



Scaling the peak and steady-state aerobic power of running and walking humans

Heather M. Bowes^{1,3} · Catriona A. Burdon¹ · Gregory E. Peoples¹ · Sean R. Notley² · Nigel A. S. Taylor¹

Received: 24 February 2021 / Accepted: 23 June 2021 / Published online: 2 July 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021, corrected publication 2021

Abstract

Purpose The first aim of this experiment was to evaluate the appropriateness of linear and non-linear (allometric) models to scale peak aerobic power (oxygen consumption) against body mass. The possibilities that oxygen consumption would scale allometrically across the complete metabolic range, and that the scaling exponents would differ significantly between basal and maximal-exercise states, were then evaluated. It was further hypothesised that the scaling exponent would increase in a stepwise manner with elevations in exercise intensity. Finally, the utility of applying the scaling exponent derived for peak aerobic power to another population sample was evaluated.

Methods Basal, steady-state walking and peak (treadmill) oxygen-consumption data were measured using 60 relatively homogeneous men (18–40 year; 56.0–117.1 kg), recruited across five mass classes. Linear and allometric regressions were applied, with the utility of each scaling method evaluated.

Results Oxygen consumption scaled allometrically with body mass across the complete metabolic range, and was always superior to both ratiometric analysis and linear regression. The scaling exponent increased significantly from rest ($\text{mass}^{0.57}$) to maximal exercise ($\text{mass}^{0.75}$; $P < 0.05$), but not between steady-state walking ($\text{mass}^{0.87}$) and maximal exercise ($P > 0.05$). When used with an historical database, the maximal-exercise exponent successfully removed the mass bias.

Conclusion It has been demonstrated that the oxygen consumption of healthy humans scales allometrically with body mass across the entire metabolic range. Moreover, only two scaling exponents (rest and exercise) were required to produce mass-independent outcomes from those data. Accordingly, ratiometric and linear regression analyses are not recommended as scaling methods.

Keywords Allometry · Body mass · Peak aerobic power · Peak oxygen consumption · Scaling · Size

Abbreviations

AIC	Akaike information criterion
r^2	Coefficient of determination
RMSE	Root-mean-square error
SD	Standard deviation

Introduction

This is the second communication from a series of experiments, wherein the relationship between whole-body metabolism and body mass was evaluated in healthy humans from basal through to maximal-exercising states. In the first experiment, the association with metabolic rate was explored

Communicated by Jean -Rene Lacour.

✉ Nigel A. S. Taylor
nigelastaylor@gmail.com

Heather M. Bowes
heatbowe@kth.se

Catriona A. Burdon
c.burdon@bigpond.com

Gregory E. Peoples
peoples@uow.edu.au

Sean R. Notley
snotley@uottawa.ca

¹ Centre for Medical and Exercise Physiology, School of Medicine, University of Wollongong, Wollongong, NSW 2522, Australia

² Human and Environmental Physiology Research Unit, School of Human Kinetics, University of Ottawa, Ottawa, Canada

³ Present Address: Department of Environmental Physiology, School of Technology and Health, Royal Institute of Technology, Stockholm, Sweden

during basal and resting states (Bowes et al. 2021). Just as others had found for animals (Sarrus and Rameaux 1838; Kleiber 1932; White and Seymour 2003), an allometric (power) model provided a biologically valid, powerful and superior description of that relationship, relative to linear scaling. Moreover, a unique intra-specific (human), body-mass exponent for metabolic rate was revealed (metabolic rate = $\text{mass}^{0.50-0.55}$). During exercise, there exists inter-specific (species) evidence that, whilst that allometric relationship is retained, the scaling exponent is elevated (Taylor et al. 1981; Bishop 1999; Weibel et al. 2004), especially during maximal-intensity exercise. Human evidence for such a change also exists, although it remains inconclusive, and possibly unconvincing. Therefore, the primary purpose of this contribution was to explore the veracity of different scaling models during maximal-intensity exercise, so that, when combined with basal observations from the same individuals (Bowes et al. 2021), those analyses would span the complete metabolic range (aerobic scope). Collectively, those descriptive observations would facilitate the derivation of mass-specific metabolic rate and oxygen consumption from existing absolute data.

During maximal exercise, whole-body oxygen consumption increases to ~ 20 times that observed at rest (Taylor et al. 1981 [animals]; Clark et al. 1983 [humans]), due to increases in the type and mass of the metabolically active tissues, and an intensity-dependent metabolic transition (Darveau et al. 2002; Weibel 2002). As a consequence, the skeletal-muscle contribution to metabolism increases from $\sim 20\%$ at rest (Elia 1992; Gallagher et al. 1998; Müller et al. 2011) to $\sim 90\%$ when exercising maximally (Mitchell and Blomqvist 1971; Hochachka 1994). Collectively, those factors could modify the relationship between oxygen consumption and body mass, relative to that observed at rest. Therefore, others have postulated that the scaling exponent would increase with elevations in metabolic intensity, perhaps in the form of an allometric cascade (Darveau et al. 2002; Hochachka et al. 2003; Suarez et al. 2004), reaching a theoretical maximum of 0.92 ($\text{mass}^{0.92}$; Darveau et al. 2002).

Since the comparative literature reveals that such an exponent increase seems to occur when moving from rest [e.g. $\text{mass}^{0.67}$ (White and Seymour 2003)] through to maximal exercise [$\text{mass}^{0.87}$ (Weibel et al. 2004)], that possibility needed to be explored within humans. Moreover, since it is plausible that the species specificity of the allometric relationship, evident at rest (Heusner 1982; White and Seymour 2003; Sieg et al. 2009), may continue across the entire metabolic range, human-specific data were required to provide a more complete evaluation of those possibilities. Ideally, data should be gathered using the same individuals used to scale basal metabolic rate (Bowes et al. 2021).

However, for human data, ratiometric scaling ($\text{mass}^{1.0}$) dominates the exercise literature, even though allometric

scaling is commonly used for both resting (Kleiber 1932; Sieg et al. 2009; White and Kearney 2014) and exercising states (Taylor et al. 1981; Bishop 1999; Weibel et al. 2004) within the comparative literature. The ratiometric scaling of peak oxygen consumption (peak aerobic power), typically converts the mass-dependent (positive) bias evident within those absolute data into a systematic negative bias, when expressed as mass-specific equivalents (Vaage and Hermansen 1977). This occurs because neither the dependent nor the independent variables change isometrically (a pre-requisite of ratiometric scaling; Packard and Boardman 1988), but disproportionately. The effect is that the validity of those mass-independent data, as well as the capacity for predicting data beyond those experimental mass ranges, remains questionable.

Recognising that limitation, some have adopted non-linear scaling to analyse human peak aerobic power (e.g. Secher et al. 1983; Bergh et al. 1991; Nevill et al. 1992; Batterham and Jackson 2003; Markovic et al. 2007; Lolli et al. 2017). Unfortunately, efforts to minimise sampling bias, by standardising inter-individual variability and obtaining evenly distributed and adequately sized samples across the mass range investigated (White and Seymour 2005), seem not to have been made. Indeed, data were often obtained from samples of convenience, or from existing databases. Not surprisingly, the resulting exponents are highly variable, even when restricted to healthy, adult males [e.g. $\text{mass}^{0.63}$ (Nevill et al. 1992); $\text{mass}^{0.94}$ (Batterham and Jackson 2003)]. Those inconsistencies complicate the adoption and application of a uniform scaling approach for human data, and will neither convince those who prefer linear scaling to consider another possibility, nor will they help clarify the relationship between body mass and oxygen consumption during exercise.

Therefore, the first aim was to evaluate how human, peak aerobic power would scale against body mass within a homogeneous sample of healthy, adult males. It was hypothesised that data collected across the metabolic range would scale allometrically with mass. Secondly, it was hypothesised that the scaling exponent for maximal (treadmill) exercise would be significantly larger than observed at rest. To test the possibility that all non-resting states would have a significantly larger scaling exponent than resting states, data obtained during steady-state walking (4.8 km h^{-1}) were also scaled. It was further hypothesised that the size of the scaling exponent would increase in a stepwise manner from rest, to steady-state walking and then to maximal running. Those hypotheses were evaluated using a single population sample ($N=60$), which varied primarily in body mass, with all individuals providing data for each metabolic intensity. Those data were scaled using both linear and allometric methods; the latter reflect the state of the art in comparative physiology. Finally, to test the hypothesis that the scaling

exponent for peak aerobic power would be applicable to broader population samples, historical data were extracted from an in-house database, and analysed.

