Check for updates

A generalized equation for predicting peak oxygen consumption during treadmill exercise testing: mitigating the bias from total body mass scaling

Everton J. Santana[1,2,3], Nicholas Cauwenberghs[3], Bettia E. Celestin[2,4], Tatiana Kuznetsova[3], Christopher Gardner[5], Ross Arena[6], Leonard A. Kaminsky[7], Matthew P. Harber[8], Euan Ashley[1], Jeffrey W. Christle[1†], Jonathan Myers[1,9†] and Francois Haddad[1,2*†]

[1]Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, United States, [2]Stanford Cardiovascular Institute, Stanford University, Stanford, CA, United States, [3]Research Unit Hypertension and Cardiovascular Epidemiology, Department of Cardiovascular Sciences, University of Leuven, Leuven, Belgium, [4]Division of Pathology, Stanford University School of Medicine, Stanford, CA, United States, [5]Department of Medicine, Stanford Prevention Research Center, Stanford University, Stanford, CA, United States, [6]Department of Physical Therapy, College of Applied Sciences, University of Illinois at Chicago, Chicago, IL, United States, [7]Fisher Institute for Health and Well-Being, Ball State University, Muncie, IN, United States, [8]Clinical Exercise Physiology, Ball State University, Muncie, IN, United States, [9]Cardiology Division, Veterans Affairs Palo Alto Healthcare System, Palo Alto, CA, United States

Background: Indexing peak oxygen uptake (VO$_2$peak) to total body mass can underestimate cardiorespiratory fitness (CRF) in women, older adults, and individuals with obesity. The primary objective of this multicenter study was to derive and validate a body size-independent scaling metric for VO$_2$peak. This metric was termed exercise body mass (EBM).

Method: In a cohort of apparently healthy individuals from the Fitness Registry and the Importance of Exercise National Database, we derived EBM using multivariable log-normal regression analysis. Subsequently, we developed a novel workload (WL) equation based on speed (Sp), fractional grade (fGr), and heart rate reserve (HRR). The generalized equation for VO$_2$peak can be expressed as VO$_2$peak = Cst × EBM × WL, where Cst is a constant representing the VO$_2$peak equivalent of one metabolic equivalent of task. This generalized equation was externally validated using the Stanford exercise testing (SET) dataset.

Results: A total of 5,618 apparently healthy individuals with a respiratory exchange ratio >1.0 (57% men, mean age 44 ± 13 years) were included. The EBM was expressed as Mass (kg)$^{0.63}$ × Height (m)$^{0.53}$ × 1.16 (if a man) × exp ($-0.39 \times 10^{-4} \times$ age^2), which was also approximated using simple sex-specific additive equations. Unlike total body mass, EBM provided body size-independent scaling across both sexes and WL categories. The generalized VO$_2$peak equation was expressed as 11 × EBM × [2 + Sp (in mph) × (1.06 + 5.22 × fGr) + 0.019 × HRR] and had an R^2 of 0.83, $p < 0.001$. This generalized equation mitigated bias in VO$_2$peak estimations across age, sex, and body mass index subgroups and was validated in the SET registry, achieving an R^2 of 0.84 ($p < 0.001$).

Conclusions: We derived a generalized equation for measuring VO$_2$peak during treadmill exercise testing using a novel body size-independent scaling metric. This approach significantly reduced biases in CRF estimates across age, sex, and body composition.

1 Introduction

Cardiorespiratory fitness (CRF) is the capacity of the circulatory and respiratory systems to deliver oxygen to skeletal muscles during physical activity and is widely recognized as one of the strongest predictors of survival in the general population (1–3). In clinical practice, CRF is usually measured by peak oxygen consumption (VO$_2$peak), which is typically indexed to total body mass, i.e., VO$_2$peak/mass (Figure 1A) (1).

While indexing VO$_2$peak to total body mass partially adjusts for body size differences, it does not fully account for variations in metabolically active tissue or differences in body composition. In fact, studies have shown that indexing to total body mass underestimates CRF in women, older adults, and individuals with obesity (4, 5). An ideal metric for scaling VO$_2$peak would provide unbiased CRF estimates after accounting for differences in exercise capacity (6). Practically, this is tested by assessing the relationship between indexed VO$_2$peak and body size (Figure 1B). To improve VO$_2$peak indexing, several approaches have been considered including allometric scaling (body size^a, where a is an exponent), the use of different body compartments, or a combination of both approaches. At the molecular level, the body can be divided into lean body mass (LBM), fat mass (FM), and mineral content; at the cellular level, the body is divided into body cell mass (BCM), extracellular solids, extracellular fluids, and fat (Figure 1C) (5, 7). Compared to LBM, BCM better accounts for age-related increases in extracellular solids or fluids, which are generally not metabolically active (5, 7).

Evidence for allometric scaling of total body mass or LBM comes from both animal and clinical studies. In a 1981 study, Taylor et al. found that maximal VO$_2$ scales allometrically with body mass, with an exponent of 0.79 (95% CI: 0.75–0.83) in wild animals and 0.77 (95% CI: 0.68–0.85) in domestic animals (8). In a recent meta-analysis of 36 studies involving 6,514 participants, Lolli et al. reported that the pooled allometric exponent for indexed aerobic capacity was 0.70 (95% CI: 0.64–0.76) for total body mass and 0.90 (95% CI: 0.83–0.96) for fat-free mass (9). In the DR's EXTRA study, Krachler et al. also confirmed that VO$_2$peak scales more closely to LBM compared to total mass and Kohler et al. showed the potential added benefit of BCM-based scaling (5).

Although the allometric studies showed that scaling to mass$^{0.70}$ provided better body size independence than indexing to total body mass, we hypothesized that including height, sex, and age in the allometric scaling would better account for metabolically active

tissue. Therefore, the first objective of this work was to develop and validate a body size-independent scaling metric for VO$_2$peak. Here, we introduce a novel metric, referred to as exercise body mass (EBM) that includes mass, height, age, and sex factors (Figure 1D). Building on the EBM concept, our second objective was to develop a well-calibrated workload (WL) equation that not only adjusts for treadmill speed (Sp) and fractional grade (fGr) (10) but also for heart rate reserve (HRR) (11). Integrating the two equations, we derived and validated a generalized VO$_2$peak equation expressed as constant × EBM × WL, where the constant represents one metabolic equivalent of task (MET). The calibration of this equation was compared to the standard equation VO$_2$peak = 3.5 × mass × WL across the spectrum of age, sex, and body mass index (BMI). Finally, we determined whether the calibration for the generalized equation also holds in patients with heart failure (HF), as these patients often have sarcopenia or frailty (12).

2 Materials and methods

2.1 Study cohorts

This study included two exercise cohorts, i.e., the Fitness Registry and the Importance of Exercise National Database (FRIEND) registry (13) and the Stanford exercise testing (SET) registry (14). A third cohort, the Diet Intervention Examining the Factors Interacting with Treatment Success (DIETFITS) (15), was used to determine associations between EBM and directly measured LBM.

