©Journal of Sports Science and Medicine (2025) **24**, 684-695
http://www.jssm.org DOI: https://doi.org/10.52082/jssm.2025.684

Research article

Quantifying Running Economy in Amateur Runners: Evaluating VO_2 and Energy Cost with Model-based Normalization

Jay Lee [1,2], Xiuli Zhang [2]✉ and Zhaowei Kong [1]✉
[1] Faculty of Education, University of Macau, Macao, China
[2] School of Physical Education & Sports Science, South China Normal University, Guangzhou, China

Abstract

Submaximal oxygen uptake (VO_{2sub}) scaled by ratio is commonly used to evaluate running economy (RE) to reflect metabolic consumption at a given submaximal-intensity velocity. However, this method is questionable due to its neglect of substrate-related issues and the inherent mathematical discrepancies in ratio scaling. This study aimed to investigate the validity of ratio-scaled VO_{2sub} as a measure of RE by comparing it with allometric-scaled energy cost (E_c, kcal/kg^b/min). Sixty-nine recreationally active college students underwent VO_{2max} tests and discontinuous submaximal running assessments at three %VO_{2max} intensities. A 1000-meter test assessed running performance. One-way repeated-measures ANOVA compared changes in VO_{2sub} or E_c with increasing running intensities. Regression analysis explored methods for metabolic data standardization. Pearson correlation coefficient evaluated the effectiveness of standardization and the correlations between sports performance and RE scaled by different measures. Magnitude-based inferences were used to assess sex differences and probabilities of RE at each running intensity. Both VO_{2sub} and E_c significantly increased with increasing intensities, suggesting that VO_{2sub} is a valid quantification of RE. Allometric scaling is more suitable than ratio scaling for removing the influence of body weight on both E_c and VO_{2sub}, with females showing better RE. Allometric-scaled E_c was sensitive in detecting correlations with performance, strongest at 65% VO_{2max}. While VO_{2sub} is a valid quantification of RE, allometric scaling, rather than ratio scaling, should be used to normalize the RE quantification before performing reliable interindividual comparisons. The 2/3 law can be considered as the exponent b value for body weight. Additionally, 65%VO_{2max} intensity is recommended as the submaximal testing intensity in the RE test. Nonetheless, more studies with diverse samples are needed to confirm the validity.

Key words: Running economy, ratio scaling, oxygen consumption, energy cost, allometric scaling.

Introduction

Running economy (RE), defined as the metabolic consumption at a given submaximal-intensity velocity, is a more robust physiological indicator for distinguishing homogenous populations and predicting endurance performance than maximal oxygen uptake (VO_{2max}) alone (Midgley et al., 2007; Shaw et al., 2014) .When combined with VO_{2max} and lactate threshold or fractional utilization at the speed corresponding to lactate threshold (sLT), RE can explain up to 90% of the variance in distance events (Abad et al., 2016; Midgley et al., 2007). Therefore, ensuring the validity of RE measurements is crucial for evaluating human locomotion and performance.