Methods

Participants

For this investigation, data from two different population samples of males were analysed. In the first instance, a homogeneous sample of healthy University students was used ($N=60$; Table 1), all of whom had already been studied under basal conditions (Bowes et al. 2021). Those participants provided written, informed consent before undertaking both steady-state ambulatory (treadmill walking) and maximal running exercise (treadmill). The resulting data enabled a comparison of scaling models that could be applied to both ends of the metabolic range, with walking representing an intermediate state, at the lower end of the exercising range. All experimental procedures were approved by a Human Research Ethics Committee (University of Wollongong, Australia: HE14/469) following national regulations (National Health and Medical Research Council), and in compliance with the Declaration of Helsinki. A second sample ($N=54$) came from an in-house, historical database of maximal-exercise data (running), collected by the authors, and for which the same level of technical precision and experimental control was known to exist. Those individuals were used to represent a random sample of physically active University men, all of whom had provided written, informed consent to similarly, and previously, approved procedures. Those data were used to evaluate the robustness of the maximal-exercise scaling model.

Experimental participants

The experimental sample came from a pool of 71 healthy men, 68 of whom had been studied under basal conditions (Bowes et al. 2021), and who now participated in steady-state walking. However, only 61 subjects performed maximal exercise, which left 60 individuals who participated in all three trials. They were the experimental participants. Of the 11 remaining participants, three suffered unrelated injuries, five withdrew and three were deemed to be outliers with regard to their relative adiposity or absolute peak aerobic power (Cook's distance), so their data were excluded. All remaining subjects were either competitive athletes or previous research participants, and engaged in endurance exercise at least three times per week. Strict recruitment criteria were used to minimise the effects of inter-individual variations on the scaling models, as described elsewhere (Bowes et al. 2021). In the first instance, sexual dimorphism was prevented. Age-dependent mass differences in the metabolically active tissues (Holliday et al. 1967; Holliday 1971; Müller et al. 2011) were minimised by recruiting adults aged 18–40 years (Table 1). To minimise data clustering, and its adverse impact upon scaling (White and Seymour 2005), 10–15 participants were initially recruited across each of five body-mass classes (Table 1; extra small: 55–65 kg; small: 66–76 kg; medium: 77–87 kg; large: 88–98 kg; extra large: > 99 kg), although the loss of three extra-large individuals represents a sampling limitation. Finally, variations in body composition were minimised by standardising subcutaneous adiposity, such that only those with average-to-low adiposity were included (i.e., height-adjusted sum of six skinfold thicknesses ≤ 88 mm; Ross and Wilson 1974).

Table 1 Physical characteristics of the experimental subjects

Participant size classifications	Sample sizes	Age (years)	Height (cm)	Body mass (kg)	Height-adjusted body mass (mm)	Sum of skinfolds (mm)	Height-adjusted adiposity (mm)
All participants	60	22.8 (3.3)	181.1 (8.6)	79.9 (14.2)	65.7 (6.9)	58.0 (23.9)	53.2 (20.2)
Ranges		18–34	163.0–201.5	56.0–117.1	50.2–84.1	28.0–129.0	28.5–112.2
Extra small	12	21.8 (3.2)	172.7 (5.2)	61.3 (2.2)	58.7 (5.6)	42.1 (10.5)	41.6 (10.8)
Small	13	21.6 (2.3)	173.9 (4.5)	71.5 (3.1)	66.9 (4.8)	56.2 (20.0)	52.1 (17.9)
Medium	19	22.7 (3.8)	184.1 (3.5)	81.4 (2.4)	64.4 (5.0)	54.3 (16.9)	50.4 (16.2)
Large	11	24.1 (2.5)	186.3 (5.6)	93.9 (2.9)	71.9 (7.2)	67.3 (22.1)	61.8 (21.0)
Extra large	5	25.2 (3.9)	196.8 (2.7)	109.3 (4.5)	70.8 (3.3)	94.0 (41.8)	76.2 (33.7)

Data are means with standard deviations in parentheses, and overall data ranges

The height-adjusted measures are anthropological data adjusted to the stature of 170.18 cm (Ross and Wilson 1974). Skinfold thicknesses are the sum of values measured at six sites (see methods). None of the within-class differences was significantly different between the current experimental sample and that used for scaling basal metabolic rate (Bowes et al. 2021; $P > 0.05$)

In-house, historical participants

Data were also extracted from an in-house database of published observations for peak aerobic power (Booth et al. 2001; Taylor et al. 2012; Peoples et al. 2016, 2017; Craddock et al. 2021). In those investigations, one-off (incremental) treadmill runs to volitional exhaustion, were performed on physically active subjects. Data were extracted for participants aged 18–40 years ($N=54$), and covering a body-mass range of 54.6–101.7 kg {age: 26 years [standard deviation (SD) 6.3]; body mass: 76.3 kg (SD 8.7); height 179.5 cm (SD 6.3); absolute peak aerobic power: 4.44 L min⁻¹ (SD 0.54)}. However, clustering occurred within the small-to-medium, body-mass classes (extra small: $N=7$; small: $N=21$; medium: $N=22$; large: $N=3$; extra large: $N=1$), as defined for the experimental sample. Those data were initially used to evaluate the robustness of the maximal-exercise scaling exponent derived from the experimental subjects, when applied to data collected for the same exercise mode, but without the physical and physiological attributes of those individuals being standardised. Secondly, to evaluate the possible impact of applying scaling procedures to a heterogeneous sample of convenience, over which no attempt was made to obtain either a wide or even distribution of body masses, the peak aerobic-power data of those 54 individuals were independently scaled.

Procedures

Experimental overview

Incremental (graded) exercise tests were administered to the experimental participants on three, non-consecutive days, starting at the same time. That approach minimised the impact of familiarisation and practice effects on performance (Shephard et al. 1968), and increased comfort and confidence during subsequent treadmill exercise. Data from the last maximal-exercise test were used to represent peak aerobic power. During each laboratory visit (~23 °C and ~50% relative humidity), subjects commenced with blocks of steady-state walking (which also served as a warm-up), followed by seated rest and then an incremental forcing function (maximal treadmill ramp) to volitional exhaustion. Subjects wore running shoes, shorts, t-shirt and a torso safety harness. Each ramp test commenced at a unique, pre-determined running speed for each person (8–12 km h⁻¹; horizontal), with only the gradient increasing thereafter (1% steps each minute: ~12 min). The criterion for test termination was volitional failure, with verbal encouragement provided. Steady-state exercise data were collected during horizontal walking (15 min at 4.8 km h⁻¹), before the third incremental exercise test.

Prior to commencing this experiment, a pilot trial was undertaken in which 10 individuals (females and males) with similar endurance-training behaviours, performed three maximal-exercise trials, as per the main experiment. Oxygen consumption was measured continuously during each trial. Peak aerobic power did not change significantly across those trials [trial 1: 3.69 L min⁻¹ (SD 0.77); trial 2: 3.71 L min⁻¹ (SD 0.75); trial 3: 3.59 L min⁻¹ (SD 0.86); $P>0.05$]. It was therefore interpreted that, for similarly trained subjects, treadmill familiarity and learning effects would not introduce measurement artefacts.

Experimental standardisation

Subjects were asked to present in a well-rested (~8 h sleep) and normally hydrated state, having abstained from strenuous physical activity over the previous 24 h. They were also requested to consume a high-carbohydrate, low-fat meal > 2 h before testing, and to maintain normal fluid intakes.

Measurements

Anthropometric and skinfold-thickness measurements were taken for the experimental subjects, as described elsewhere (Bowes et al. 2021). From those indices, height-adjusted body-mass and skinfold-thickness data were derived, following the method of Ross and Wilson (1974): mass (kg) × [170.18/height (cm)]³; and skinfold sum (cm) × [170.18/height (cm)]. Those values were used to minimise inter-individual variability in relative muscularity and adiposity during recruitment.

Oxygen consumption was measured during rest and exercise (open-circuit respirometry), with expired gas fractions and expiratory flows sampled from a two-way, low-resistance, non-rebreathing valve (7400 series, Hans Rudolph Inc., Kansas, USA). Those data were analysed continuously (TrueOne 2400, ParvoMedics Inc., Utah, USA) and used to derive oxygen consumption, carbon dioxide production and minute ventilation (15-s averages). Two-point gas [room air plus *alpha* standard gases (16.00% oxygen, 4.00% carbon dioxide, 80.00% nitrogen)] and a range of flow calibrations (50 to > 300 L min⁻¹) preceded every trial. Steady-state oxygen consumption was averaged over the last 5 min, whilst the highest oxygen consumption observed during incremental running was taken as each individual's peak value.

