometric model intercepts for different age groups (by decade).

$\text{Ln}(VO_{2max}) = \ln(a) + b \cdot \text{Ln}(M) + c \cdot BF\% + \text{Ln}(\varepsilon)$. (Eq.4)

In effect, log-transformed VO_{2max} becomes the dependent variable in the ANCOVA, with sex and age introduced as the fixed factors and log(M), and BF% as the covariates. In the extended models that also incorporated central adiposity (ln waist circumference (ln(WC))), waist circumference (WC) was not available for all participants. Sensitivity analyses using multiple imputation by chained equations (m=20, maxit=30) confirmed that complete-case estimates were robust to 9.7% missingness in WC.

Although the log-transformation stabilises variance and supports the allometric error structure, it does not correct measurement error in VO_{2max} or covariates. Therefore, the practical accuracy of the prediction equations depends on the reliability and standardisation of the underlying CPET and body-composition measurements; the prediction error reported here reflects the combined biological and measurement variability present in the FRIEND registry. As repeated tests under identical conditions were not available, we could not estimate intrasession reliability or external responsiveness directly within this dataset.[14 15] For the final prediction equations (table 2), the fitted residual SD on the log scale was $\sigma \approx 0.162$ (95% CI 0.158 to 0.167), corresponding to a typical multiplicative prediction error of approximately 17.6% on the original VO_{2max} scale ($100 \times [\exp(\sigma) - 1]$)

To assess potential multicollinearity among predictors, variance inflation factors (VIFs) were calculated. To assess the potential for multicollinearity among the predictor variables in the ANCOVA models, VIFs were calculated.

RESULTS

The strong, positive association between VO_{2max} (L.min^{-1}) and FFM (kg) is observed in figure 1a. Indeed, ignoring age and sex, the simple Pearson's correlation between VO_{2max} and FFM was r=0.704, compared with the correlation between VO_{2max} and M given by r=0.385. The fitted simple allometric model parameters (log-transformed, see Eq.2) are given in table 3a.

However, this simple allometric model (table 3a) overlooks/ignores the sex and age differences in VO_{2max} (see figure 1b, c and d). When we introduce the association between VO_{2max} and FFM allowing for these sex and age differences using Eq.2, we obtained the parameters in table 3b.

Table 2b reveals a systematic decline in the age group intercepts (by decade), a feature clearly seen in figure 1d. Note also that the FFM exponent is b=0.658 (95% CIs 0.614 to 0.701) with an explained variance of R^2=0.718 or 71.8%.

To compare the quality of fit predicting VO_{2max} reported in table 3b using FFM alone (Eq.2), with the quality of fit predicting VO_{2max} using M and BF% separately, we reanalysed the data using both Eq. 2 and Eq.4. The fitted parameters for FFM alone and M and BF% separately are given in table 3a and b, respectively. Note that the discrete age decline observed in table 3b was confirmed in table 3a and b but rather using the discrete age decade category reported in table 3b, we adopted a continuous quadratic age^2 term, as recommended by Nevill and Cooke[16] (see table 3a and b).

Table 3a gives the FFM exponent to be b=0.643 (95% CI 0.60 to 0.685) with an explained variance R^2=0.727 or 72.7% and an AIC=−1986.63. Table 3b identified the M exponent to be similar b=0.636 (95% CI 0.595 to 0.678) together with a negative BF% decline c=−0.012. The explained variance in table 4b was greater than that reported in table 4a, R^2=0.733 or 73.3% versus R^2=0.727 or 72.7%. This difference in R^2 may appear trivial; however, a more convincing model comparison (or model fit) was obtained using the AIC=−2077.4 reported in table 4b, which was considerably lower than that reported in table 3a, given as AIC=−1882.52 (difference in AIC=194.88) or in table 4a, given by AIC=−1986.63 (difference in AIC=90.77). Note that a difference in AIC between two competing models of less than 2 is insignificant; differences >2 ≤ 6, evidence for the lower model is somewhat positive; differences >6 ≤ 10, the evidence for the model with the lower AIC value is strong; differences>10, the evidence for the model with the lower AIC is very strong.[17 18] VIFs were examined for all ANCOVA specifications; for the final table 4 models, all VIFs were below 3 (with the highest being 2.95 (Sex), followed by 2.64 (BF%), 2.35 (ln[mass]) and 1.13 (Age^2), indicating no evidence of numerically problematic collinearity in those fitted models.