The FRIEND registry is a multi-institutional initiative established in 2014 with a primary goal of establishing reference CRF values in the United States across the adult life span (13). Cardiopulmonary exercise testing (CPX) laboratories from California, Connecticut, Indiana, Illinois, Louisiana, Maryland, North Carolina, Tennessee, and Texas contributed data to the registry (16). The procedures used for acquiring and managing the FRIEND registry data have been previously reported (17, 18). In brief, all the CPX laboratories contributing data to the FRIEND registry used valid and reliable calibration and testing procedures, and experienced personnel qualified to conduct exercise tests to maximal exertion (17). For the current study, participants who completed a graded exercise treadmill test with a peak respiratory exchange ratio (RER) ≥1.0, age ≥18 and ≤79 years, and BMI between 18.5 and 40 kg·m^{-2} were considered for inclusion. To select an apparently healthy group, participants with cardiovascular disease (CVD), HF, chronic lung disease, chronic kidney disease (CKD), chronic liver disease,
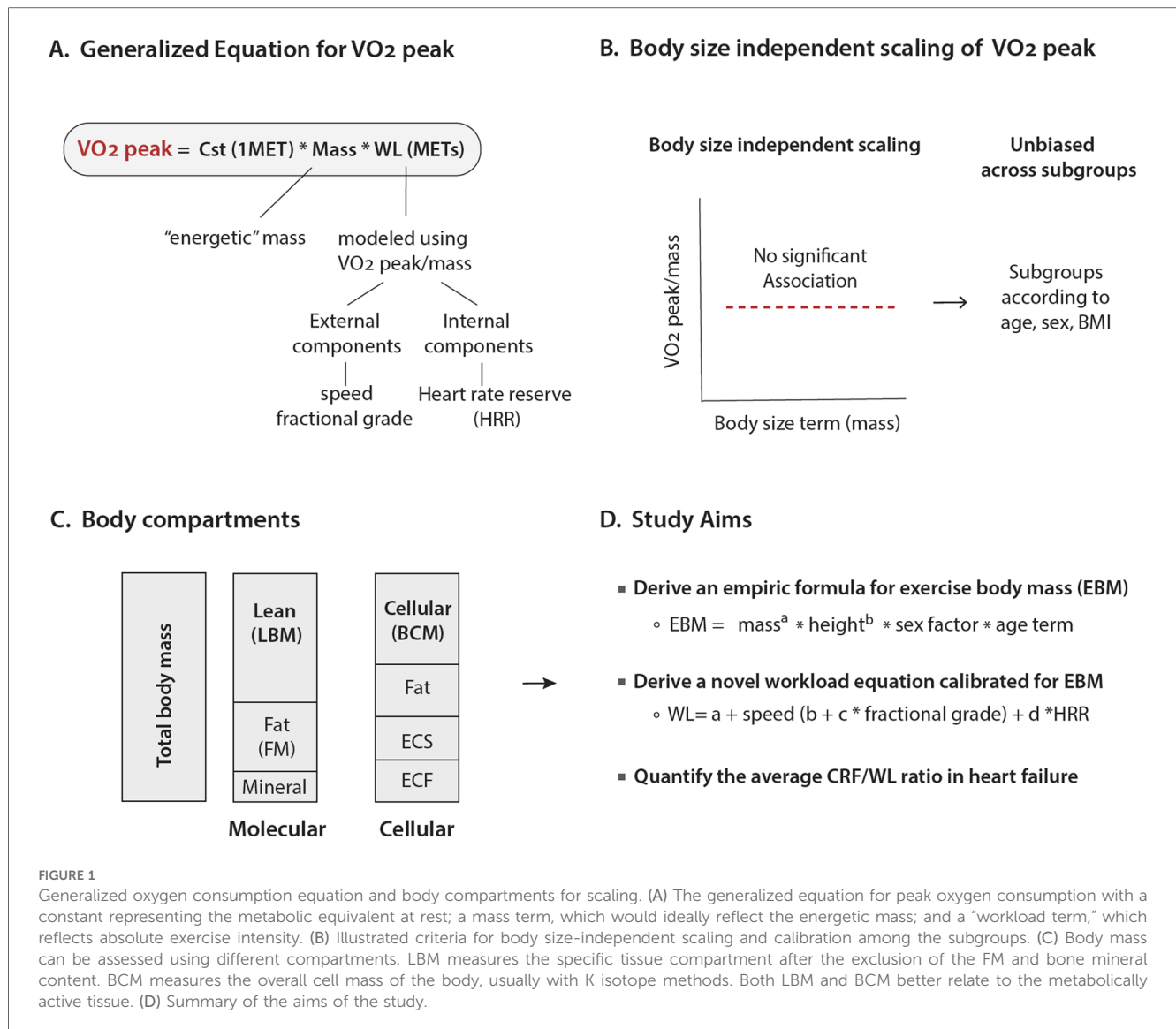
FIGURE 1
Generalized oxygen consumption equation and body compartments for scaling. **(A)** The generalized equation for peak oxygen consumption with a constant representing the metabolic equivalent at rest; a mass term, which would ideally reflect the energetic mass; and a "workload term," which reflects absolute exercise intensity. **(B)** Illustrated criteria for body size-independent scaling and calibration among the subgroups. **(C)** Body mass can be assessed using different compartments. LBM measures the specific tissue compartment after the exclusion of the FM and bone mineral content. BCM measures the overall cell mass of the body, usually with K isotope methods. Both LBM and BCM better relate to the metabolically active tissue. **(D)** Summary of the aims of the study.

endocrine disorders, neurological disorders, or diabetes mellitus were excluded.

The SET registry includes individuals who underwent CPX at Stanford University Medical Center since April 2007. Stanford's CPX personalized ramp is based on an initial assessment of the Veterans Specific Activity Questionnaire (VSAQ) (19). For the validation cohort, we selected two groups of individuals: (1) a lower-risk group without a diagnosis of CVD or HF with an age-predicted CRF >80% according to the Wasserman formula and a minute ventilation/carbon dioxide production (VE/VCO$_2$) slope <32 ($n = 198$); and (2) patients with HF or cardiomyopathy ($n = 1,339$).

The DIETFITS was a randomized controlled weight loss trial comprising 609 participants assigned to either healthy low-carbohydrate or healthy low-fat diets (15). Dual-energy x-ray absorptiometry (DXA) scans were performed on 466 participants at the baseline visit using a Hologic QDR-4500 W fan-beam scanner (Bedford, MA, USA) based on the manufacturer's guidelines. A three-compartment model quantifying LBM, fat mass, and bone mineral content was used.

2.2 Data collection

For each cohort, data collection included demographics (age, sex, and race), vital signs (blood pressure and heart rate), and anthropomorphic data (height, body mass, and BMI). Exercise variables included Sp, fGr, HRR, and VO$_2$peak (measured using open-circuit spirometry). The literature-based WL was calculated using Kokkinos et al.'s formula derived from the FRIEND registry: VO$_2$peak = Mass (kg) × [Sp (m/min) × (0.17 + fGr × 0.79) + 3.5] divided by one MET ($\sim 3.5 \, \text{ml} \, O_2 \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$) (10).

2.3 Statistical analyses

The analyses were performed using Python 3.10. Continuous variables were presented as mean and standard deviation (SD) or interquartile range as appropriate and categorical variables as percentages. Upset plots were used to present the prevalence of co-morbidities in the FRIEND registry. For regression analyses,

linear multivariable backward-weighted regression was used, in which variables with $p > 0.05$ were removed sequentially. For the log-linear modeling, the natural logarithmic of VO$_2$peak or mass and height were used and then transformed to their multiplicative expression.

2.3.1 Deriving EBM to index Vo$_2$peak and its body compartment correlates

2.3.1.1 Deriving the EBM terms using a multivariable multiplicative model

The natural logarithm (ln) of VO$_2$peak was used as the dependent variable while the independent variables consisted of ln(mass), ln (height), sex, age terms (age, age^2, and age^3), Sp, fGr, and HRR. The coefficients obtained in this model were used to express EBM as Mass$^a \times$ Height$^b \times$ Male factor $(e^c) \times \exp(d \times$ Age factor), where a, b, c, and d are coefficients. Interactions with sex were tested and retained if they significantly differed from a common allometric model. An **approximation** of the allometric EBM equation was also derived using weighted least square regression, where EBM was the dependent variable and mass, height, and age were the independent variables; such expression would be considered valid if it presents an R^2 close to unity and minimal variance.

2.3.1.2 Testing body size independence for EBM scaling of Vo$_2$peak

Body size independence for scaling occurs when no significant relationship is observed between an indexed variable and the body size metric. Since VO$_2$peak also depends on exercise intensity, we not only tested body size independence in the entire cohort but also in different WL categories and in men and women separately.

2.3.1.3 Biological plausibility of EBM assessment using body compartment correlates

The DIETFITS has an age range of 20–50 years old and allows the testing of mass and height allometric coefficients as well as sex factor in this age range. Using the DIETFITS cohort, we assessed the linear relationship between EBM and LBM$^{0.9}$ with an intercept at the origin. We chose an allometric coefficient of 0.9 for LBM based on the meta-analysis of Lolli et al. [9]. To visualize whether EBM could theoretically approximate BCM, we converted EBM to EBM$^{1.11}$ and plotted the resulting values over the age range; EBM was adjusted to a power of 1.11 (the inverse of 0.9) to provide an equivalent non-allometric expression of BCM. In this plot, we superposed the measured BCM values of the study of St-Onge et al. in healthy individuals with a BMI <35 (digitized curves from the original study) [20].

2.3.2 Deriving a novel WL equation based on EBM

2.3.2.1 Deriving a WL equation integrating external and internal factors

Using VO$_2$peak/EBM, we derived a WL equation using an additive linear backward regression, entering as independent variables the horizontal (Sp) and vertical (Sp $\times$ fGr) components (treadmill external factors) as well as the internal factor of HRR. A constant of 11 was used as the equivalent of 1 MET assuming a height for men of 1.75 m and a BMI of 25 kg$\cdot$m^{-2} (i.e., 76.6 kg); this constant was derived by assuming 1 MET was equivalent to 3.5 ml$\cdot$min$^{-1}\cdot$kg^{-1}, dividing by the EBM formula and an age of 20 years old, and rounding to the closest integer (more details in Supplementary Methods).