Design and analysis

To minimise subsequent scaling errors, assumptions underlying both the sampling and scaling procedures were evaluated. Those processes began with an assessment of the sample homogeneity. One-way Analysis of Variance was used

to compare both raw and height-adjusted morphological data across the body-mass groups, with differences evaluated for statistical significance using Tukey's HSD post hoc procedure. If between-group differences existed, Analysis of Covariance was used to evaluate whether or not that variable, when unadjusted, had a significant interaction with the scaling relationship; such interactions did not occur. For all statistical analyses, α was set at the 0.05 level. Data are reported as means with standard deviations, and 95% confidence intervals for the scaling coefficients.

The relationship between peak aerobic power and body mass was modelled both linearly and non-linearly (hypothesis one). Indeed, for those scaling models to be valid, each had to satisfy both biological and statistical assumptions, because best fit does not necessarily equate with the correct solution (Sholl 1948; Motulsky and Ransnas 1987; Nevill et al. 1992). The biological criterion was an origin intercept (zero-zero), which was appraised using untransformed data, and evaluated using t tests. Since a non-zero oxygen consumption when the body mass is zero is biologically impossible (Krogh 1916; Kleiber 1950), that criterion was heavily weighted.

To provide an evaluation of the model shape, non-linear data were initially log-transformed [log-linear (\log_{10} base); as described elsewhere; Bowes et al. 2021] so that the assumptions underlying linear regression could be appraised (normality, linearity, homoscedasticity and an absence of first-order correlations). Normality was assessed using skewness and kurtosis z scores (± 3.29 when $50 < N < 300$; Kim 2013), as well as the Shapiro–Wilk test. Linearity and homoscedasticity were evaluated using scatter and residual plots (respectively). The Durbin–Watson test was used to assess first-order correlations between variables.

To enhance the accuracy of each scaling model, the following statistical procedures were followed. Firstly, Cook's distance test was used to identify potential outliers within the absolute oxygen-consumption data that might disproportionately influence scaling, with the cut-off set to $4/(N - k - 2)$; where N was the sample size and k the number of coefficients. Data from three individuals were omitted (Analysis of Covariance or an abnormally low absolute value). The scaling models were then bootstrapped (1000 repetitions using resampling) to generate coefficient means, and Bias-Corrected and Accelerated analyses were used to yield 95% confidence intervals that were independent of skewness produced during bootstrapping.

The relationship between peak aerobic power and body mass was modelled using both linear (first-order polynomial: Eq. 1A) and non-linear scaling (allometric: Eq. 2A), with ratiometric analysis also performed. For steady-state walking, both scaling models were again used, as described elsewhere (Bowes 2018), with only the latter being reported here. Non-linear data were also

log-transformed for use with analyses that assumed linearity (log-linear; Eq. 3A). For all equations, body masses are in kg and oxygen-consumption data are in absolute units (L min^{-1}). The following equation numbers appear throughout the manuscript, but with the alphabetical qualifiers varying according to their application.

Linear regression: Oxygen consumption

$$= a + b \times \text{mass} \quad [\text{L min}^{-1}], \quad (1A)$$

where a = ordinate intercept [L min^{-1}], b = linear coefficient [slope; $\text{L kg}^{-1} \text{min}^{-1}$].

Allometric regression: Oxygen consumption

$$= b \times \text{mass}^k \quad [\text{L min}^{-1}], \quad (2A)$$

where b = linear coefficient [$\text{L kg}^{-1} \text{min}^{-1}$], k = coefficient of allometry (scaling exponent) for body mass [kg].

Log-linear : Log_{10} (oxygen consumption)

$$= \text{Log}_{10}(b) + k \times \text{Log}_{10}(\text{mass}) \quad [\text{L min}^{-1}], \quad (3A)$$

where b = linear coefficient [$\text{L.kg}^{-1}.\text{min}^{-1}$], k = scaling exponent for body mass [kg].

A limitation of the present study, which was primarily descriptive, rather than mechanistic in nature, is that, whilst both the linear coefficient and the coefficient of allometry (Eq. 2A) will vary, only variations in the latter have been explored. Therefore, it would be remiss to ignore the impact of variations in the linear coefficient, which can independently, and in combination with changes in the scaling exponent, significantly affect the prediction model. For example, when using Eq. 2A to predict absolute peak aerobic power from body mass alone, a 10% variation in the linear coefficient would result in an equivalent relative change across the complete body-mass range. One suspects that few would have that objective. Instead, it is much more likely that others would choose to derive mass-specific values for peak aerobic power, having first measured absolute oxygen consumption. That conversion only requires the coefficient of allometry (Eq. 4), which is the central focus of this manuscript.

$$Y_{\text{adjusted}} = Y_{\text{measured}} / \text{mass}^k \quad [\text{mL kg}^{-k} \text{min}^{-1}], \quad (4)$$

where Y_{adjusted} = mass-specific oxygen consumption, Y_{measured} = absolute oxygen consumption [L min^{-1}], k = scaling exponent for body mass [kg].

Differences in the goodness-of-fit among those models were compared using the root-mean-square error (RMSE), the Akaike Information Criterion (AIC), coefficients of determination (r^2) and F tests (Motulsky and Ransnas 1987). Pearson's correlation coefficient was used to evaluate the remaining mass bias (residuals) within the

scaled oxygen-consumption data, with correlations that approached zero indicating successful removal of that bias. Between-model comparisons were made using Analysis of Covariance, with a significant interaction effect demonstrating that those models had unique slopes (linear: linear coefficients; non-linear: exponent). Differences among the non-linear exponents for oxygen consumption during basal conditions, steady-state walking and maximal exercise were evaluated using Analysis of Covariance. That test required a sample of at least 52 participants to detect between-group slope (exponent) differences of a large effect size (Cohen's $f=0.40$), with a statistical power of at least 80% (Faul et al. 2007). Thus, both the experimental and historical samples were adequately powered (a priori) for these analyses. The exponent for peak aerobic power was also compared with both the peak animal exponent ($\text{mass}^{0.87}$; Weibel et al. 2004) and the proposed theoretical maximal exponent ($\text{mass}^{0.92}$; Darveau et al. 2002) using t tests.

To evaluate the robustness of the final experimental scaling model, when applied to data obtained from an independent population sample, an in-house sample of convenience was used. Those data were also allometrically scaled (Eq. 2A) to evaluate the impact of performing that analysis without controlling for variations in the physical and physiological characteristics, or the range and distributions of the body masses, of those subjects.

Results

Homogeneity of the experimental sample

Whilst the experimental participants were distributed across a more than twofold elevation in body mass (Table 1), inter-class differences were not observed for either their height-adjusted body mass or adiposity (Table 1; $P>0.05$), except for comparisons between the lightest and heaviest groups ($P<0.05$). When the impact of those differences on the scaling models was evaluated, significant effects were not observed for either measure ($P>0.05$). Those outcomes support the interpretation that the experimental sample was relatively homogeneous, and that the influences of relative body mass (proportionality) and adiposity had been sufficiently minimised for the current scaling purposes.

Integrity of the experimental data

The absolute oxygen consumption of the experimental subjects increased ~ 16 -fold from the supine, basal state [0.27 L min^{-1} (SD 0.03)] through to volitional exhaustion on the third maximal-exercise trial [4.25 L min^{-1} (SD 0.71)]. The respective heart rates were $59 \text{ beats min}^{-1}$ (SD

1) and $194 \text{ beats min}^{-1}$ (SD 10). Those data are consistent with values expected at the end of an incremental (running) forcing function, and, in combination with the non-significant changes in peak aerobic power across the three pilot trials, permit one to suggest that they represented valid measures of peak aerobic power. As expected, those data displayed a positive mass bias (Fig. 1), with values ranging from 3.01 L min^{-1} (extra-small subject: 59.8 kg) to 6.27 L min^{-1} (extra-large subject: 107.6 kg). When expressed ratiometrically, a negative mass bias became evident: $56.0 \text{ mL kg}^{-1} \text{ min}^{-1}$ (mean of the extra-small subjects) and $48.8 \text{ mL kg}^{-1} \text{ min}^{-1}$ (extra-large subjects).

Scaling-model assumptions

Both the body-mass and peak aerobic-power data were normally distributed (Table 2; Shapiro–Wilk test, $P>0.05$), as confirmed from the z scores for skewness and kurtosis, which were well within the pre-defined range for samples of this size (Table 2; ± 3.29 , $P>0.05$; Kim 2013). The raw and log-transformed independent and dependent variables were linearly related [Fig. 2A (linear: $r^2=0.65$) and Fig. 2C (log-linear: $r^2=0.67$)], with an even distribution of residuals (homoscedasticity) about those regressions (Fig. 2B and D, respectively). Indeed, the sum of those residuals approached zero (Fig. 2D; $r=-0.001$; $P>0.05$). Notwithstanding the intrinsic relationship between body mass and peak aerobic power, auto-correlation was not observed within either model (Table 2; Durbin–Watson test, $P>0.05$). Accordingly, both linear and allometric scaling were statistically justified.