When we repeat the analysis reported in table 4a predicting Ln (VO_{2max}) using Ln(FFM) BUT also incorporating BF% (see fitted parameters in table 4c), the FFM exponent is similar to that reported in table 4a, b=0.646 (95% CI 0.604 to 0.688), but the BF% term remained significantly negative (B=−0.003; 95% CI −0.004 to −0.003). The explained variance (table 4c) is very similar R^2=0.733 or 73.3% to that reported in table 4b. This suggests that by incorporating FFM alone in the models given in table 3b, both models fail to remove all the negative/adverse effects of excessive BF%, and that a further/additional component of BF% in addition to FFM is required to be removed (−0.003) to optimally predict VO_{2max} (see the parameters in table 4c).

To help explain the likely cause of this 'additional' negative BF% effect reported in table 3b and c, and to answer our secondary aim outlined in the Introduction, that is, to explore the effect that other variables have to help optimise the prediction of VO_{2max}, we introduced some additional variables into our prediction models, some already reported by Nevill *et al*.[19] These included three additional categorical factors (type of body fat measure, type of exercise test and active vs inactive participants) plus the two continuous predictors of FVC and WC. These 'more inclusive' prediction models are given

Table 2 The fitted allometric model parameters associated with predicting Ln(VO_{2max}) (a) using Ln(M), BF%, LN(FVC), Ln(WC) and age^2 as covariates, by sex, CPX method, PA and BF% method, and (b) using Ln(FFM), BF%, LN(FVC), Ln(WC) and age^2 as covariates, by sex, CPX method, PA and BF% method

					95% CI	
Parameter	**B**	**SE**	**t**	**Sig.**	**Lower bound**	**Upper bound**
Table 2a						
Intercept (Male) Ln(a)	−0.927	0.127	−7.282	<0.001	−1.176	−0.677
(Sex=Female)ΔLn(a)	−0.266	0.026	−10.408	<0.001	−0.316	−0.216
(max_rer=CY)ΔLn(a)	−0.121	0.013	−9.148	<0.001	−0.146	−0.095
(Phys_Active)	0.142	0.007	20.593	<0.001	0.129	0.156
(bodyfat_method=BP)	−0.001	0.008	−0.141	0.888	−0.017	0.015
(bodyfat_method=DXA)	−0.018	0.008	−2.105	0.035	−0.034	−0.001
Lnmasskg	0.673	0.034	19.623	<0.001	0.606	0.74
bodyfat	−0.008	0.001	−13.196	<0.001	−0.009	−0.007
(Sex=Female) × bodyfat	0.002	0.001	2.609	0.009	0	0.003
LnFVC	0.283	0.02	14.049	<0.001	0.243	0.322
LnWC	−0.229	0.043	−5.299	<0.001	−0.314	−0.144
Age2	−8.19E- 05	2.89E−06	−28.374	<0.001	−8.75E−05	−7.62E−05
R squared=0.796 (adjusted R squared=0.796); AIC=−2190.25						
Table 2b						
Intercept (Male) Ln(a)	−0.981	0.127	−7.714	<0.001	−1.231	−0.732
(Sex=Female)ΔLn(a)	−0.306	0.025	−12.172	<0.001	−0.356	−0.257
(max_rer=CY)ΔLn(a)	−0.119	0.013	−9.026	<0.001	−0.145	−0.093
(PhysInactive=0)	0.143	0.007	20.756	<0.001	0.13	0.157
(bodyfat_method=BP)	−0.001	0.008	−0.08	0.936	−0.017	0.016
(bodyfat_method=DXA)	−0.017	0.008	−2.079	0.038	−0.034	−0.001
LnFFFkg	0.677	0.034	19.633	<0.001	0.609	0.745
bodyfat	0.001	0.001	1.712	0.087	0	0.002
(Sex=Female) × bodyfat	0.003	0.001	4.527	<0.001	0.002	0.005
LnFVC	0.279	0.02	13.827	<0.001	0.239	0.318
LnWC	−0.226	0.043	−5.249	<0.001	−0.311	−0.142
Age^2	−8.27E−05	2.88E−06	−28.699	<0.001	−8.83E−05	−7.70E−05
R squared=0.796 (adjusted R squared=0.796); AIC=−2190.6						