2.3.2.2 Calibration between CRF and WL among different subgroups according to age, sex, and BMI, and external validation

We tested the calibration of the generalized equation in subgroups according to age, sex, and BMI by evaluating the mean slope between VO$_2$peak and the VO$_2$peak generalized equation $(11 \times$ EBM $\times$ WL). Radar plots were used to present the average slope according to age (<40, 40–60, >60 years) and BMI (<25, 25–30, >30 kg$\cdot$m^{-2}) in men and women. A slope of 1 indicates a good calibration; a slope >1 represents a higher VO$_2$peak index than the WL estimates of CRF while a slope <1 represents a lower VO$_2$peak index than the WL estimates of CRF. The equation was then tested externally in the SET registry in patients without HF and within the CRF range according to the Wasserman formula.

2.3.3 Quantifying the difference in CRF when scaling to total body mass or EBM

We compared the differences between VO$_2$peak indexed to total body mass or that indexed to the EBM standard in the men and women (BMI < 25 kg$\cdot$m^{-2}), normal weight and obese (BMI < 25 vs. >30 kg$\cdot$m^{-2}), and younger and older age (<40 vs. >60 years) subgroups. The effect size was measured using relative mean ratios (%) and the Cohen's D effect size. Partition plots were used to visualize VO$_2$peak indexed to total body mass or EBM, with values standardized for one MET. For total body mass, the constant used was 3.5 ml$\cdot$min$^{-1}\cdot$kg^{-1} and for EBM, the constant was 11.

2.3.4 The CRF/WL ratio in patients with HF

In the patients with HF or cardiomyopathy from the FRIEND registry, the VO$_2$peak index may be lower than WL due to the influence of anaerobic metabolism, sarcopenia, or fluid overload. We, therefore, assessed the average slope VO$_2$peak/WL relationship in patients with HF. We also validated the average underestimation in the SET registry.

3 Results

3.1 The FRIEND cohort

A total of 5,618 individuals met the inclusion criteria for the apparently healthy group (Figure 2) and 7,240 patients with CVD were included, of which 1,007 had a diagnosis of HF. The most common co-morbidities were CVD (52.6%) followed by diabetes mellitus (11.9%) and pulmonary disease (5.2%) (Supplementary Figure S1). The characteristics of the FRIEND cohort are presented in Table 1. In the apparently
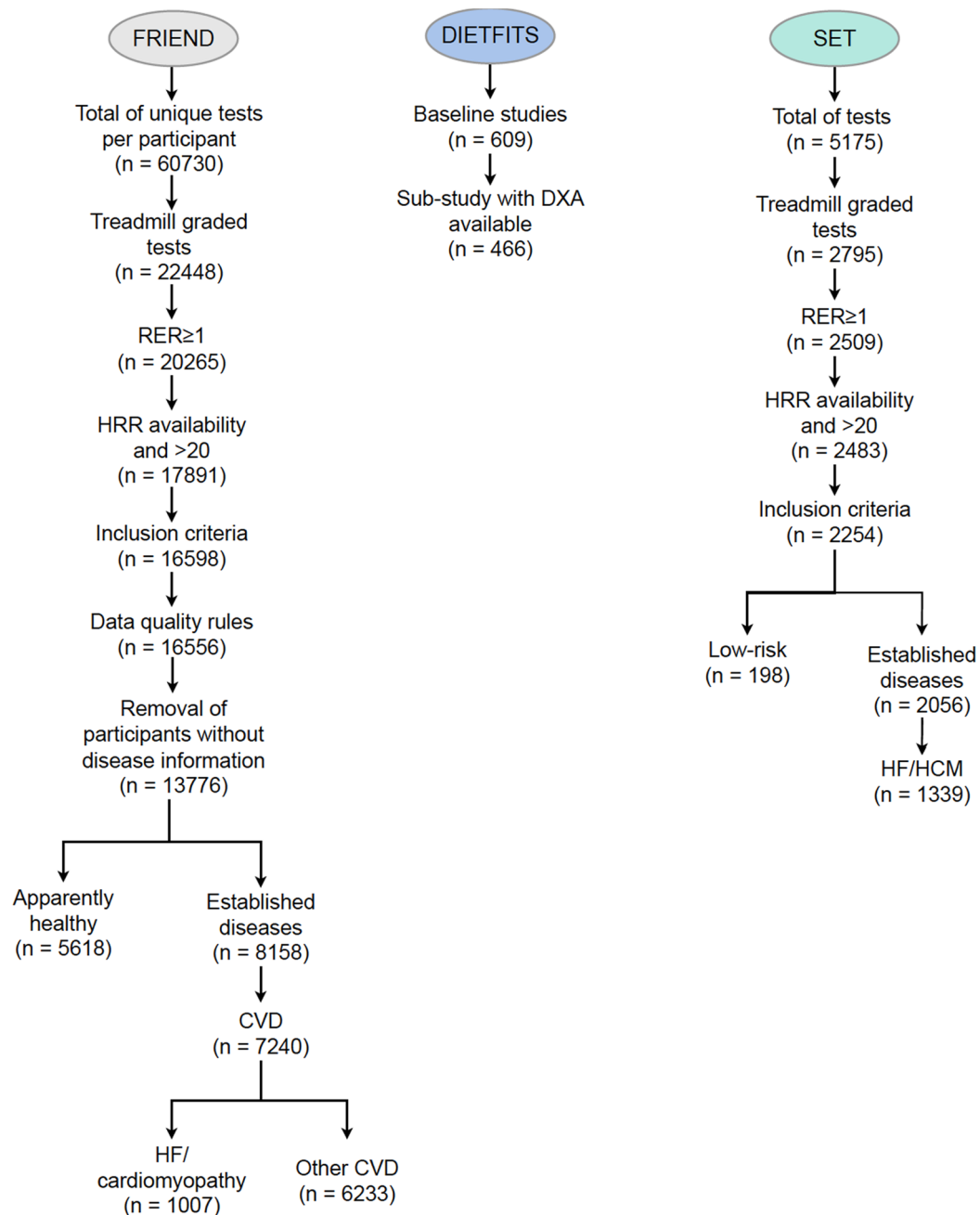
FIGURE 2
Consort diagram for the FRIEND (derivation cohort), SET (external validation cohort), and DIETFITS (contextualization cohort) datasets. CVD, cardiovascular disease; DIETFITS, Diet Intervention Examining the Factors Interacting with Treatment Success; DXA, dual-energy x-ray absorptiometry; FRIEND, Fitness Registry and the Importance of Exercise National Database; HCM, hypertrophic cardiomyopathy; SET, Stanford exercise testing.

healthy individuals, the average age was 44 ± 13 years, with 57% men and a mean BMI of $26 \pm 4 \, \text{kg} \cdot \text{m}^{-2}$. The mean RER was 1.17 ± 0.10, with a percentage age-predicted heart rate (21) of $100 \pm 7\%$ and VO_2peak of $35.6 \pm 10.6 \, \text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$. When compared to the apparently healthy group, the individuals with CVD were older, had higher BMIs, and had lower VO_2peak values.

3.2 Deriving EBM for indexing Vo₂peak and showing its clinical implications

3.2.1 Deriving the EBM equation

The multivariable model allowing the derivation of EBM terms had a coefficient of determination of 0.83, $P < 0.001$, with normally distributed residuals (Supplementary Figure S2). Considering

TABLE 1 Clinical characteristics of the FRIEND population.

Characteristic	Apparently healthy			Group with CVD		
	Total[a] (n = 5,618)	Men (n = 3,214)	Women (n = 2,404)	Total[a] (n = 7,240)	Men (n = 5,324)	Women (n = 1,916)
Demographics and anthropometrics						
Age (years)	44 ± 13	44 ± 12	44 ± 13	57 ± 13	58 ± 13	56 ± 14
Height (m)	1.73 ± 0.10	1.78 ± 0.07	1.65 ± 0.06	1.72 ± 0.09	1.75 ± 0.07	1.63 ± 0.07
Mass (kg)	78 ± 16	84 ± 14	70 ± 14	81 ± 16	84 ± 15	73 ± 15
BMI (kg·m⁻²)	26 ± 4	26 ± 4	26 ± 5	28 ± 4	28 ± 4	28 ± 5
Baseline vitals and test measurements						
Max RER	1.17 ± 0.10	1.18 ± 0.10	1.17 ± 0.10	1.15 ± 0.09	1.15 ± 0.09	1.13 ± 0.09
VO$_2$peak (ml·min⁻¹)	2,747 ± 899	3,286 ± 731	2,025 ± 515	2,098 ± 700	2,273 ± 696	1,610 ± 424
VO$_2$peak (ml·kg⁻¹·min⁻¹)	35.6 ± 10.6	40.0 ± 10.0	29.9 ± 8.3	25.9 ± 7.7	27.2 ± 7.9	22.5 ± 6.0
VO$_2$peak (METs)	10.2 ± 3.0	11.4 ± 2.9	8.5 ± 2.4	7.4 ± 2.2	7.8 ± 2.3	6.4 ± 1.7
Speed (mph, m·min⁻¹)	4.8 ± 1.5	5.4 ± 1.6	4.0 ± 1.1	3.5 ± 1.0	3.6 ± 1.0	3.2 ± 0.9
	129 ± 41	144 ± 42	108 ± 30	94 ± 27	97 ± 27	86 ± 24
Fractional grade (%)	0.11 ± 0.05	0.11 ± 0.05	0.12 ± 0.04	0.13 ± 0.03	0.14 ± 0.03	0.12 ± 0.04
Resting SBP (mmHg)	120 ± 14	124 ± 13	115 ± 14	124 ± 17	124 ± 17	125 ± 18
Resting DBP (mmHg)	77 ± 10	80 ± 9	74 ± 10	76 ± 10	76 ± 10	76 ± 10
Max HR (bpm)	176 ± 16	177 ± 15	175 ± 16	142 ± 23	141 ± 23	146 ± 23
HRR (bpm)	108 ± 19	112 ± 17	103 ± 20	71 ± 22	71 ± 22	72 ± 22
% Predicted HR (%)	100.3 ± 7.3	100.9 ± 7.0	99.4 ± 7.6	87.5 ± 12.6	87.0 ± 12.5	89.2 ± 12.6