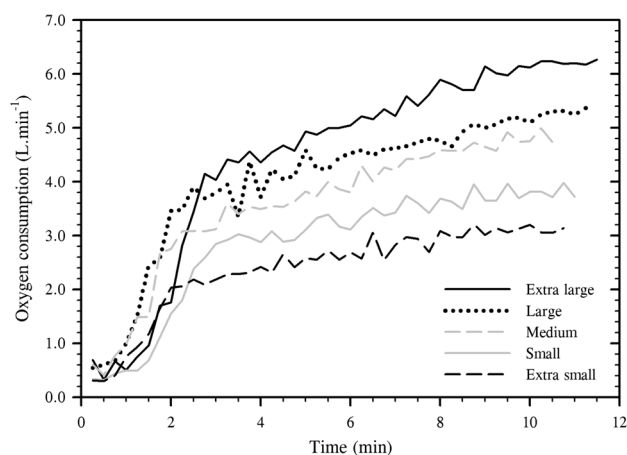


Fig. 1 Time-series oxygen-consumption data from five experimental participants, one from each body-mass class, during incremental treadmill running to volitional exhaustion. Data are 15-s averages measured continuously. Oxygen-consumption plateaux are infrequently observed, even in experienced and familiarised subjects

Table 2 Normality and auto-correlation (first-order correlation) tests performed on peak aerobic power (dependent variable) and body mass (raw and log-linear forms), as well as the model residuals (linear and log-linear regressions; $N=60$)

	Dependent and independent variables				Regression model residuals	
	Peak aerobic power		Body mass		Linear	Log-linear
	Raw (L min ⁻¹)	Transformed [Log ₁₀ (L min ⁻¹)]	Raw (kg)	Transformed [Log ₁₀ (kg)]		
Skewness statistic	0.18	-0.18	0.47	0.11	0.26	0.11
z score	0.59	-0.57	1.53	0.35	0.84	0.35
Kurtosis statistic	-0.29	-0.67	-0.27	-0.65	-0.46	-0.52
z score	-0.46	-1.10	-0.45	-1.06	-0.75	-0.86
Shapiro–Wilk statistic	0.97	0.96	0.96	0.98	0.98	0.98
Durbin–Watson statistic	–	–	–	–	2.04	2.00

Significant differences from a normal distribution were not observed, nor were there significant autocorrelations between those variables ($P > 0.05$)

A Shapiro–Wilk test statistic of 1.00 represents a perfect, normal distribution (Shapiro and Wilk 1965). A Durbin–Watson test statistic of 2.00 corresponds to the absence of auto-correlation between the dependent and independent variables (Durbin and Watson 1950, 1951)

Scaling peak aerobic power

The ratio standard demonstrated a visually [Figs. 2E and 3 (dashed line)] and statistically inferior fit (RMSE = 0.456; AIC = 80.033), when compared with the linear (Eq. 1B) and allometric models (Eq. 2B; the corresponding statistical parameters appear in subsequent paragraphs). The positive, absolute body-mass bias for peak aerobic power (Fig. 2A; $r = 0.81$) was converted into a negative bias following ratiometric scaling ($r = -0.40$; $P < 0.05$; Fig. 2E), resulting in the mass-specific data for the ten heaviest participants (48.77 mL kg⁻¹ min⁻¹) differing significantly from the corresponding data for the ten lightest individuals (56.03 mL kg⁻¹ min⁻¹; $P < 0.05$). Thus, ratiometric analysis did not remove the inherent mass bias, but merely changed its form, so that method was not further explored.

When least-squares, linear regression was used to model that relationship [Figs. 2A, 3 (dotted line); Eq. 1B], a statistically strong linear model was obtained (RMSE = 0.418; AIC = 71.457), which accounted for approximately 65% (r^2) of the variation in oxygen consumption on the basis of body mass. Whilst the goodness-of-fit indices (RMSE and AIC) for both the linear and ratiometric models were similar, only the linear-regression model satisfactorily normalised data for variations in body mass (Fig. 2B; $r = 0.005$; $P > 0.05$). Thus, goodness-of-fit, on its own, was insufficient for evaluating the appropriateness of those models. Nevertheless, predictive errors were evident when the linear model was applied to masses lower than the experimental body-mass range. That was epitomised by a significant, non-zero ordinate intercept (Fig. 3; 1.05 L min⁻¹; $P < 0.05$). Therefore, whilst linear scaling was statistically suitable, it was unable to model

that relationship for individuals lighter than the smallest experimental subject, and its rejection was theoretically justified.

$$\text{Peak aerobic power} = 1.049\{0.325 - 1.677\} + 0.040\{0.032 - 0.050\} \times \text{mass}, \quad (1B)$$

Where oxygen consumption units are L min⁻¹ and body mass is in kg, 95% confidence intervals are shown within the braces.

In contrast, allometric scaling implicitly satisfied the biological criterion [Fig. 3 (solid line); Eq. 2B], whilst simultaneously providing a statistically strong fit to the experimental data ($P < 0.05$). That model did not differ significantly from the linear-regression outcome [$r^2 = 0.67$ (log-linear format; Eq. 3B); RMSE = 0.414; AIC = 70.560; F test > 0.05]. Accordingly, allometric modelling was superior to linear scaling for both describing and normalising the relationship between peak aerobic power and body mass in healthy, adult males. The allometric model was also expanded to include an ordinate intercept (metabolic rate = $b \times \text{mass}^k + c$), as was performed when analysing basal and resting data (Bowes et al. 2021), and to evaluate the possibility that it might further improve the model fit. However, that model could not be confidently applied, due to the absence of data below 56 kg. Therefore, just as had been found with the resting data and body masses > 10 kg, the addition of an intercept did not improve the model fit.

$$\text{Peak aerobic power} = 0.160\{0.075 - 0.292\} \times \text{mass}^{0.750\{0.615 - 0.924\}}, \quad (2B)$$

where oxygen consumption units are L min⁻¹ and body mass is in kg, 95% confidence intervals are shown within the braces.