Inactive male participants whose BF% was assessed using skinfolds and whose exercise test (CPX) was measured on a treadmill were estimated as the baseline measure Ln(a) and all other categories were compared with it, indicated by ΔLn(a).
AIC, Akaike information criterion; BF%, percentage body fat; BP, Bod Pod; CPX, cardiopulmonary exercise tests; FFM, fat-free mass; FVC, forced vital capacity; M, body mass; PA, physical activity; VO_{2max}, maximum oxygen uptake; WC, waist circumference.

in table 4a and b, the former with M and the latter with FFM as the body size dimension variables.

It appears that the estimated M exponent and FFM exponents in table 4a and b are very similar, given as 0.673 and 0.677, respectively. This supports the exponent anticipated by Åstrand and Rodahl[2] who provided a plausible explanation as to why VO_{2max} (L.min^{-1}) of trained athletes should scale to body mass $M^{2/3}$. In the current dataset, where most participants were certainly not trained athletes, provided we can adjust M and FFM for the excess body fat, the adjusted FFM or M exponents should also scale to the same exponent (b=2/3).

Clearly, the models reported in table 4a and b are reported primarily for information purposes only, providing the readers with insight into the mechanisms associated with predicting VO_{2max}. However, these 'more inclusive' models are much too intricate for practical purposes. For example, details of lung function are unlikely to be available in most population studies. As such, much simpler prediction models are required to predict VO_{2max}, see more practical additive (linear) models by Myers *et al* or the equivalent allometric (curvilinear) models by Nevill *et al*, but to include FFM or BF%, given the focus of the current study.[20 21]

Table 3 The fitted allometric model parameters (Eq. 2) for (a) Ln(VO_{2max}) using Ln(FFM) as the covariate assuming a single common intercept and (b) for Ln(VO_{2max}) using Ln(FFM) as the covariate allowing separate intercepts by sex and age group (decade)

					95% CI	
Parameter	**B**	**SE**	**T**	**Sig.**	**Lower bound**	**Upper bound**
Table 3a						
Intercept Ln(a)	−3.804	0.07	−54.224	<0.001	−3.941	−3.666
Ln(FFM)(b)	1.178	0.017	67.348	<0.001	1.143	1.212
R squared=0.536 (adjusted R squared=0.536)						
Table 3b						
Intercept (Male) Ln(a)	−1.992	0.091	−21.782	<0.001	−2.172	−1.813
(Sex=Female)ΔLn(a)	−0.265	0.01	−27.009	<0.001	−0.284	−0.246
Ln(FFM)(b)	0.658	0.022	29.699	<0.001	0.614	0.701
(Agegroup=20.00)	0.601	0.021	28.793	<0.001	0.56	0.642
(Agegroup=30.00)	0.53	0.02	26.338	<0.001	0.491	0.57
(Agegroup=40.00)	0.463	0.02	23.428	<0.001	0.424	0.502
(Agegroup=50.00)	0.37	0.02	18.942	<0.001	0.332	0.409
(Agegroup=60.00)	0.273	0.02	13.714	<0.001	0.234	0.312
(Agegroup=70.00)	0.137	0.021	6.441	<0.001	0.095	0.179
(Agegroup=80.00)	0	.	.	.	.	.
R squared=0.718 (adjusted R squared=0.717), AIC=−1882.52						

Male participants aged 80 are used as the baseline measure Ln(a), and female participants are compared with it, indicated by ΔLn(a).
AIC, Akaike information criterion; FFM, fat-free mass; VO_{2max}, maximum oxygen uptake.