CVD, cardiovascular disease; DBP, diastolic blood pressure; HR, heart rate; HRR, heart rate reserve; RER, respiratory exchange ratio; SBP, systolic blood pressure.
Where not specified, the value is given as mean ± standard deviation.
[a]$p < 0.001$ for all characteristics when considering a two-sided t-test between total columns for apparently healthy and for CVD.

fAge = Age/100, EBM was expressed as (Figure 3):

$$EBM = Mass (kg^2)^{0.63} \times Height (m^2)^{0.53} \times 1.16 \text{ (if male)} \times \exp(-0.39 \times fAge^2)$$

This equation can be approximated to the following additive with a correlation close to 1 and minimal residuals (Supplementary Figure S3):

$$\textbf{EBM}_{\textbf{males}} = -0.58 + 0.18 \text{ Mass (kg)} + 7.20 \text{ Height (m)} - 0.08 \text{ Age}$$

$$\textbf{EBM}_{\textbf{females}} = -0.26 + 0.16 \text{ Mass (kg)} + 5.63 \text{ Height (m)} - 0.06 \text{ Age}$$

3.2.2 Testing body size independence for EBM indexing of Vo$_2$peak

Scaling to EBM provided body size-independent scaling of VO$_2$peak with no relationship with EBM ($p < 0.01$) (Table 2). In contrast, scaling to total body mass was significantly associated with the majority of WL intervals in men and women. When compared to scaling to Mass$^{0.7}$, scaling to EBM resulted in lower absolute correlations with several intervals, particularly with 8–8.9 METs for men and with the 6–6.9 and 7–7.9 MET intervals for women.

3.2.3 Biological plausibility of the EBM formula

The characteristics of the DIETFITS cohort are summarized in Supplementary Table S1. The average age was 39 ± 7 years and the BMI ranged from 25 to 40 kg·m⁻². A strong linear correlation was observed between LBM$^{0.9}$ and EBM with $r = 0.95$, $p < 0.001$, where **LBM$^{0.9}$** = 1.61·EBM with similar slopes for men and women

(Figure 3). To determine whether the age factors in the EBM formula were biologically plausible, we superposed the EBM$^{1.11}$ and BCM values over time from the study of St-Onge et al. (20) using isotope ^{40}K (Supplementary Figure S4 for the raw data), as shown in Figure 3. This assumes that BCM would also scale to a power of 0.9 during treadmill exercise.

3.3 Deriving a new WL equation based on EBM

3.3.1 Deriving a WL equation integrating external and internal factors

Using multivariable weighted regression, we derived the generalized equation of VO$_2$peak expressed as VO$_2$peak = 11 × EBM × WL, with the WL term given by

$$WL (METs) = 2 + Sp (mph) [1.06 + 5.22 \times fGr] + 0.019 \times HRR (bpm).$$

The overall relationship between the observed and predicted VO$_2$peak had an R^2 of 0.85, $p < 0.001$. The WL formula based on EBM yielded, on average, a higher estimated WL than the one derived using Kokkinos et al.'s formula (WL$_{Kokkinos}$ = 0.88·WL$_{EBM}$) with higher values at a lower WL (Supplementary Figure S5).

3.3.2 Calibration between the observed and predicted Vo$_2$peak

Radar plots show the average calibration slope stratified according to the age, sex, and BMI subgroups for total body
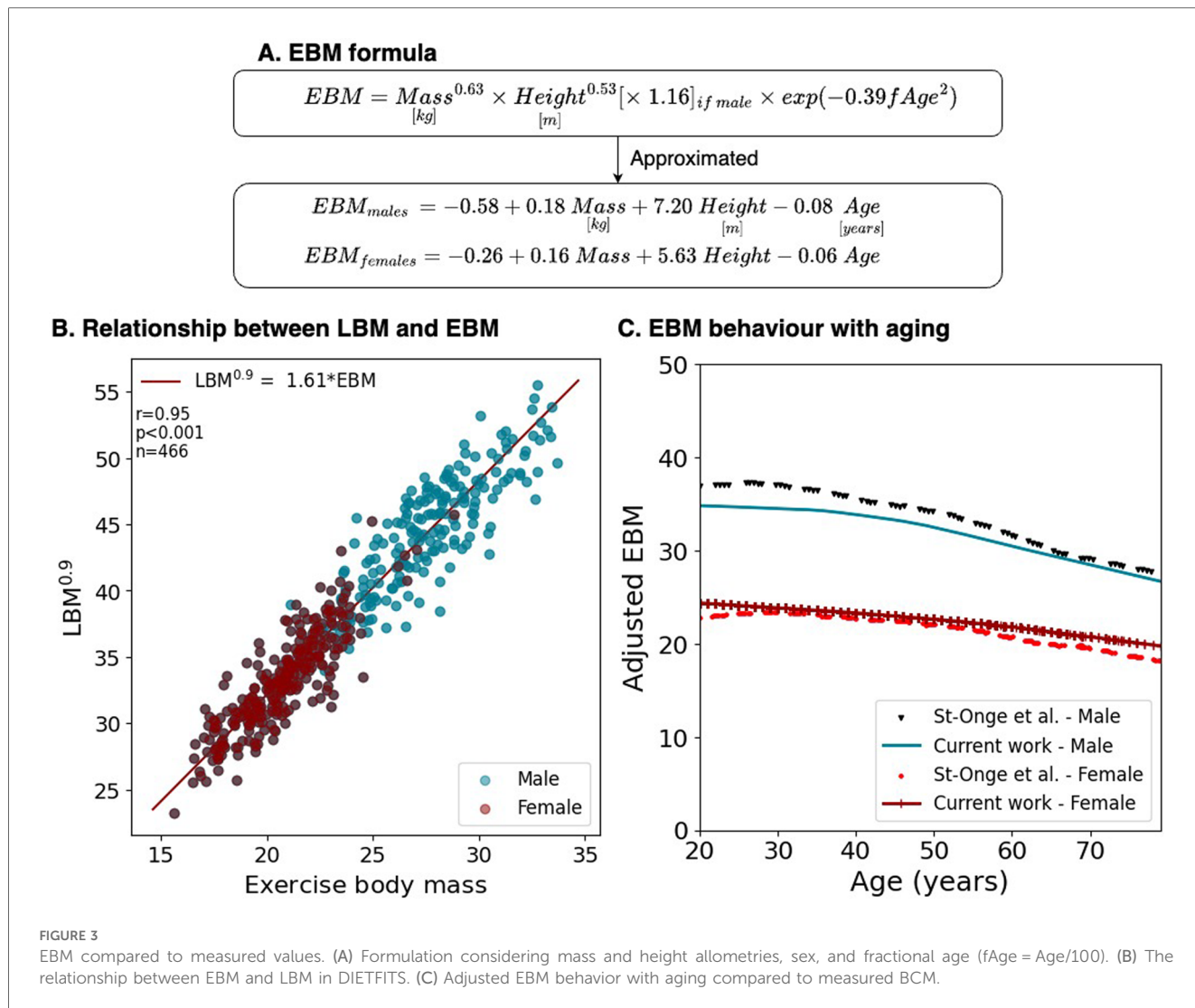
A. EBM formula

$$EBM = \underset{[kg]}{Mass}^{0.63} \times \underset{[m]}{Height}^{0.53} [\times 1.16]_{if\,male} \times exp(-0.39 fAge^2)$$

Approximated

$$EBM_{males} = -0.58 + 0.18 \underset{[kg]}{Mass} + 7.20 \underset{[m]}{Height} - 0.08 \underset{[years]}{Age}$$

$$EBM_{females} = -0.26 + 0.16 Mass + 5.63 Height - 0.06 Age$$

B. Relationship between LBM and EBM

C. EBM behaviour with aging

FIGURE 3

EBM compared to measured values. **(A)** Formulation considering mass and height allometries, sex, and fractional age (fAge = Age/100). **(B)** The relationship between EBM and LBM in DIETFITS. **(C)** Adjusted EBM behavior with aging compared to measured BCM.