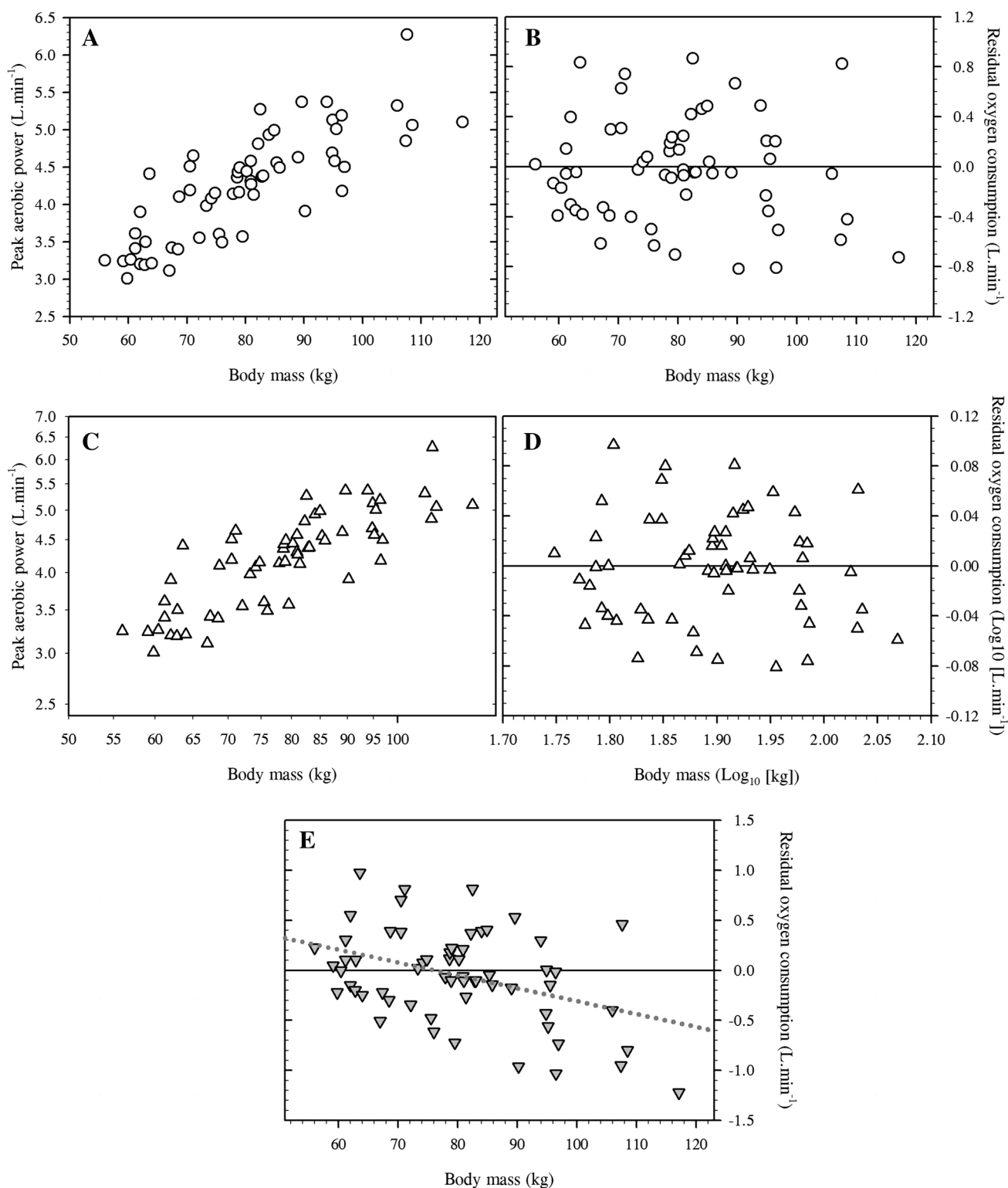


Fig. 2 Scatter plots of the independent (body mass) and dependent (peak aerobic power) variables from the experimental participants ($N=60$) obtained at volitional exhaustion. Those data were used to evaluate the assumptions of linearity between those variables, and are presented as absolute values on both linear (A) and logarithmic scales (C). The respective residual plots for both formats are presented in B

and D, with E containing the residuals resulting from the ratiometric analysis of the data presented in A. Each residual represents the difference between the measured peak aerobic power and the predicted value derived for each body mass, as determined from linear-regression analysis. Figures 2B, D and E were then used to evaluate homoscedasticity

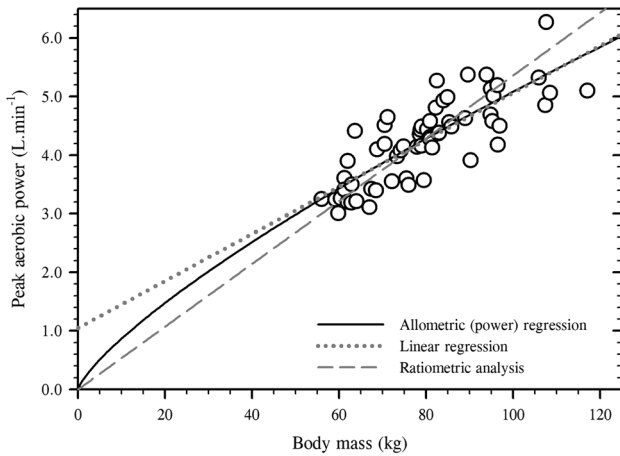


Fig. 3 Allometric (power: Eq. 2B) and linear regression (least-squares, first-order polynomial: Eq. 1B) models of the relationship between absolute, peak aerobic power and body mass within the experimental subjects ($N=60$; circles). Also presented is the commonly used, ratiometric treatment of those data. The non-zero ordinate intercept (1.05 L min^{-1} ; $r^2=0.65$; $P<0.05$) for the linear model violates the biological assumption underpinning all scaling models (presence of an origin intercept), and is thermodynamically impossible. The allometric model revealed a mass exponent of 0.75 ($r^2=0.67$; $P<0.05$)

$$\begin{aligned} \text{Log}_{10}(\text{peak aerobic power}) = & -0.863\{-1.140 \text{ to } -0.595\} \\ & + 0.784\{0.641 - 0.935\} \\ & \times \text{Log}_{10}(\text{mass}), \end{aligned} \quad (3B)$$

where oxygen consumption units are L min^{-1} and body mass is in kg, 95% confidence intervals are shown within the braces.

In combination with previous analyses of basal metabolism (Bowes et al. 2021), it was evident that the metabolic rate of men scaled allometrically with body mass across the complete metabolic range (aerobic scope), and hypothesis one was accepted. The human exponent for peak aerobic power ($\text{mass}^{0.75}$) did not differ significantly from the interspecific, mammalian exponent ($\text{mass}^{0.87}$; $P>0.05$; Weibel et al. 2004), but was significantly smaller than the maximal exponent postulated by Darveau et al. (2002; $\text{mass}^{0.92}$; $P<0.05$).

Allometric scaling of steady-state oxygen consumption

Having demonstrated elsewhere the appropriateness of allometrically scaling steady-state oxygen consumption (Bowes 2018), data obtained from the experimental subjects during horizontal walking (4.8 km h^{-1}) were scaled in both allometric (Eq. 2C) and log-linear formats (Eq. 3C). A statistically strong allometric model was again obtained (RMSE = 0.071).

Steady-state oxygen consumption

$$= 0.023\{0.014 - 0.037\} \times \text{mass}^{0.865\{0.759-0.982\}}, \quad (2C)$$

where oxygen consumption units are L min^{-1} and body mass is in kg, 95% confidence intervals are shown within the braces.

$\text{Log}_{10}(\text{steady-state oxygen consumption})$

$$\begin{aligned} = & -1.630\{-1.855 \text{ to } -1.421\} \\ & + 0.857\{0.748 - 0.976\} \times \text{Log}_{10}(\text{mass}), \end{aligned} \quad (3C)$$

where oxygen consumption units are L min^{-1} and body mass is in kg, 95% confidence intervals are shown within the braces.

Scaling oxygen consumption across metabolic intensities

Since 60 individuals participated in all three trials, and since allometric modelling was superior for each metabolic state, a unique opportunity arose to evaluate the effect of a change in metabolic intensity on the scaling exponent. To achieve that, absolute oxygen-consumption data for the basal state were scaled (Eq. 2D; RMSE = 0.019; AIC = -301.908), rather than using the metabolic rates reported elsewhere (MJ day^{-1} ; Bowes et al. 2021), and compared with both the steady-state and peak scaling models. The basal scaling exponent ($\text{mass}^{0.57}$) differed slightly, but not significantly ($P>0.05$), from the exponent reported for basal metabolic rate ($\text{mass}^{0.50-0.55}$), due to differences in subject numbers between investigations (60 and 68, respectively). Relative to the basal exponent, the exponents for both steady-state (Eq. 2C; $\text{mass}^{0.87}$) and maximal exercise (Eq. 2B; $\text{mass}^{0.75}$) were significantly greater ($P<0.05$), so hypothesis two was accepted. However, the exponents for steady-state and maximal exercise did not differ significantly ($P>0.05$). Therefore, a stepwise elevation in the body-mass exponent appeared not to accompany increases in metabolic rate, and hypothesis three was unable to be accepted.

$$\begin{aligned} \text{Basal oxygen consumption} = & 0.023\{0.015 - 0.038\} \\ & \times \text{mass}^{0.567\{0.458-0.671\}}, \end{aligned} \quad (2D)$$

where oxygen consumption units are L min^{-1} and body mass is in kg, 95% confidence intervals are shown within the braces.

To evaluate the significance of the exponent change when moving from resting to exercising states, and also to assess the interchangeability of those exponents, peak data from the experimental sample were re-scaled using the basal exponent ($\text{mass}^{0.57}$). The residuals from that analysis revealed a

significant, positive mass bias ($r=0.34$; $P<0.05$), thereby demonstrating the importance of using the scaling exponent appropriate to the metabolic intensity of interest, even when analysing data from the same individuals. However, since the exercising exponents were not significantly different ($P>0.05$), it seems that only two exponents (basal and exercising) are necessary to account for mass-dependent variations in oxygen consumption across the metabolic range of healthy men.

An assessment of scaling robustness: peak aerobic power

The robustness of the allometric model for maximal treadmill exercise (Eq. 2C) was evaluated by applying it to data from the in-house database. Both sets of data were overlaid (Fig. 4), revealing good visual agreement, with the possible exception of two to three historical subjects (~ 80 kg).