Based on the fitted parameters in table 4b, the log-transformed allometric model becomes Ln (VO_{2max}) = −1.173–0.235 × sex+0.636 · Ln(M) −0.0012 · BF% − 0.0000984 × age^2, where sex is entered as a [0,1] indicator variable (male=0 and female=1). The prediction equation to predict VO_{2max} (L.min^{-1}) becomes

VO_{2max} (L · min^{-1}) = exp(−1.173–0.235 × sex) × M$^{0.636}$ × exp(−0.012 × BF%−0.0000984 × age^2). (Eq.5)

As described by Nevill *et al* because we are fitting a log-linear regression model (Eq. 4), if we wished to predict the ratio standard VO_{2peak} (L·kg^{-1}·min^{-1}), all the fitted parameters in table 4b would remain the same with the exception of the log-transformed body mass term, which would simply become Ln(Mass) = (0.636–1) = −0.364.[22] Algebraically, this result can be derived simply by dividing both sides of Eq. 5 by Mass, or by taking Ln(Mass) from both sides of Eq. 4, resulting in the following allometric prediction equations,

VO_{2max} (L · kg^{-1} · min^{-1}) = exp(−1.173–0.235 × sex) × M$^{-0.364}$ × exp(−0.012 × BF%−0.0000984 × age^2),

or weight-related VO_{2max} (mL · kg^{-1} ·min^{-1}) as follows:

VO_{2max} (mL · kg^{-1} ·min^{-1}) = 1000 × exp(−1.173–0.235 × sex) × M$^{-0.364}$ × exp(−0.012 × BF%−0.0000984 × age^2). (Eq. 6)

Repeated 10-fold cross-validation (50 repeats) confirmed stable out-of-sample performance of this prediction model (mean cross-validated R^2 = 0.794 and RMSE=0.163 on the ln(VO_{2max}) scale). For example, a 20-year-old male who weighs 80 kg (mid-range weight) and has 25% body fat, the equation (Eq. 5 and Eq. 6) predicts his VO_{2max} to be 3.58 (L · min^{-1}) and 44.72 (mL.kg^{-1}.min^{-1}), respectively. Similarly, a 50-year-old female weighing 120 kg (overweight) with 35% body fat, the equations predict her VO_{2max} to be 2.64 (L · min^{-1}) and 22.0 (mL.kg^{-1}.min^{-1}), respectively. Download the following excel sheet to illustrate these examples interactively (https://doi.org/10.6084/m9.figshare.30306631). Note that the additive model reported by Myers *et al* predicted the same two individuals to have VO_{2max} = 49.75 and 13.17 (mL.kg^{-1}.min^{-1}), respectively.[20] Using the linear model proposed by Myers *et al*, results in an over-estimation of the mid-range or average weight male but an under-estimation of the (overweight) female.[20] The need for non-linear models such as the power-function curves can be clearly seen in figure 2a and b, respectively.

DISCUSSION

FFM is clearly a strong predictor of VO_{2max} (figure 1a). Ignoring the sex differences, the fitted FFM exponent was b=1.178 (95% CI 1.143 to 1.212). It was only by incorporating both sex and age (by decade) did we obtain a biologically plausible FFM exponent of b=0.658 (95% CI 0.614 to 0.701, see table 3b).