TABLE 2 Body size independence of VO$_2$peak indexed to Mass compared to Mass$^{0.7}$ and EBM in terms of correlation coefficients.

	WL (METs)	*N*	VO$_2$peak/Mass	VO$_2$peak/Mass$^{0.7}$	VO$_2$peak/EBM
Men					
	<7	76	−0.263 (*p* = 0.022)	−0.002 (*p* = 0.988)	0.100 (*p* = 0.389)
	7–7.9	117	**−0.365 (*p* < 0.001)**	−0.128 (*p* = 0.170)	−0.190 (*p* = 0.040)
	8–8.9	296	**−0.168 (*p* = 0.004)**	**0.191 (*p* < 0.001)**	0.150 (*p* = 0.010)
	9–9.9	318	**−0.261 (*p* < 0.001)**	0.065 (*p* = 0.246)	0.049 (*p* = 0.387)
	10–10.9	849	**−0.244 (*p* < 0.001)**	0.069 (*p* = 0.046)	−0.000 (*p* = 1.000)
	11–11.9	418	**−0.285 (*p* < 0.001)**	0.033 (*p* = 0.507)	−0.019 (*p* = 0.701)
	12–12.9	505	**−0.336 (*p* < 0.001)**	−0.061 (*p* = 0.169)	−0.044 (*p* = 0.318)
	13–13.9	256	**−0.306 (*p* < 0.001)**	−0.002 (*p* = 0.974)	−0.076 (*p* = 0.224)
	14–14.9	180	**−0.398 (*p* < 0.001)**	−0.135 (*p* = 0.071)	−0.125 (*p* = 0.096)
	≥ 15	199	**−0.300 (*p* < 0.001)**	−0.013 (*p* = 0.860)	−0.051 (*p* = 0.470)
Women					
	<6	60	−0.292 (*p* = 0.024)	0.031 (*p* = 0.812)	0.097 (*p* = 0.462)
	6–6.9	275	**−0.230 (*p* < 0.001)**	**0.195 (*p* = 0.001)**	0.087 (*p* = 0.148)
	7–7.9	357	**−0.263 (*p* < 0.001)**	**0.154 (*p* = 0.004)**	0.101 (*p* = 0.056)
	8–8.9	601	**−0.327 (*p* < 0.001)**	0.006 (*p* = 0.891)	0.068 (*p* = 0.094)
	9–9.9	310	**−0.403 (*p* < 0.001)**	−0.140 (*p* = 0.014)	−0.012 (*p* = 0.834)
	10–10.9	496	**−0.318 (*p* < 0.001)**	−0.081 (*p* = 0.070)	0.044 (*p* = 0.329)
	11–11.9	139	**−0.284 (*p* < 0.001)**	−0.052 (*p* = 0.546)	0.032 (*p* = 0.706)
	≥ 12	166	**−0.206 (*p* = 0.008)**	0.010 (*p* = 0.903)	0.089 (*p* = 0.254)

p < 0.01 are in bold due to multiple comparison.

mass and EBM standards (Figure 4). These demonstrate better calibration when using the EBM-based generalized equations. Underestimation of CRF using VO₂peak/total body mass for a given WL was worst ($p < 0.001$) in women, in older individuals, and in individuals with obesity.
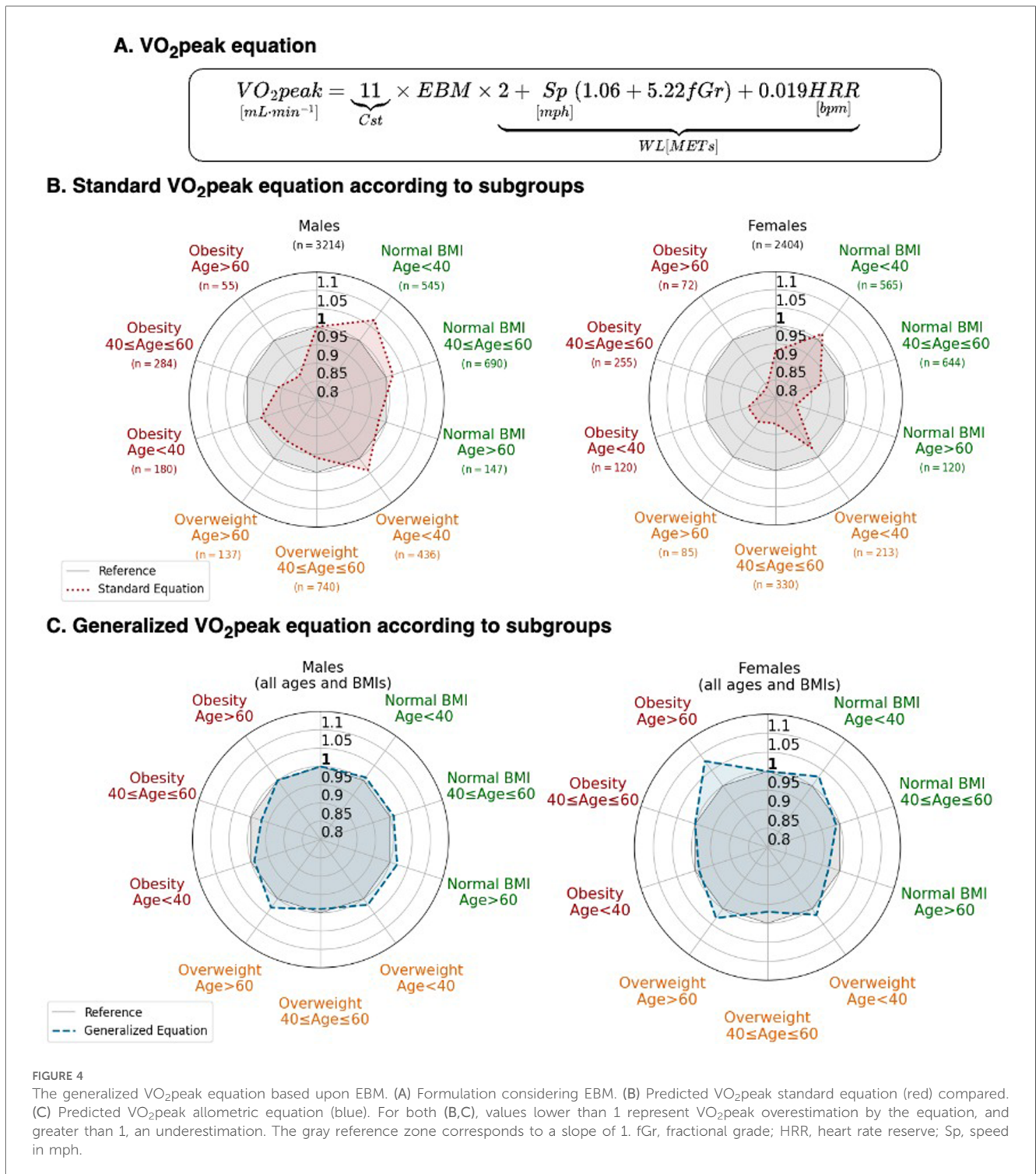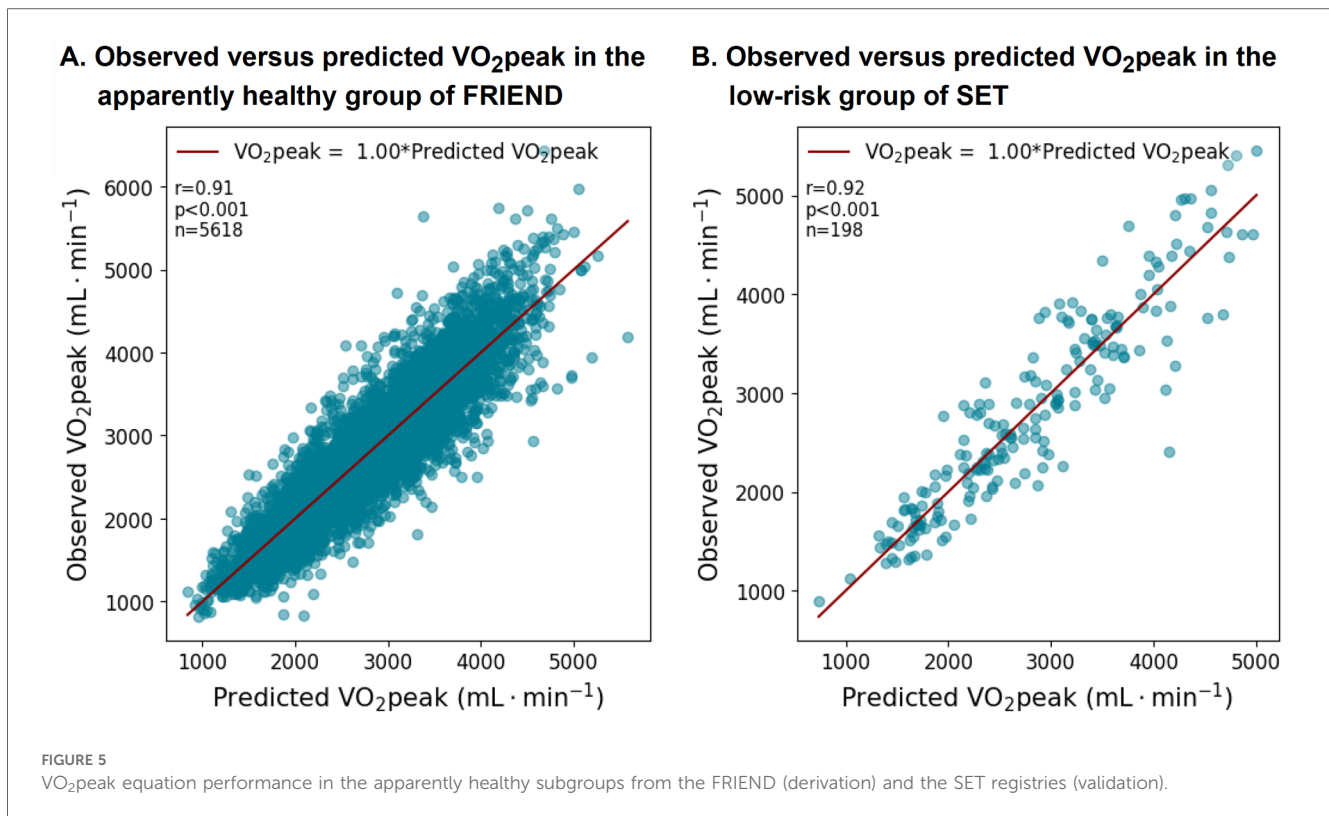
3.3.3 Validation in the SET registry