Nevertheless, data clustering was evident, with $\sim 80\%$ of the historical participants having masses between 66 and 87 kg, so the body-mass distribution was uneven. Furthermore, 76% of those values were above the allometric regression (Fig. 4), resulting in under-predictions when using Eq. 2C. Whilst the average prediction error (residual) was 0.32 L min^{-1} (SD 0.47), the greatest absolute difference was 1.55 L min^{-1} (range -0.84 to 1.55), with 71% of the absolute residuals exceeding 0.25 L min^{-1} . For the experimental participants, the residuals ranged from -0.77 to 0.92 L min^{-1} , with 65% of the absolute residuals exceeding 0.25 L min^{-1} . The similarity of those outcomes indicates that the experimental model was sufficiently robust to scale the historical data, and hypothesis four was cautiously accepted.

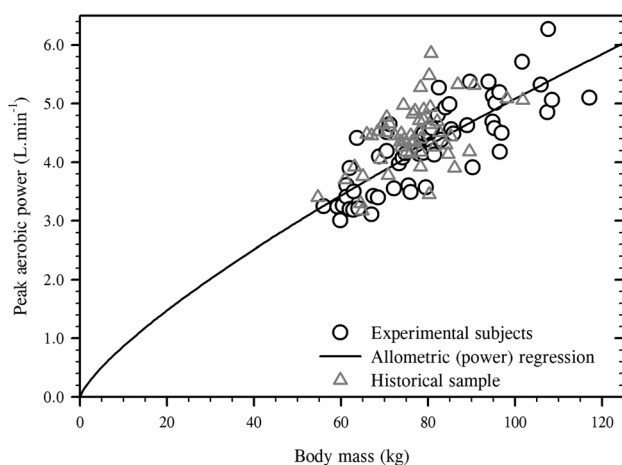


Fig. 4 Peak aerobic-power and body-mass data for the experimental subjects ($N=60$; circles), with the corresponding allometric regression, overlaid with the co-ordinates obtained from the in-house, historical sample ($N=54$; triangles)

Finally, to evaluate the impact of applying a unique allometric regression to the historical database, those data were independently modelled (Eq. 2E). Since those data did not conform to the criteria necessary for generating a meaningful, independent mathematical scaling model (White and Seymour 2005). Therefore, the purpose of that exercise was to evaluate the implications of that practice. From a statistical perspective, the in-house model (Eq. 2E) was similar to the experimental model (Eq. 2B) in its goodness-of-fit (ANCOVA = $P>0.05$; F test = $P>0.05$; RMSE = 0.458). However, the body-mass exponent ($\text{mass}^{0.55}$) approximated that observed for the basal oxygen consumption of the experimental subjects ($\text{mass}^{0.57}$; $P>0.05$). That outcome was theoretically unsound, and was not further considered.

$$\text{Peak aerobic power} = 0.404\{0.154 - 1.116\} \times \text{mass}^{0.553\{0.323-0.788\}}, \quad (2E)$$

where oxygen consumption units are L min^{-1} and body mass is in kg, 95% confidence intervals are shown within the braces.

Discussion

In this manuscript, the scaling of whole-body metabolism against body mass has been re-visited in healthy men during two steady-states (basal rest and walking), as well as during one of the most unsteady of physiological states; volitional exhaustion. Four significant outcomes have arisen. Firstly, whilst both the linear and non-linear scaling of peak aerobic power were supported statistically, ratiometric scaling merely converted a positive mass bias into a negative bias, and linear regression yielded a biologically impossible ordinate intercept. This confirms that the peak aerobic power of male adults scales allometrically with body mass ($\text{mass}^{0.75}$), as it does for other mammals ($\text{mass}^{-0.87}$; Taylor et al. 1981; Bishop 1999; Weibel et al. 2004). Therefore, across the entire metabolic range (aerobic scope) of these participants, oxygen consumption scaled allometrically. Secondly, the exponent for peak aerobic power was significantly larger than observed at rest ($\text{mass}^{0.57}$). Thirdly, whilst steady-state oxygen consumption also scaled allometrically with body mass ($\text{mass}^{0.87}$), its exponent did not differ significantly from that derived for peak aerobic power. Hence, the scaling exponent did not increase in a stepwise manner with increments in metabolic activity. Finally, the exponent for peak aerobic power could be used to normalise similar historical data, albeit with the expected noise. Nonetheless, the independent application of allometric scaling to those historical data, which infringed some sampling pre-requisites, resulted in an improbable relationship with body mass.

Linear versus non-linear scaling of peak aerobic power

The principal objective behind modelling these relationships is to remove the influence of inter-individual mass variations. The resulting mass-specific data may then be used to explore mechanistic questions, to make comparisons across individuals of varying size or to predict the effects of changes in overall mass during ambulatory activities, regardless of the reason for such changes (e.g., adiposity changes, load carriage). However, the validity of all models depends upon the mathematical precision of those procedures and the adherence of that relationship to biological principles. Whilst both are essential, the latter is of greater importance, since goodness-of-fit has little meaning if the relationship violates first principles (Sholl 1948; Motulsky and Ransnas 1987; Nevill et al. 1992; Rogers et al. 1995).

As shown by others, the peak aerobic power of humans (Vaage and Hermansen 1977; Secher et al. 1983; Bergh et al. 1991; Nevill et al. 1992; Rogers et al. 1995; Batterham and Jackson 2003; Markovic et al. 2007; Lolli et al. 2017) and other mammals (Taylor et al. 1981; Bishop 1999; Weibel et al. 2004) scales allometrically. Whilst linear analyses are frequently used for human data, ratiometric scaling is unable to remove the body-mass bias (Fig. 2E), and a single linear regression cannot be applied to the full body-mass range without violating the *First Law of Thermodynamics* as body mass approaches zero (Fig. 3). Thus, both linear procedures provided inaccurate descriptions of the underlying biological relationships, and yielded spurious outcomes when applied to morphologically diverse samples. Therefore, neither method is recommended for scaling peak aerobic power.

The novelty of the current procedures centres around the methods used by comparative physiologists (Taylor et al. 1981; Bishop 1999; Weibel et al. 2004; White and Seymour 2005; Sieg et al. 2009; White and Kearney 2014), for it is they who have most advanced this research, and developed recruitment and analytical principles for others to follow. Most notably, those principles require the recruitment of a sufficiently large sample of homogeneous participants covering a wide body-mass range, with those individuals distributed somewhat equally across that range. To our knowledge, the current procedures constitute the first attempt to thoroughly evaluate that relationship within humans. Collectively, these animal and human experiments provide convincing evidence that peak aerobic power varies non-linearly with body mass, as does basal metabolic rate (Bowes et al. 2021). Therefore, across the complete metabolic range, oxygen consumption for humans can be described using a simple allometric relationship with mass.

Nevertheless, a reduced sensitivity of the allometric model for peak aerobic power was observed, relative to data obtained from the two steady-state walking trials, as

indicated by the poorer goodness-of-fit, the broader confidence intervals for the linear coefficient and greater discrepancies between the raw and log-transformed regression coefficients. Since peak data are inherently more variable than steady-state data, and since the experimental samples covered only a relatively narrow body-mass range, the reduced sensitivity was not unsurprising. Consequently, it is possible that a small change in the regression coefficients, but within the 95% confidence intervals, could produce an equally good model fit. However, bootstrapping was used within these analyses to maximise the likelihood of determining the most appropriate coefficients. Notwithstanding these limitations, it is recommended that the current maximal-exercise exponent ($\text{mass}^{0.75}$) be used to derive mass-independent, oxygen-consumption data from data obtained during maximal treadmill exercise in men. However, there exists considerable variability in the scaling exponents reported for peak aerobic power, which must lead to some disquiet amongst students and applied physiologists.

What then is the allometric scaling exponent for peak aerobic power in humans?

The exponent for the peak aerobic power of the experimental subjects ($\text{mass}^{0.75}$) did not differ significantly from the inter-specific exponent for maximally exercising, non-human mammals ($\text{mass}^{0.87}$; $P > 0.05$; Weibel et al. 2004). Under resting conditions, the non-human, mammalian exponent ($\text{mass}^{0.67}$; White and Seymour 2003) was also larger than that observed for the basal metabolic rate of the current experimental subjects ($\text{mass}^{0.50-0.55}$; $P < 0.05$; Bowes et al. 2021). That difference was attributable to a species specificity of those resting exponents (White et al. 2009; White and Kearney 2014), which may not obtain during maximal exercise. However, within the human literature, one finds scaling exponents for peak aerobic power varying between 0.67 and 0.94 (e.g. Secher et al. 1983 [$\text{mass}^{0.67}$]; Bergh et al. 1991 [$\text{mass}^{0.71}$]; Nevill et al. 1992 [$\text{mass}^{0.67}$]; Rogers et al. 1995 [$\text{mass}^{0.75}$]; Batterham and Jackson 2003 [$\text{mass}^{0.94}$]; Markovic et al. 2007 [$\text{mass}^{0.67-0.89}$]; Lolli et al. 2017 [$\text{mass}^{0.70}$]).