The FFM exponent predicted in table 3b, b=0.658 (ie, FFM$^{0.658}$), was similar to the M exponent 2/3 anticipated by Åstrand and Rodahl[2] based on sound biological/physiological grounds and our knowledge of muscle physiology. In their chapter on 'Body Dimensions and

Table 4 The fitted allometric parameters modelling the association between VO_{2max} and (a) FFM and age^2 by sex; (b) body mass, BF% and age^2 by sex; and (c) FFM, BF% and age^2 by sex

					95% CI	
Parameter	**B**	**SE**	**t**	**Sig.**	**Lower bound**	**Upper bound**
Table 4a						
Intercept Ln(a)	−1.303	0.091	−14.263	<0.001	−1.482	−1.124
Ln(FFMkg)	0.643	0.022	29.683	<0.001	0.6	0.685
Age^2	−1.02E04	2.11E−06	−48.644	<0.001	−1.07E−04	−9.83E−05
(Sex=Female) ΔLn(a)	−0.268	0.01	−27.688	<0.001	−0.287	−0.249
(Sex=Male)	0	.	.	.	.	.
R squared=0.727 (adjusted R squared=0.727); AIC=−1986.63						
Table 4b						
Intercept Ln(a)	−1.173	0.088	−13.359	<0.001	−1.345	−1
Ln(mass.kg)	0.636	0.021	30.19	<0.001	0.595	0.678
Bodyfat	−0.012	0.00045	−27.283	<0.001	−0.013	−0.011
Age^2	−9.84E−05	2.12E−06	−46.38	<0.001	−1.03E−04	−9.42E−05
(Sex=Female) ΔLn(a)	−0.235	0.01	−23.084	<0.001	−0.255	−0.215
(Sex=Male)	0	.	.	.	.	.
R squared=0.733 (adjusted R squared=0.733); AIC=−2077.4						
Table 4c						
Intercept	−1.252	0.09	−13.83	<0.001	−1.429	−1.074
Ln(FFMkg)	0.646	0.021	30.151	<0.001	0.604	0.688
Bodyfat	−0.003	0.00033	−9.582	<0.001	−0.004	−0.003
Age^2	−9.85E−05	2.12E−06	−46.458	<0.001	−1.03E−04	−9.44E−05
(Sex=Female)	−0.233	0.01	−22.736	<0.001	−0.253	−0.213
(Sex=Male)	0	.	.	.	.	.
R squared=0.733 (adjusted R squared=0.733); AIC=−2075.485						

Male participants are used as the baseline measure Ln(a) and female participants are compared with it, indicated by ΔLn(a).
AIC, Akaike information criterion; BF%, percentage body fat; FFM, fat-free mass; VO_{2max}, maximum oxygen uptake.

Muscular Exercise', Astrand and Rodahl[2] provided a plausible physiological/biological explanation as to why the VO_{2max} of 'trained' athletes should scale to body mass $M^{2/3}$ or L^2, where L is a linear dimension of body size. Body mass $M^{2/3}$ and $FFM^{2/3}$ are equivalent to a body surface or cross-sectional area, and their interpretation suggests that maximal oxygen uptake should be associated with, or proportional to, a cross-sectional area of M.

As Nevill and Wyon[10] anticipated, "The allometric models used to predict VO_{2max} using fat-free mass as the predictor variable will only be explicable if the exponent is sound/unbiased and all-important confounding variables have been explained and removed." When the alternative allometric model (Eqs. 3 and 4) was used to optimally predict VO_{2max}, allowing the allometric model to assess/adjust body mass for differences in BF% in