The characteristics of the patients in the SET registry are presented in Supplementary Table S3. The equation was well-validated in the SET registry with an average slope of 1.0

($r = 0.92$, $p < 0.001$) (Figures 5A,B). The subgroup analyses by sex, age, and BMI also demonstrate a well-calibrated equation (Supplementary Table S4).

3.4 Quantifying the difference in CRF when using the total body mass or EBM standard

The partition plots of the VO₂peak index for total mass or EBM (presented as METs) are presented in Figure 6. Smaller

FIGURE 4
The generalized VO₂peak equation based upon EBM. **(A)** Formulation considering EBM. **(B)** Predicted VO₂peak standard equation (red) compared. **(C)** Predicted VO₂peak allometric equation (blue). For both **(B,C)**, values lower than 1 represent VO₂peak overestimation by the equation, and greater than 1, an underestimation. The gray reference zone corresponds to a slope of 1. fGr, fractional grade; HRR, heart rate reserve; Sp, speed in mph.

FIGURE 5
VO$_2$peak equation performance in the apparently healthy subgroups from the FRIEND (derivation) and the SET registries (validation).

differences between men and women, older and younger, and obese and normal individuals were observed when using the EBM standard compared to the total mass standard (Figure 6, Supplementary Figure S6).

3.5 The CRF/WL relationship in patients with HF

In both the FRIEND and SET registries, patients with HF had on average a lower slope compared to apparently healthy individuals as presented in Figure 7 and Supplementary Table S2. Supplementary Table S5 summarizes the main equations related to this work.

4 Discussion

In this study, we developed a body size-independent scaling factor for VO$_2$peak. EBM provided body size-independent scaling with robust calibration across different ages, sexes, and BMI categories. This has important implications for accurate reporting of CRF in the general population. In addition, our novel WL equation incorporates HRR, which can provide better estimates of energy expenditure during treadmill exercise testing. Finally, we found that in patients with HF, VO$_2$peak was often lower than the WL-predicted values, which may be due to an increased reliance on anaerobic metabolism, the presence of sarcopenia, or fluid overload.

Scaling plays a central role in comparative biology as it reduces differences associated with body size and composition (6). This is particularly relevant when comparing men and women, younger and older individuals, or individuals with normal weight and obesity. Several studies have shown that VO$_2$peak scales to body size according to a log-linear relationship, providing the basis for allometric scaling (6, 8, 9, 22). In 1981, Taylor et al. conducted a landmark study to determine whether VO$_2$max is scaled proportionally to mass (8). Their study included both domestic and wild animals spanning several orders of magnitude of mass (7.2 g–263 kg) and testing for VO$_2$max following training sessions in all the animals. They found that VO$_2$max scaled to 0.79 (0.75–0.83) in wild animals and to 0.77 (0.68–0.85) in domestic animals. In a recent meta-analysis of 36 studies involving 6,514 participants, Lolli et al. found that the pooled allometric exponent for indexed aerobic capacity was 0.70 (95% Cl of 0.64–0.76) for whole body mass and 0.90 (95% Cl of 0.83–0.96) for fat-free mass (9). A similar mass coefficient of 0.71 was also found in groups of prepubertal, circumpubertal, and adult subjects after accounting for differences in height (23).

Original to our study, we validated, in a large cohort, a body size-independent scaling metric for VO$_2$peak. EBM was closely related to LBM and followed the age-associated decrease in BCM described by in previous works (20, 24). Scaling to EBM attenuated differences in VO$_2$peak associated with sex, age, or obesity. Scaling to Mass$^{0.7}$ also yielded overall good body size-independent scaling but had the disadvantage of not considering stature or integrating sex differences in body composition. In addition, when scaling to EBM, the sex gap in younger individuals of approximately