Previous studies have examined inter-individual differences in specific anthropological, mechanical, and physiological aspects to identify factors contributing to an economical runner (Midgley et al., 2007; Segers et al., 2008; Shaw et al., 2014). A commonly used indirect measure of RE is submaximal oxygen uptake (VO_{2sub}), expressed in milliliters per kilogram per minute (ml/kg/min) during steady-state exercise. This metric is based on the assumption that oxygen consumption represents adenosine triphosphate turnover during submaximal exercise, thereby serving as a proxy for overall energy expenditure (Shaw et al., 2014). However, the validity of VO_{2sub} as a reliable quantification of RE has been questioned. Several studies have reported inconsistencies between changes in VO_{2sub} and actual energy cost (E_c) as running velocity increases (Fletcher et al., 2009; Shaw et al., 2014). One plausible explanation for this discrepancy is the metabolic shift toward greater carbohydrate utilization at higher exercise intensities. As intensity rises from approximately 55% to 75% of VO_{2max}, there is increased recruitment of type II muscle fibers, thereby accompanied by a reduction in fat oxidation (from ~60% to ~35%) and a corresponding increase in carbohydrate oxidation (from ~35% to ~60%) (Jeukendrup and Wallis, 2005; Rapoport, 2010). This shift has important implications for interpreting oxygen consumption. Although carbohydrate oxidation requires less oxygen per gram (0.81 L) compared to fat oxidation (2.01 L), it also yields less energy (4.07 kcal vs. 9.75 kcal per gram) (Jeukendrup and Wallis, 2005). Consequently, during moderate to high-intensity exercise, total oxygen consumption may remain relatively stable despite increased energy output, potentially leading to misleading conclusions about RE when quantified by VO_{2sub}. This phenomenon may partly explain the observed divergence between VO_{2sub} and E_c in previous studies (di Prampero et al., 2009; Fletcher et al., 2009; Jeukendrup and Wallis, 2005). Another factor that complicates the use of VO_{2sub} as a quantification of RE is the variability in substrate utilization among individuals. For instance, if athlete A is more efficient at metabolizing fat than athlete B, athlete A may exhibit higher oxygen consumption due to the greater oxygen demand of fat oxidation (Hawley and Leckey, 2015). This could lead to the erroneous conclusion that athlete B has better RE than athlete A. Therefore, if oxygen consumption does not reflect the metabolic consumption well, ridiculous findings may occur, and the validity of VO_{2sub} as a measure of RE could be compromised. Directly calculating energy consumption seems more accurate and reasonable, as it avoids issues related to substrate utilization and aligns with the definition

Received: 06 June 2025 / Accepted: 01 August 2025 / Published (online): 01 September 2025

of RE. However, limited studies have employed E_c as a primary measure of RE (Blagrove et al., 2017; di Prampero et al., 2009; Fletcher et al., 2009), and only a relatively small cohort of highly trained runners in these studies may weaken the power of relevant findings or not be suitable to be applied to other populations (e.g., non-athletes).

Moreover, accurate interindividual comparisons of RE require proper normalization of physiological variables, independent of potential confounding factors. Studies on the quantitative relationship between VO_{2max} and body weight (Armstrong and Welsman, 2019; Segers et al., 2008; Shaw et al., 2014) have observed a non-linear relationship and a significant negative correlation between VO_{2max} / body weight, namely the normalized VO_{2max}, and body weight. These findings challenge the validity of ratio scaling in removing the influence of body weight. Instead, allometric scaling ($y = ax^b$) seems more accurate to normalize physiological variables, where y represents VO_{2max} and x represents body weight (Shingleton and Frankino, 2018). If ratio scaling is valid, the exponent b would equal 1, indicating a linear relationship. However, empirical studies on VO_{2max} rejected the assumed linear relationship ($b \neq 1$), with calculated b values clustering around either 2/3 or 3/4. These values correspond to established biological scaling laws (i.e., the 2/3 or 3/4 power scaling laws), leading scholars to suggest that an exponent b for body weight of 2/3 or 3/4 is acceptable when scaling VO_{2max} (Chamari et al., 2005; Helgerud, 1994; Lolli et al., 2017). If a similar curvilinear relationship exists in scaling RE, findings related to RE expressed by ratio scaling should be viewed skeptically. For instance, results based on ratio scaling might lead to underestimation of an individual's aerobic capacity while overestimating its RE (Chamari et al., 2005; Davies et al., 1997; Lee et al., 2024; Lee and Zhang, 2021). Moreover, inconsistent results in RE differences between male and female athletes have been reported (Ariëns et al., 1997; Armstrong et al., 1999; Davies et al., 1997; Helgerud, 1994; Janz et al., 1998), which can be partly explained by the statistical adjustment method that is incapable of removing the influence of body weight when conducting inter-individual comparison. Despite the importance of this issue, relatively few studies(Armstrong et al., 1999; Fletcher et al., 2009; Lee et al., 2024; Rogers et al., 1995; Shaw et al., 2014; Welsman and Armstrong, 2000) have explored the relationship between body weight and RE measures. Moreover, it is inappropriate to apply the non-linear scaling relationship observed in VO_{2max} directly to RE (Chamari et al., 2005), as the exponent b may vary depending on the metabolic state—shifting from resting to maximal conditions(Lee et al., 2024). Furthermore, these studies have focused on elite athletes, limiting the generalizability of their findings. The exploration in moderately-trained runners with diverse athletic abilities may reflect more general characteristics of RE, specifically the universality of allometric-scaled E_c.