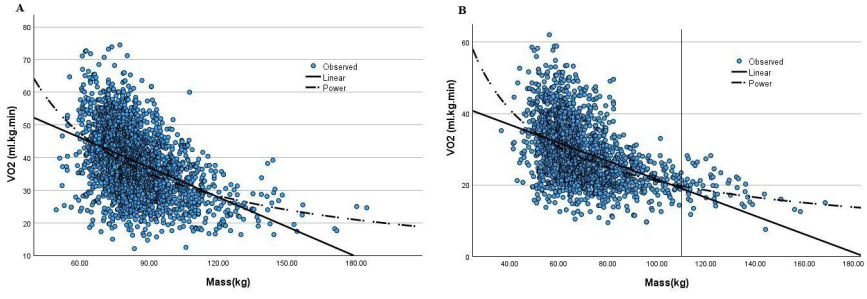

Figure 2 The association between maximum oxygen uptake (VO_{2max}) (mL.kg^{-1}.min^{-1}) and body mass (kg) using linear and power function models for (a) male and (b) female participants.

table 4a, the M exponent was b=0.636very similar to the FFM exponent in table 3a, but with a significant negative %BF term (c=−0.012; 95% CI −0.013 to −0.011). This alternative allometric model also explained more of the variance when predicting VO_{2max} (R^2=0.733). Note that the negative BF% term remained significant when both ln(FFM) and BF% were used to predict ln(VO_{2max}) in the reanalysis of the data incorporating both covariates in table 4b. Clearly, FFM alone fails to remove all of the detrimental effect of excessive BF% when attempting to optimally predict VO_{2max}. Note that this additional negative BF% term, reported in table 4a, has already been identified by Nevill *et al* (see Supplementary Table 4), who suggest that the likely cause of the additional negative BF% term might be due to the negative effect of central adiposity or excessive WC when predicting VO_{2max}.[19]

This was confirmed in the 'more inclusive' models reported in table 4a and b, when three additional categorical factors (type of BF% measure, type of exercise test and active vs inactive participants) and two continuous predictors of FVC and WC were incorporated into the prediction models, the first including M as the body size dimension (see table 2a) and the second incorporating FFM as the body size dimension (see table 2b). Both models found WC to be negative (p<0.001) but in the case of table 2b, the introduction of WC in addition to FFM meant the BF% coefficient became small and positive and was not statistically distinguishable from zero (p=0.087) *for males* BUT for *females* (see the sex-by-body fat interaction) the BF% term was significantly positive (p<0.001, differing from the conclusion reported by Goran *et al*.[3]

The main practical implication of this study is the importance of including M and body fat (%) separately when predicting VO_{2max}. Body fat has a negative contribution to aerobic capacity as the mass adds no metabolic benefit to oxygen consumption. With the rise in obesity worldwide, especially in Western countries, the inclusion of body fat % provides a more robust prediction model which is particularly important in large population studies where direct measurement is not realistic.

The main limitations of the study are within how the data were initially measured for the FRIEND registry. CPX included in the FRIEND registry used both treadmills and cycle ergometers, the former has a larger physiological response patterns than cycling[23] mainly due to the treadmill exercise being weight-bearing. This was partially overcome by applying an equation developed by de Souza e Silva *et al* to calculate norm standards for VO_{2max}, for the treadmill or cycle ergometer.[24] Sensitivity analyses allowing protocol-specific slopes did not alter the substantive conclusions (see online supplemental table S1-S2), which report the adjusted treadmill–cycle contrast and protocol-specific ln(mass) and BF% slopes). BF% was assessed using either SF, BP or DXA, as mentioned in the Introduction and Methods, the use of different methods[8] will add noise to the estimation of FFM and BF%. Sensitivity analyses indicated that the main conclusion, that modelling M and BF% as separate terms provides better performance than relying on FFM alone, was consistent across methods, although method-related differences in BF% levels and fitted slopes were detectable (eg, the estimated ln(mass) exponent ranged from 0.57 for SF to 0.85 for DXA; online supplemental table S3a, b).