A. VO₂peak partitions according to mass standard ratios in males
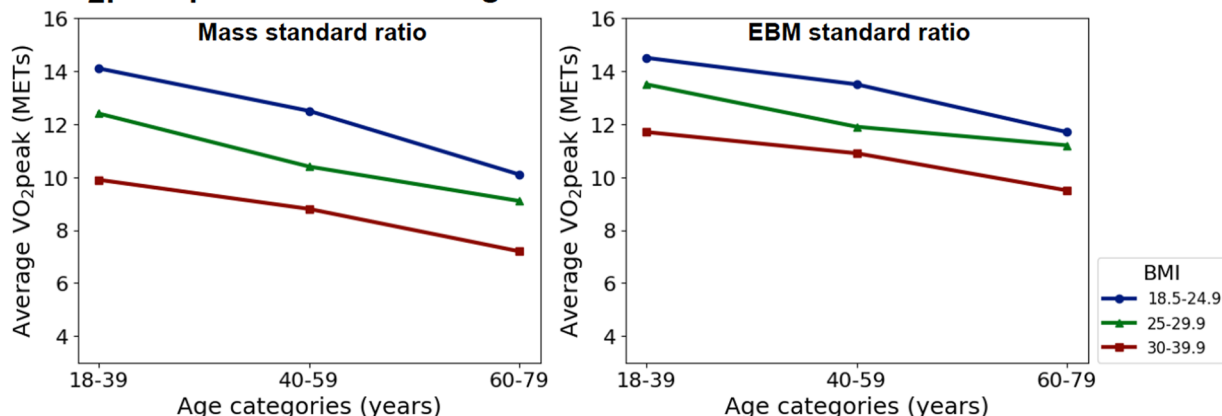
B. VO₂peak partitions according to mass standard ratios in females
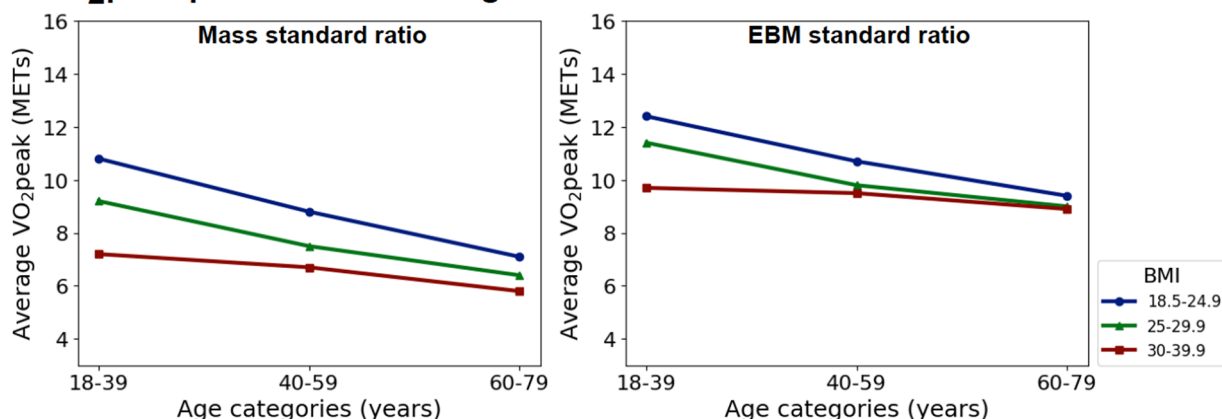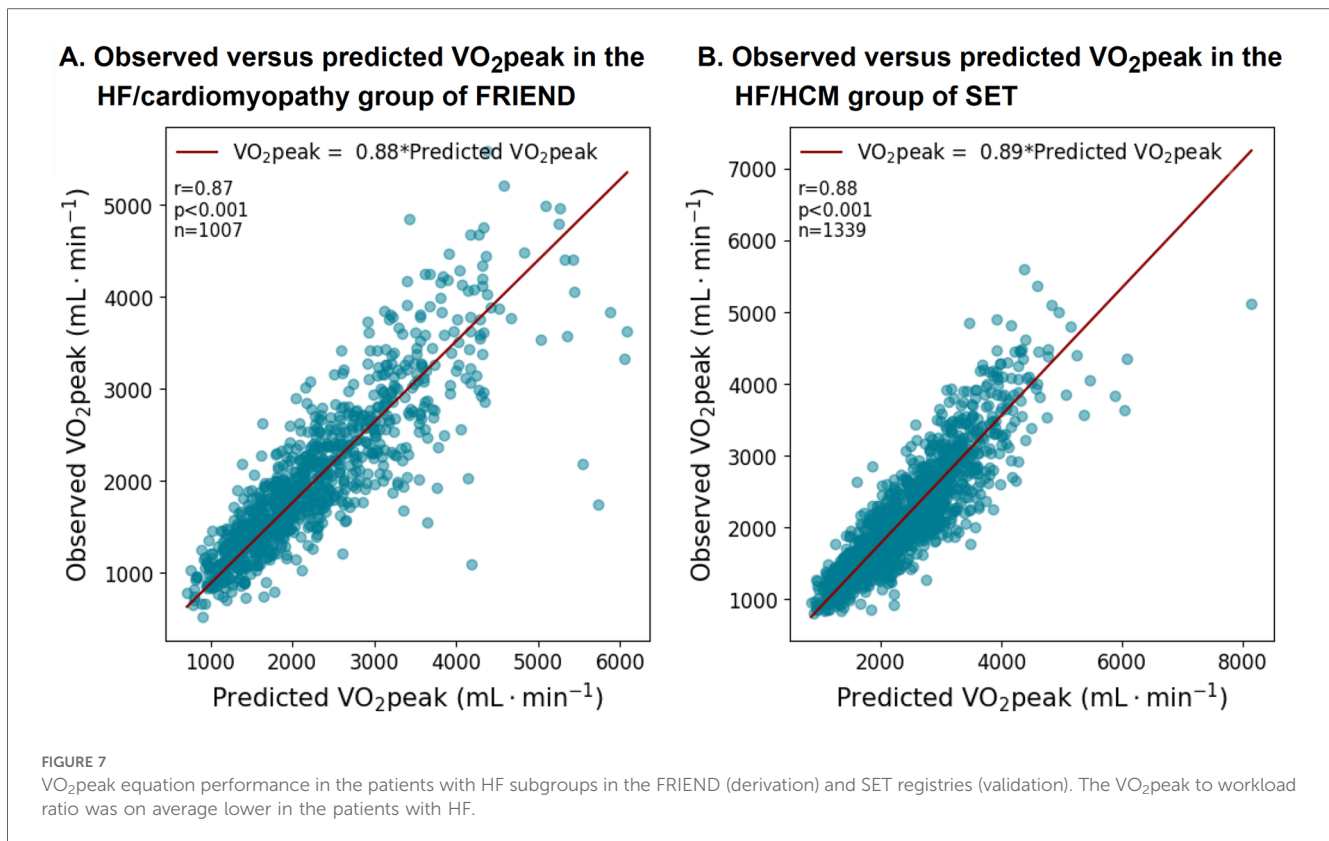

C. Implications of mass standard ratios

	VO₂peak ratio		Cohen's D	
	Mass standard	EBM standard	Mass standard	EBM standard
Male vs Female*	76.8	85.9	-1.35	-0.80
BMI <25 vs >30**				
Male	70.0	80.4	-1.92	-1.24
Female	66.2	78.1	-1.86	-1.19
Age<40 vs >60***				
Male	71.1	80.7	-1.74	-1.14
Female	65.1	74.8	-2.04	-1.48

FIGURE 6

Indexing VO₂peak to EBM instead of mass reduces age, sex, and BMI differences. **(A)** Partition analysis for men. **(B)** Partition analysis for women. **(C)** Comparison of mean VO₂peak ratios and Cohen's D effect size (*BMI < 25 and age <40 years; **age <40 years; ***BMI < 25; comparison of all subgroups in Supplementary Figure S5).

10%–15% was consistent with performance difference reported in athletes (25). Using the simplified additive equation for EBM will likely facilitate clinical adoption and have a similar expression as the national health and nutrition examination survey (NHANES) equations for LBM (26). The second original contribution of our study was the development of a WL equation that incorporates HRR; this will provide more personalized estimates of aerobic capacity (11). Compared to the Kokkinos equation (10), the new WL equation provides higher estimates at lower Sp and fGr,

reflecting the reduced efficiency of slower locomotion (27–29). This explains why the intercept of the equation was likely higher than unity with a higher early increment.

The goal underpinning the generalized equation is to improve the consistency of CRF reporting across the general population. For instance, at the same Sp, fGr, or HRR, apparently healthy individuals will, on average, have similar aerobic capacity regardless of age, sex, or BMI. The EBM equation is an average scaling tool and not a substitute for direct measurements of LBM or BCM.

FIGURE 7
VO$_2$peak equation performance in the patients with HF subgroups in the FRIEND (derivation) and SET registries (validation). The VO$_2$peak to workload ratio was on average lower in the patients with HF.

Specific considerations are needed for trained athletes and frail individuals with sarcopenia: athletes generally have a higher proportion of BCM, while frail individuals with sarcopenia have a lower proportion (30, 31). Without these adjustments, CRF may be overestimated in athletes and underestimated in individuals with sarcopenia. The lower calibration slope observed in patients with HF from both the FRIEND and SET registries could be due to variations in body composition (such as sarcopenia or fluid overload) or increased reliance on anaerobic metabolism. Tanabe et al. (32) previously reported a discrepancy in measured and estimated VO$_2$peak (or WL) in patients with HF using an ergocycle (32). Although we were able to confirm the 20% higher estimates, these were reduced to 10% differences after adjusting for HRR.

While our study is the largest to empirically validate a scaling metric for VO$_2$peak, several limitations need to be highlighted. First, the EBM formula is based on a United States population of mainly white participants and further adjustment for different races and ethnicity will need to be studied. Second, we did not account for duration of exercise or holding the handrails of the treadmill in our WL estimates. Third, other factors such as peak blood pressure, environment, and diurnal rhythm can also influence VO$_2$peak and will require further study. We were also not able to compare EBM to measure LBM in the FRIEND cohort itself due to a lack of DXA, requiring another cohort (DIETFITS) for this task. In addition, the developed equations are specific for treadmill CPX. Finally, future refinements in an EBM formula should consider correction factors for athletic status and sarcopenia/frailty.

In conclusion, the EBM formula enables body size-independent scaling of VO$_2$peak and more consistent and equitable reporting of

CRF. Incorporating HRR in the WL equations also has the added advantage of a more personalized estimate of aerobic capacity during exercise.

Data availability statement

The datasets presented in this article are not readily available because of ethical restrictions. Requests to access the datasets should be directed to Francois Haddad, fhaddad@stanford.edu.

Ethics statement

The studies involving humans were approved by FRIEND: Ball State University; SET and DIETFITS: Stanford University. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because each contributing site obtained local research ethics board approval (or waiver of consent) prior to submitting data.