Given that the relationship between body weight and measures of RE, as well as the application of these measures in sports performance, has yet to be examined in a larger cohort of recreationally active runners, the appropriate scaling of E_c relative to body weight in this population remains to be elucidated. Therefore, this study aimed to assess the validity of VO_{2sub} expressed by ratio scaling as a measure of RE in comparison with E_c. We hypothesized that allometric-scaled E_c would be a more appropriate RE measurement than the traditional ratio-scaled VO_{2sub}.

Methods

Experimental design

This study comprised different measurements assigned across four non-consecutive testing days spaced 1-7 days apart. On Day 1, anthropometric data (stature, body weight, gender, and age) and VO_{2max} were assessed. On Day 2, treadmill velocities eliciting 55%, 65%, and 75%VO_{2max} for the RE test were verified. On Day 3, the formal RE assessment was conducted at these set speeds. On Day 4, a 1000-meter running field test was performed.

All tests were conducted in the afternoon, with the laboratory temperature maintained at 22-25 °C. Participants were required to avoid high-intensity exercises 24 hours before the tests to prevent fatigue. In addition, participants were instructed to fast for 2 hours before the tests and to wear the same sports shoes for all tests. Consumption of caffeine and tea beverages was prohibited before the tests.

Participants

Ethical approval of this study was obtained from the Human Research Ethics Committee of South China Normal University (SCNU-SPT-2022-040). The research was conducted in accordance with the principles of the Declaration of Helsinki and the local statutory requirements. Participants were recruited through campus advertisements according to the following criteria: (1) aged 18-30 years; (2) regular running habits, engaging in at least moderate-intensity running exercise three times a week for at least one hour per session; (3) free from any sports-related injuries or respiratory problems; (4) non-smoker and not habitual alcohol drinkers. Exclusion criteria included: experiencing any sports injury within the past 3 months, being a student-athlete, having physiological or psychological defects or diseases, regularly participating in sports/exercise activities other than running, or lacking exercise habits. Ultimately, sixty-nine recreationally active college students (34 males, age, 23.9 ± 4.7 years, stature, 174.4 ± 5.4 cm, body weight, 66.5 ± 6.6 kg, BMI, 21.4 ± 3.1 kg/m^2; 35 females, age, 22.0 ± 2.1 years, stature, 159.8 ± 5.7 cm, body weight, 53.3 ± 6.9 kg, BMI, 20.7 ± 2.0 kg/m^2) who met the criteria volunteered for the experiment. All participants were informed about the study procedures and expectations and provided written informed consent before participation.

Measurements

VO_2max test: Based on our pre-test, a continuous incremental treadmill test adapted from (Werneck et al., 2019) was applied to measure VO_{2max}. Participants performed a 10-minute warmup exercise at a velocity of 6 km/h and 0% slope before the formal test. After a 10-minute break, the formal test began with the velocity set at 8km/h and a slope of 4%. The velocity increased by 0.6 km/h every minute

until the destination of the test. When the respiratory exchange ratio (RER) approached 1, blood lactate was measured from the fingertip using EKF Lactate Scout+ (Barleben, Germany) in the last 30 seconds of each stage. A portable COSMED K5 metabolic analyzer (Rome, Italy) was used to collect relevant gas data. The criteria for VO_{2max} were based on the new proposals by (Edvardsen et al., 2014): 1) for females aged 20-49, blood lactate ≥ 7.0 mmol/L and the RER ≥ 1.1; 2) for males aged 20-49, blood lactate ≥ 9.0 mmol/L and RER ≥ 1.1. For safety reasons, the test was also terminated if the participants reached volitional exhaustion and could not continue. Average oxygen uptake for the last 30 seconds was considered VO_{2max}.