CONCLUSIONS

Subtracting FM from M to obtain FFM appears to be less than optimal when predicting VO_{2max}. Using M and BF% as separate predictors provides a superior fit, allowing a positive contribution of M and a negative contribution of BF% to improve the prediction of VO_{2max}. Our analysis identified a larger BF% weighting that improved prediction of VO_{2max}; this weighting was markedly reduced when WC was included, consistent with Nevill *et al* in suggesting it may reflect the influence of central adiposity not captured by BF% when FFM is used as the body size term.[19] Furthermore, the M and FFM exponents of the more inclusive/final prediction models (table 2) were found to be approximately $M^{2/3}$ or L^2, suggesting that VO_{2max} should be normalised using preferably VO_{2max} ($mL{\cdot}M^{-2/3}{\cdot}min^{-1}$), as recommended by Nevill *et al*, or VO_{2max} ($mL{\cdot}kg\ FFM^{-2/3}{\cdot}min^{-1}$) rather than VO_{2max} ($mL{\cdot}kg\ FFM^{-1}{\cdot}min^{-1}$) as advised by Goran *et al* and Imboden *et al*.[1 3 4]

Author affiliations
[1]Faculty of Education, Health and Wellbeing, University of Wolverhampton, Wolverhampton, UK
[2]Health Sciences, Taylor University, Upland, Indiana, USA
[3]Sport and Physical Activity Research Centre, University of Wolverhampton Faculty of Education Health and Wellbeing, Walsall, UK
[4]National Institute of Dance Medicine and Science, Birmingham, UK
[5]Cardiology, VA Palo Alto Health Care System, Palo Alto, California, USA
[6]Department of Physical Therapy, University of Illinois Chicago, Chicago, Illinois, USA
[7]Newman University, Birmingham, UK
[8]Ball State University, Muncie, Indiana, USA

Acknowledgements The authors would like to thank Ball State University for access to the FRIEND database.

Contributors AMN is the guarantor of the work and accepts full responsibility for the integrity of the data and the accuracy of the analysis. AMN is also the lead author; substantial contribution to the intellectual content of the manuscript including statistical analysis and interpretation of the data. MW: substantial contribution to drafting and revising intellectual content of the manuscript including interpretation of the data, formatting the manuscript, final approval of submitted version. JM: substantial contribution to revising intellectual content of the manuscript including interpretation of the data, final approval of submitted version. MPH: substantial contribution to revising intellectual content of the manuscript including interpretation of the data, final approval of submitted version. RA: substantial contribution to revising intellectual content of the manuscript including interpretation of the data, final approval of submitted version. TDM: substantial contribution to revising intellectual content of the manuscript including analysis and preparation of figures and tables, final approval of submitted version. LAK: substantial contribution to revising intellectual content of the manuscript including interpretation of the data, access to dataset, final approval of submitted version. N/A.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, conduct, reporting or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants. The current data, recently published by Nevill *et al*,[13] were originally recorded in accordance with research ethics procedures at each participating location, including informed consent, independent of the FRIEND registry. Each contributing laboratory then obtained approval from their local research ethics board prior to submitting deidentified, coded data to the data coordinating centre at Ball State University, which has institutional review board approval for maintaining the database. All data were collected in accordance with the Helsinki Declaration. Data from each contributing laboratory were reviewed to ensure values were within expected normal ranges before being added to FRIEND. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data available at the FRIENDS registry.

ORCID iDs
Alan Nevill https://orcid.org/0000-0003-0506-3652
Matthew Wyon https://orcid.org/0000-0003-0942-2333
Jonathan Myers https://orcid.org/0000-0003-2592-136X