It seems that the highest design priority for those last seven (human) studies might have been to maximise their sample sizes, with all but two reports (Secher et al. 1983; Rogers et al. 1995) using large databases, or data extracted from the literature. Whilst large samples are beneficial, the quality of those data must be weighted more highly (White and Seymour 2005). Indeed, in none of those studies does one find evidence for obtaining an even distribution of body masses or for minimising inter-subject heterogeneity for variables known to independently influence peak aerobic power. Instead, two groups used both genders, as well as adolescents and adults (Rogers et al. 1995; Lolli et al. 2017), two used data from both mass-supported and unsupported

exercise (Secher et al. 1983; Lolli et al. 2017), two groups studied elite athletes (Secher et al. 1983; Bergh et al. 1991), with two others using data from elderly (Batterham and Jackson 2003; Lolli et al. 2017), and even obese participants (Lolli et al. 2017). It is suggested that differences in the supply of, and demand for, energy among such participants might independently modify the scaling exponents.

To illustrate that variability, it is helpful to consider the largest of those investigations (Lolli et al. 2017); 6514 individuals across 36 primary investigations. Twenty-two of those studies focussed on, or included a significant number of adolescents; one study used obese boys and three emphasised elderly adults. If the current subject exclusion criteria had been applied, 72% of those studies would have been eliminated (2005 participants). Similarly, women would have been excluded ($N=916$). To standardise health status, people with a high risk of metabolic syndrome ($N=571$) and physical disabilities ($N=45$; spinal-cord injuries), would have been eliminated. Thus, > 50% of the participants would have been excluded if steps were taken to minimise heterogeneity. The authors were aware of that variability, and attempted to counteract that effect statistically.

Another important consideration was the inclusion of data from mass-supported exercise. In the first instance, a significant proportion of the metabolic rate during treadmill exercise must be attributable to supporting an upright posture and to vertical displacements of the body mass. Thus, running activates a greater muscle mass than does wheelchair ergometry ($N=45$), and, to a lesser extent, cycling ($N=22$). Rowing is different again, possibly activating the greatest muscle mass, but it is unclear how many rowers were tested. The inclusion of those data within analyses aimed at producing mass-independent data seems counter-intuitive, as does the derivation of fat-free scaling exponents, when adipose tissue, just like any load, contributes to the overall carried mass and its metabolic cost during load-bearing activities.

Nevertheless, the scaling exponent for peak aerobic power derived by Lolli et al. (2017; $\text{mass}^{0.70}$) is close to that obtained from the current experimental subjects ($\text{mass}^{0.75}$), both of which are centrally positioned within the human data range. What are the implications of that similarity? Those authors counselled against using their exponent for scaling. Instead, they recommended scaling to fat-free mass. The current authors, whilst not challenging those analyses, have difficulty accepting the inclusion of such heterogeneous samples. The similarity of the two exponents is difficult to reconcile, and is perhaps a chance occurrence.

To provide some insight into that outcome, in-house, historical data ($N=54$) were analysed (Eq. 2E). Although those data were both statistically and visually similar to the experimental data (Fig. 4), the equivalence of the scaling exponent ($\text{mass}^{0.55}$) to the basal oxygen-consumption exponent ($\text{mass}^{0.57}$) demonstrates that, unless appropriate recruitment criteria are

followed (White and Seymour 2005), the generation and subsequent use of scaling models should be approached with caution. Indeed, this highlights the sensitivity of the scaling exponent to even slight violations of those criteria, with data from similarly aged, morphologically alike, endurance-trained males yielding a significantly different outcome.

Is the scaling exponent sensitive to an elevation in metabolic intensity?

The scaling exponent for peak aerobic power ($\text{mass}^{0.75}$) increased significantly relative to the basal state ($\text{mass}^{0.57}$). Since that difference was observed within the same sample, then it supports the hypothesis that the exponent for oxygen consumption is sensitive to an elevation in metabolic intensity, at least when moving from a basal state to maximal exercise. The mechanisms responsible for that elevation include proportional increases in the metabolically active body mass, in combination with an intensity-dependent progression through the metabolic pathways (Weibel 2002; Darveau et al. 2002).

The exponent change between rest and maximal exercise was anticipated. It was also hypothesised that the exponent would increase in a stepwise manner with increments in metabolic rate. That did not occur; the walking and maximal running were equivalent ($P>0.05$). Therefore, a progressive elevation in the exponent (an allometric cascade) was not evident. It is possible that, unlike the change in skeletal-muscle activity between rest and running, the recruitment of inactive tissues was minimal when moving from walking to maximal running, with the primary change being a variation in the metabolic rate of already active muscles. It is also possible that the variability of those peak data, or even gait differences, may have masked an exercise-intensity effect. Resolution of this issue awaits evidence from a similarly robust experimental design with a much larger sample.

Nonetheless, within each scaling model (Eqs. 1–3), both coefficients could change. In the current analyses, the principal emphasis was upon the exponent (k), with the linear coefficient (b) not being analysed statistically. Inspection of Eqs. 2B–2D reveals that the linear coefficients for the basal and walking conditions were smaller and less variable. However, within the current mass range, those differences did not have a meaningful effect on the capacity of the maximal-exercise exponent to generate mass-independent data across both exercise intensities. This implies that unloaded, horizontal exercise could be scaled across all intensities using a common exponent ($\text{mass}^{0.75-0.87}$), and the recommended exponent for that purpose is 0.75. What remains to be established is whether or not those same relationships obtain when loads are carried and when ambulatory gradients are modified.

Conclusion

From this experiment, it has been demonstrated that, across the entire metabolic range (aerobic scope) of healthy men, oxygen consumption scales allometrically with body mass. During maximal exercise, ratiometric scaling was unable to remove the body-mass bias, whilst linear regression produced biologically inappropriate data for low body masses. Neither of those methods is recommended. The size of the scaling exponent increased significantly between resting ($\text{mass}^{0.57}$) and maximal-exercise states ($\text{mass}^{0.75}$), but remained relatively stable when changing from steady-state walking ($\text{mass}^{0.87}$) to maximal running. Therefore, across all non-resting, metabolic intensities, it is recommended that the exponent for peak aerobic power ($\text{mass}^{0.75}$) be used to derive mass-specific oxygen consumption for adult men [$Y_{\text{adjusted}} = Y_{\text{measured}} / \text{mass}^k$ ($k = \text{exponent}$)]. For those remaining unconvinced, it is recommended that, instead of linear analysis, absolute oxygen consumption (L min^{-1}) and body masses should be reported (including data ranges), with individual data provided whenever possible.

Acknowledgements This project was supported by grants from the Defence Science and Technology Group (Australia). The opinions expressed in this paper are those of the authors, and do not reflect the official policy or position of the Defence Science and Technology Group, or the Australian Government. HMB was supported by an International Postgraduate Tuition Award (University of Wollongong, Australia). SRN is supported by a Postdoctoral Fellowship from the Human and Environmental Physiology Research Unit (University of Ottawa, Canada). Parts of this manuscript have been published in abstract form (Bowes et al. 2017).

Author contributions These experiments were undertaken at the Centre for Human and Applied Physiology (University of Wollongong, Australia). NAST conceived, and NAST, CAB, HMB and GEP designed and planned this research. HMB ran the experiments and, with CAB, was responsible for the collection and analysis of data from the experimental participants. GEP and SRN were responsible for the historical data. All authors contributed to data analysis and its interpretation. NAST and HMB wrote the manuscript, with all authors contributing. All authors read and approved the final version of the manuscript, and its submission for publication, and they agree to be accountable for this work, ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All persons designated as authors qualify for authorship, and all who qualify for authorship are listed as authors.

Declarations

Conflict of interest There are no conflicts of interest.