Determining velocities for RE test: Given that VO_{2max} is a widely recognized and standardized measure of aerobic capacity, using VO_{2max} as a reference point allows for comparability across studies and populations (Mann et al., 2013). Furthermore, VO_{2max} is generally more reliable than threshold measurements, such as lactate threshold, which can be complex, time-consuming, and challenging to implement consistently (Mann et al., 2013). Consequently, moderate-intensity exercise set at a percentage of VO_{2max} is less likely to elicit individual variation in blood lactate accumulation (Mann et al., 2013). Therefore, many researchers favor using $\%VO_{2max}$, and the RE test set relative to VO_{2max} in this study is reasonable. Given that submaximal intensity exercise typically corresponds to a range of 50% to 85% of VO_{2max} (Jeukendrup and Wallis, 2005; Williams et al., 1991), we applied 55%, 65%, and 75% of VO_{2max} to represent the submaximal intensities in RE testing. The velocities corresponding to 55%, 65%, and $75\%VO_{2max}$ were referred to the method of (Morgan and Daniels, 1994). Based on the data of velocities and corresponding oxygen consumption (60%-$90\%VO_{2max}$) collected from the VO_{2max} test, a linear model was established. Once the linear model was created, velocities corresponding to 55%, 65%, and $75\%VO_{2max}$ were initially calculated, followed by adjustment and determination in the laboratory. Participants were required to run at each calculated velocity for 6 minutes. The average oxygen consumption for the last 2 minutes was used to determine if the intensity triggered by the calculated velocities falls within 55% ± 5%, 65% ± 5%, and $75\% \pm 5\%VO_{2max}$. If the intensity was outside the expected range, the calculated velocity was adjusted by 1km/h, and another 6-minute run was required to verify. The total trials should be no more than 5 times, and each 6-minute run was followed by a 10-minute break. The treadmill slope was set at 1% to simulate outdoor running conditions (Hung et al., 2019).

RE test: Comparisons based on absolute velocity neglect the differences in relative intensity (Fletcher et al., 2009) and therefore, $\%VO_{2max}$ was applied to control the intensity in this study. Participants performed a 10-minute warm-up at $50\%VO_{2max}$ on a treadmill with a 0% slope. After a 5-minute resting period, the formal RE test began. Participants should run at the velocities corresponding to 55%, 65%, and 75% of VO_{2max} on a treadmill set at a 1% slope for 6 minutes. The average oxygen uptake and carbon dioxide production during the last 2 minutes of each run were calculated and used to calculate the energy cost based on the updated conversion formulae proposed by (Jeukendrup and Wallis, 2005). Blood lactate was collected immediately after each 6-minute run. A 5-minute standing rest was given before the next run. Steady-state criteria were based on suggestions by (Fletcher et al., 2009; McMiken and Daniels, 1976): 1) RER ≤ 0.95; 2) an increase of < 100 ml O_2 over the final 2 minutes of each stage. If the increase exceeded 100 ml, this run was extended by another 30 seconds.

Field test: A break of at least 1 day was given after the RE test, and the 1000-meter test was conducted on a standard athletic field within 7 days after the completion of the RE test. Before the test, participants performed a 5-minute 'quick warm-up cardio workout' followed by another 5-minute freestyle warm-up exercise. This 'quick warm-up cardio workout' included 10 different exercises (boxer shuffle, overhead reach + stretch, high knee march, torso twists, toe touch kicks, full torso circles, lateral step toe touches, squats, jumping jacks, high knees) with 30 seconds apiece. To simulate a competitive environment during the test, participants who had completed the laboratory tests competed in the field test together rather than individually. Strong verbal encouragement was provided by the same assistant throughout the test. Time was recorded by the same qualified referee using a stopwatch. The average outdoor temperature and relative humidity were 27.3 ± 3.0 °C and 78.2 ± 2.9 %, respectively.

Explanation for tests employed in this study

All the laboratory tests worked for the RE test. Although the slope used in VO_{2max} testing (4%) differed from that employed in the RE test (1%), thereby complicating the determination of appropriate speed for the RE test, this protocol was adopted to ensure participant safety and to obtain reliable VO_{2max} measurements, which is critical to determine relative exercise intensities. In the RE test, the intensity was controlled to ensure consistency among individuals. Therefore, $\%VO_{2max}$ was applied to control the intensity, and the treadmill velocity varied for each participant. A continuous incremental treadmill test of VO_{2max} was applied for each participant (Day 1), in which the treadmill velocity of each stage and the corresponding VO_{2sub} were also collected. Using the method described by (Morgan and Daniels, 1994), a simple linear model was established, with the velocity as the predictor and the corresponding VO_{2sub} as the outcome variable. This allowed for the preliminary calculation of treadmill velocities corresponding to $\%VO_{2max}$. However, due to limited data and different settings between the VO_{2max} test and the RE test, these calculated velocities were only for reference and facilitated the selection of required velocities for each participant to reach the target intensities. They were not directly applied in the RE test. Therefore, a pre-test for adjusting and verifying these preliminary velocities was necessary (Day 2). Once adjusted and confirmed, these velocities were used in the formal RE test (Day 3).