Author contributions

ES: Conceptualization, Formal Analysis, Visualization, Writing – original draft. NC: Writing – review & editing. BC: Writing – review & editing. TK: Writing – review & editing. CG: Data

curation, Writing – review & editing. RA: Data curation, Writing – review & editing. LK: Data curation, Writing – review & editing. MH: Data curation, Writing – review & editing. EA: Writing – review & editing. JC: Conceptualization, Data curation, Writing – review & editing. JM: Conceptualization, Data curation, Writing – review & editing. FH: Conceptualization, Formal Analysis, Visualization, Writing – original draft.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2024.1393363/full#supplementary-material

References

1. Ross R, Blair SN, Arena R, Church TS, Despres JP, Franklin BA, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation*. (2016) 134:e653–99. doi: 10.1161/CIR.0000000000000461

2. Guazzi M, Adams V, Conraads V, Halle M, Mezzani A, Vanhees L, et al. Clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. *Circulation*. (2012) 126:2261–74. doi: 10.1161/CIR.0b013e31826fb946

3. Kaminsky LA, Arena R, Myers J, Peterman JE, Bonikowske AR, Harber MP, et al. Updated reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing. *Mayo Clin Proc*. (2022) 97:285–93. doi: 10.1016/j.mayocp.2021.08.020

4. Krachler B, Savonen K, Komulainen P, Hassinen M, Lakka TA, Rauramaa R. Cardiopulmonary fitness is a function of lean mass, not total body weight: the DR's EXTRA study. *Eur J Prev Cardiol*. (2015) 22:1171–9. doi: 10.1177/2047487314557962

5. Kohler A, King R, Bahls M, Gross S, Steveling A, Gartner S, et al. Cardiopulmonary fitness is strongly associated with body cell mass and fat-free mass. *Scand J Med Sci Sports*. (2018) 28:1628–35. doi: 10.1111/sms.13057

6. Nevill AM, Holder RL. Scaling, normalizing, and per ratio standards: an allometric modeling approach. *J Appl Physiol (1985)*. (1995) 79:1027–31. doi: 10.1152/jappl.1995.79.3.1027

7. Wang Z, Heymsfield SB, Ying Z, Pierson RN Jr, Gallagher D, Gidwani S. A cellular level approach to predicting resting energy expenditure: evaluation of applicability in adolescents. *Am J Hum Biol*. (2010) 22:476–83. doi: 10.1002/ajhb.21020

8. Taylor CR, Maloiy GM, Weibel ER, Langman VA, Kamau JM, Seeherman HJ, et al. Design of the mammalian respiratory system. III scaling maximum aerobic capacity to body mass: wild and domestic mammals. *Respir Physiol*. (1981) 44:25–37. doi: 10.1016/0034-5687(81)90075-X

9. Lolli L, Batterham AM, Weston KL, Atkinson G. Size exponents for scaling maximal oxygen uptake in over 6500 humans: a systematic review and meta-analysis. *Sports Med*. (2017) 47:1405–19. doi: 10.1007/s40279-016-0655-1

10. Kokkinos P, Kaminsky LA, Arena R, Zhang J, Myers J. New generalized equation for predicting maximal oxygen uptake (from the fitness registry and the importance of exercise national database). *Am J Cardiol*. (2017) 120:688–92. doi: 10.1016/j.amjcard.2017.05.037

11. Swain DP, Leutholtz BC, King ME, Haas LA, Branch JD. Relationship between% heart rate reserve and% VO2 reserve in treadmill exercise. *Med Sci Sports Exerc*. (1998) 30:318–21. doi: 10.1097/00005768-199802000-00022

12. Lena A, Anker MS, Springer J. Muscle wasting and sarcopenia in heart failure—the current state of science. *Int J Mol Sci*. (2020) 21:6549. doi: 10.3390/ijms21186549

13. Kaminsky LA, Myers J, Arena R. Determining cardiorespiratory fitness with precision: compendium of findings from the FRIEND registry. *Prog Cardiovasc Dis*. (2019) 62:76–82. doi: 10.1016/j.pcad.2018.10.003

14. Moneghetti KJ, Giraldeau G, Wheeler MT, Kobayashi Y, Vrtovec B, Boulate D, et al. Incremental value of right heart metrics and exercise performance to well-validated risk scores in dilated cardiomyopathy. *Eur Heart J Cardiovasc Imaging*. (2018) 19:916–25. doi: 10.1093/ehjci/jex187

15. Stanton MV, Robinson JL, Kirkpatrick SM, Farzinkhou S, Avery EC, Rigdon J, et al. DIETFITS study (diet intervention examining the factors interacting with treatment success)—study design and methods. *Contemp Clin Trials*. (2017) 53:151–61. doi: 10.1016/j.cct.2016.12.021

16. Kaminsky LA, Arena R, Myers J. Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing. *Mayo Clin Proc*. (2015) 90:1515–23. doi: 10.1016/j.mayocp.2015.07.026

17. Pelliccia A, Sharma S, Gati S, Back M, Borjesson M, Caselli S, et al. 2020 ESC guidelines on sports cardiology and exercise in patients with cardiovascular disease. *Eur Heart J*. (2021) 42:17–96. doi: 10.1093/eurheartj/ehaa605

18. Kaminsky LA, Arena R, Beckie TM, Brubaker PH, Church TS, Forman DE, et al. The importance of cardiorespiratory fitness in the United States: the need for a national registry. *Circulation*. (2013) 127:652–62. doi: 10.1161/CIR.0b013e31827ee100

19. Maeder M, Wolber T, Atefy R, Gadza M, Ammann P, Myers J, et al. A nomogram to select the optimal treadmill ramp protocol in subjects with high exercise capacity. *J Cardiopulm Rehabil*. (2006) 26:16–23. doi: 10.1097/00008483-200601000-00004

20. St-Onge M-P, Wang J, Shen W, Wang ZM, Allison DB, Heshka S, et al. Dual-energy x-ray absorptiometry-measured lean soft tissue mass: differing relation to body cell mass across the adult life span. *J Gerontol A Biol Sci Med Sci.*. (2004) 59(8): B796–800. doi: 10.1093/gerona/59.8.b796

21. Arena R, Myers J, Kaminsky LA. Revisiting age-predicted maximal heart rate: can it be used as a valid measure of effort? *Am Heart J*. (2016) 173:49–56. doi: 10.1016/j.ahj.2015.12.006

22. Nevill AM, Ramsbottom R, Williams C. Scaling physiological measurements for individuals of different body size. *Eur J Appl Physiol Occup Physiol.* (1992) 65:110–7. doi: 10.1007/BF00705066

23. Welsman JR, Armstrong N, Nevill AM, Winter EM, Kirby BJ. Scaling peak VO2 for differences in body size. *Med Sci Sports Exerc.* (1996) 28:259–65. doi: 10.1097/00005768-199602000-00016

24. Fleg JL, Morrell CH, Bos AG, Brant LJ, Talbot LA, Wright JG, et al. Accelerated longitudinal decline of aerobic capacity in healthy older adults. *Circulation.* (2005) 112:674–82. doi: 10.1161/CIRCULATIONAHA.105.545459

25. Thibault V, Guillaume M, Berthelot G, Helou NE, Schaal K, Quinquis L, et al. Women and men in sport performance: the gender gap has not evolved since 1983. *J Sports Sci Med.* (2010) 9:214–23.

26. Lee DH, Keum N, Hu FB, Orav EJ, Rimm EB, Sun Q, et al. Development and validation of anthropometric prediction equations for lean body mass, fat mass and percent fat in adults using the National Health and Nutrition Examination Survey (NHANES) 1999–2006. *Br J Nutr.* (2017) 118:858–66. doi: 10.1017/S0007114517002665

27. Taylor CR, Heglund NC. Energetics and mechanics of terrestrial locomotion. *Annu Rev Physiol.* (1982) 44:97–107. doi: 10.1146/annurev.ph.44.030182.000525

28. Heglund NC, Taylor CR. Speed, stride frequency and energy cost per stride: how do they change with body size and gait? *J Exp Biol.* (1988) 138:301–18. doi: 10.1242/jeb.138.1.301

29. Askew GN, Formenti F, Minetti AE. Limitations imposed by wearing armour on medieval soldiers' locomotor performance. *Proc Biol Sci.* (2012) 279:640–4. doi: 10.1098/rspb.2011.0816

30. Curcio F, Testa G, Liguori I, Papillo M, Flocco V, Panicara V, et al. Sarcopenia and heart failure. *Nutrients.* (2020) 12:211. doi: 10.3390/nu12010211

31. Brioche T, Pagano AF, Py G, Chopard A. Muscle wasting and aging: experimental models, fatty infiltrations, and prevention. *Mol Aspects Med.* (2016) 50:56–87. doi: 10.1016/j.mam.2016.04.006

32. Tanabe Y, Nakagawa I, Ito E, Suzuki K. Hemodynamic basis of the reduced oxygen uptake relative to work rate during incremental exercise in patients with chronic heart failure. *Int J Cardiol.* (2002) 83:57–62. doi: 10.1016/S0167-5273(02)00013-X