References

- Batterham AM, Jackson AS (2003) Validity of the allometric cascade model at submaximal and maximal metabolic rates in exercising men. *Respir Physiol Neurobiol* 135:103–106
- Bergh U, Sjödin B, Forsberg A, Svedenhag J (1991) The relationship between body mass and oxygen uptake during running in humans. *Med Sci Sports Exerc* 23:205–211
- Bishop CM (1999) The maximum oxygen consumption and aerobic scope of birds and mammals: getting to the heart of the matter. *Proc R Soc Lond B* 266:2275–2281
- Booth JD, Wilsmore BR, MacDonald AD, Zeyl A, McGhee S, Calvert D, Marino FE, Storlien LH, Taylor NAS (2001) Whole-body pre-cooling does not alter human muscle metabolism during submaximal exercise in the heat. *Eur J Appl Physiol* 84:587–590
- Bowes HM (2018) Human metabolic allometry from basal to maximal ambulatory states, including load carriage and its distribution. University of Wollongong, Doctor of Philosophy
- Bowes HM, Burdon CA, Taylor NAS (2017) The allometric scaling of aerobic power in adult humans, across the physiological range. *Proc Aust Physiol Soc* 48:117P
- Bowes HM, Burdon CA, Taylor NAS (2021) The scaling of human basal and resting metabolic rates. *Eur J Appl Physiol* 121:193–208
- Clark JM, Hagerman FC, Gelfand R (1983) Breathing patterns during submaximal and maximal exercise in elite oarsmen. *J Appl Physiol* 55:440–446
- Craddock JC, Neale EP, Probst YC, Peoples GE (2021) A cross-sectional comparison of the whole blood fatty acid profile and omega-3 index of male vegan and omnivorous endurance athletes. *J Am Coll Nutr*. <https://doi.org/10.1080/07315724.2021.1886196>
- Darveau CA, Suarez RK, Andrews RD, Hochachka PW (2002) Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature* 417:166–170
- Durbin J, Watson GS (1950) Testing for serial correlation in least squares regression: I. *Biometrika* 37:409–428
- Durbin J, Watson GS (1951) Testing for serial correlation in least squares regression: II. *Biometrika* 38:159–177
- Elia M (1992) Organ and tissue contribution to metabolic rate. In: Kinney JM, Tucker HN (eds) *Energy metabolism: tissue determinants and cellular corollaries*. Raven Press, New York, pp 61–80
- Faul F, Erdfelder E, Lang A-G, Buchner A (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 39:175–191
- Gallagher D, Belmonte D, Deurenberg P, Wang Z, Krasnow N, Pi-Sunyer FX, Heymsfield SB (1998) Organ-tissue mass measurement allows modeling of REE and metabolically active tissue mass. *Am J Physiol* 275:E249–E258
- Heusner AA (1982) Energy metabolism and body size. I. Is the 0.75 mass exponent of Kleiber's equations a statistical artifact? *Respir Physiol* 48:1–12
- Hochachka PW (1994) *Muscles as molecular and metabolic machines*. CRC Press, Boca Raton
- Hochachka PW, Darveau C-A, Andrews RD, Suarez RK (2003) Allometric cascade: a model for resolving body mass effects on metabolism. *Comp Biochem Physiol Mol Integr Physiol* 134:675–691
- Holliday MA (1971) Metabolic rate and organ size during growth from infancy to maturity and during late gestation and early infancy. *Pediatrics* 47:169–179
- Holliday MA, Potter D, Jarrah A, Bearg S (1967) The relation of metabolic rate to body weight and organ size. *Pediatr Res* 1:185–195
- Hoppeler H, Kayar SR, Claassen H, Uhlmann E, Karas RH (1987) Adaptive variation in the mammalian respiratory system in relation to energetic demand: III. Skeletal muscles: setting the demand for oxygen. *Respir Physiol* 69:27–46

- Kim H-Y (2013) Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restor Dent Endod* 38:52–54
- Kleiber M (1932) Body size and metabolism. *Hilgardia* 6:315–353
- Kleiber M (1950) Physiological meaning of regression equations. *J Appl Physiol* 2:417–423
- Krogh A (1916) *The respiratory exchange of animals and man*. Longmans Green and Company, London
- Lolli L, Batterham AM, Weston KL, Atkinson G (2017) Size exponents for scaling maximal oxygen uptake in over 6500 humans: a systematic review and meta-analysis. *Sports Med* 47:1405–1419
- Markovic G, Vucetic V, Nevill AM (2007) Scaling behaviour of O_2 in athletes and untrained individuals. *Ann Hum Biol* 34:315–328
- Mitchell JH, Blomqvist G (1971) Maximal oxygen uptake. *N Engl J Med* 284:1018–1022
- Motulsky HJ, Ransnas LA (1987) Fitting curves to data using nonlinear regression: a practical and nonmathematical review. *FASEB J* 1:365–374
- Müller MJ, Langermann D, Gehrke I, Later W, Heller M, Glüer CC, Heymsfield SB, Bosy-Westphal A (2011) Effect of constitution on mass of individual organs and their association with metabolic rate in humans—a detailed view on allometric scaling. *PLoS ONE* 6:22721
- Nevill AM, Ramsbottom R, Williams C (1992) Scaling physiological measurements for individuals of different body size. *Eur J Appl Physiol* 65:110–117
- Packard GC, Boardman TJ (1988) The misuse of ratios, indices, and percentages in ecophysiological research. *Physiol Zool* 61:1–9
- Peoples GE, Lee DS, Notley SR, Taylor NAS (2016) The effects of thoracic load carriage on maximal ambulatory work tolerance and acceptable work durations. *Eur J Appl Physiol* 116:635–646
- Peoples GE, Hingley L, Caldwell JN, Taylor NAS (2017) The effects of increasing thoracic load carriage, using a backpack and body-armor ensemble, on peak aerobic power and exercise tolerance. *J Sci Med Sports* 20:S105–S106
- Rogers DM, Olson BL, Wilmore JH (1995) Scaling for the VO_2 -to-body size relationship among children and adults. *J Appl Physiol* 79:958–967
- Ross WD, Wilson NC (1974) A strategem for proportional growth assessment. *Acta Paediatr Belg* 28:169–182
- Saltin B, Gollnick PD (1983) Skeletal muscle adaptability: significance for metabolism and performance. In: Peachey LD (ed) *Handbook of physiology*. Section 10 skeletal muscle. American Physiological Society, Washington DC, pp 555–631
- Sarrus F, Rameaux J (1838) Rapport sur un mémoire adressé à l'Académie Royale de Médecine. *Bull De L'académie R De Méd* 3:1094–1100
- Secher NH, Vaage O, Jensen K, Jackson RC (1983) Maximal aerobic power in oarsmen. *Eur J Appl Physiol* 51:155–162
- Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52:591–611
- Shephard RJ, Allen C, Benade AJS, Davies CTM, Di Prampero PE, Hedman R, Merriman JE, Myhre K, Simmons R (1968) The maximum oxygen intake: An international reference standard of cardio-respiratory fitness. *Bull World Health Organ* 38:757–764
- Sholl D (1948) The quantitative investigation of the vertebrate brain and the applicability of allometric formulae to its study. *Proc R Soc Lond B* 135:243–258
- Sieg AE, O'Connor MP, McNair JN, Grant BW, Agosta SJ, Dunham AE (2009) Mammalian metabolic allometry: do intraspecific variation, phylogeny, and regression models matter? *Am Nat* 174:720–733
- Suarez RK, Darveau C-A, Childress JJ (2004) Metabolic scaling: a many-splendoured thing. *Comp Biochem Physiol B Biochem Mol Biol* 139:531–541
- Taylor CR, Maloiy GMO, Weibel ER, Langman VA, Kamau JMZ, Seeherman HJ, Heglund NC (1981) Design of the mammalian respiratory system. III. Scaling maximum aerobic capacity to body mass: wild and domestic mammals. *Respir Physiol* 44:25–37
- Taylor NAS, Lewis MC, Notley SR, Peoples GE (2012) A fractionation of the physiological burden of the personal protective equipment worn by firefighters. *Eur J Appl Physiol* 112:2913–2921
- Vaage O, Hermansen L (1977) Figure 11–4. In: Åstrand PO, Rodahl K (eds) *Textbook of work physiology: physiological bases of exercise*. McGraw-Hill, New York
- Weibel ER (2002) The pitfalls of power laws. *Nature* 417:131–132
- Weibel ER, Bacigalupe LD, Schmitt B, Hoppeler H (2004) Allometric scaling of maximal metabolic rate in mammals: muscle aerobic capacity as determinant factor. *Respir Physiol Neurobiol* 140:115–132
- White CR, Kearney MR (2014) Metabolic scaling in animals: methods, empirical results, and theoretical explanations. *Compr Physiol* 4:231–256
- White CR, Seymour RS (2003) Mammalian basal metabolic rate is proportional to body mass^{2/3}. *Proc Natl Acad Sci USA* 100:4046–4049
- White CR, Seymour RS (2005) Sample size and mass range effects on the allometric exponent of basal metabolic rate. *Comp Biochem Physiol Mol Integr Physiol* 142:74–78
- White CR, Blackburn TM, Seymour RS (2009) Phylogenetically informed analysis of the allometry of mammalian basal metabolic rate supports neither geometric nor quarter power scaling. *Evol Int J Org Evol* 63:2658–2667

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.