REFERENCES

1 Nevill AM, Ramsbottom R, Williams C. Scaling physiological measurements for individuals of different body size. *Europ J Appl Physiol* 1992;65:110–7.
2 Astrand P, Rodahl K. *Textbook of work physiology - physiological bases of exercise*. 3rd edn. New York: McGraw-Hill International Editions, 1986.
3 Goran M, Fields D, Hunter G, *et al*. Total body fat does not influence maximal aerobic capacity. *Int J Obes* 2000;24:841–8.
4 Imboden MT, Kaminsky LA, Peterman JE, *et al*. Cardiorespiratory Fitness Normalized to Fat-Free Mass and Mortality Risk. *Med Sci Sports Exerc* 2020;52:1532–7.
5 Lolli L, Batterham AM, Weston KL, *et al*. Size Exponents for Scaling Maximal Oxygen Uptake in Over 6500 Humans: A Systematic Review and Meta-Analysis. *Sports Med* 2017;47:1405–19.
6 Wang Z, Ying Z, Bosy-Westphal A, *et al*. Specific metabolic rates of major organs and tissues across adulthood: evaluation by mechanistic model of resting energy expenditure. *Am J Clin Nutr* 2010;92:1369–77.
7 Yamagishi T, Saito A, Kawakami Y. Determinants of whole-body maximal aerobic performance in young male and female athletes: The roles of lower extremity muscle size, strength and power. *PLoS One* 2022;17:e0262507.
8 Lukaski HC. Methods for the assessment of human body composition: traditional and new. *Am J Clin Nutr* 1987;46:537–56.
9 Nana A, Slater GJ, Stewart AD, *et al*. Methodology review: using dual-energy X-ray absorptiometry (DXA) for the assessment of body composition in athletes and active people. *Int J Sport Nutr Exerc Metab* 2015;25:198–215.
10 Nevill AM, Wyon M. Comment on: Size Exponents for Scaling Maximal Oxygen Uptake in over 6500 Humans-A Systematic Review and Meta-analysis. *Sports Med* 2024;54:2965–7.
11 Nevill AM, Holder RL. Identifying population differences in lung function: results from the Allied Dunbar national fitness survey. *Ann Hum Biol* 1999;26:267–85.
12 Darveau C-A, Suarez RK, Andrews RD, *et al*. Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature New Biol* 2002;417:166–70.
13 Nevill AM, Holder RL, Stewart AD. Modeling elite male athletes' peripheral bone mass, assessed using regional dual x-ray absorptiometry. *Bone* 2003;32:62–8.
14 Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998;26:217–38.
15 ATS/ACCP Statement on Cardiopulmonary Exercise Testing. *Am J Respir Crit Care Med* 2003;167:211–77.
16 Nevill AM, Cooke CB. The Dangers of Estimating V·O2max Using Linear, Nonexercise Prediction Models. *Med Sci Sports Exerc* 2017;49:1036–42.
17 Anderson D, Burnham K. *Model selection and multi-model inference*. 63. . Second NY: Springer-Verlag, 2004:10.
18 Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. *Sociol Methods Res* 2004;33:261–304.
19 Nevill AM, Wyon M, Myers J, *et al*. Predicting VO_{2max} Using Lung Function and Three-Dimensional (3D) Allometry Provides New Insights into the Allometric Cascade ($M^{0.75}$). *Sports Med* 2025;55:1757–67.
20 Myers J, Kaminsky LA, Lima R, *et al*. A Reference Equation for Normal Standards for VO_2 Max: Analysis from the Fitness Registry and the Importance of Exercise National Database (FRIEND Registry). *Prog Cardiovasc Dis* 2017;60:21–9.
21 Nevill AM, Myers J, Kaminsky LA, *et al*. 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. *Prog Cardiovasc Dis* 2019;62:515–21.
22 Nevill AM, de Castro Silveira JF, Gaya AR, *et al*. The dangers of predicting ratio standards using multiple regression: A case study using the ratio standard VO2peak (ml·kg−1·min−1). *J Sci Med Sport* 2026;29:85–90.
23 Abrantes C, Sampaio J, Reis V, *et al*. Physiological responses to treadmill and cycle exercise. *Int J Sports Med* 2012;33:26–30.
24 de Souza e Silva CG, Kaminsky LA, Arena R, *et al*. A reference equation for maximal aerobic power for treadmill and cycle ergometer exercise testing: Analysis from the FRIEND registry. *Eur J Prev Cardiolog* 2018;25:742–50.