Energy consumption

Energy consumption during the RE test was calculated using energy conversion formulae recommended by (Jeukendrup and Wallis, 2005). These formulae are based on indirect calorimetry, which assumes that the RER adequately reflects the respiratory quotient (RQ) (Jeukendrup and Wallis, 2005). In plain language, oxygen consumption and carbon dioxide production are assumed to result solely from oxidative processes, and the gas composition measured in exhaled breath reflects gas exchanged from fuel metabolism at the tissue level. The indirect calorimetry technique for substrate oxidation determination, as described by (Jeukendrup and Wallis, 2005), has been proven reliable and valid for low- to moderate-intensity exercise using stable isotope techniques. Consequently, these energy conversion formulae can accurately calculate energy consumption from different substrates and provide a reliable measure of metabolic consumption during the RE test.

$$\text{Carbohydrate oxidation (g/min)} = 4.210VCO_2 - 2.962VO_2 \quad \mathbf{1}$$

$$\text{Fat oxidation (g/min)} = 1.695\ VO_2 - 1.701\ VCO_2 \quad \mathbf{2}$$

$$\text{Energy cost (kcal/min)} = 0.550\ VCO_2 + 4.471\ VO_2 \quad \mathbf{3}$$

Statistical analysis

Normality distribution was assessed by the Shapiro-Wilk test. Regression analysis based on the following equations was applied to explore the quantitative relationship between VO_{2sub} or E_c (averaged for the three intensities, $55\%VO_{2max}$, $65\%VO_{2max}$, and $75\%VO_{2max}$) and body weight.

$$\text{Linear function: } y = k\,x + d \quad \mathbf{4}$$

$$\text{Allometric scaling: } y = a\,x^b \quad \mathbf{5}$$

A natural logarithm transformation ($\ln y = \ln a + b \ln x$) was performed for equation **5** before conducting linear regression to calculate the b value. Specifically, the natural logarithm can be taken on both sides of equation **5**, which gives:

$$\ln y = \ln (a\,x^b) \quad \mathbf{6}$$

According to the sum and difference formulas for logarithmic functions ($\log_c m + \log_c n = \log_b (mn)$; $\log_b m - \log_b n = \log_b (m/n)$), equation **6** can be changed to

$$\ln y = \ln a + b \ln x \quad \mathbf{7}$$

Let $Y = \ln y$, $X = \ln x$, $C = \ln a$, the equation **7** can be written as

$$Y = b\,X + C \quad \mathbf{8}$$

Therefore, a linear regression was applied to calculate the b values. The probability of future similar studies observing the group differences and the probability of the exponent b being lower than 1 or the constant d being greater than 0 were interpreted according to the following scale (Hopkins and Batterham, 2016): <0.5%, most unlikely; 0.5-5%, very unlikely; 5-25%, unlikely; 25-75%, possible; 75-95%, likely; 95%-99.5%, very likely; >99.5%, most likely. Pearson Correlation Coefficient was applied to verify the influence of body weight on measures of RE and their correlations with running performance. If the model could succeed in accounting for body weight, the correlation between body weight and the assessed variable scaled appropriately to body weight should not exist (Lolli et al., 2017). The correlation coefficient r was interpreted according to the following ratings (Landis and Koch, 1977): no correlation or trivial correlation, $|r| < 0.2$; fair, $0.2 \le |r| < 0.4$; moderate, $0.4 \le |r| < 0.6$; substantial, $0.6 \le |r| < 0.8$; almost perfect, $|r| \ge 0.8$.

A one-way repeated-measures ANOVA was used to evaluate differences in VO_{2sub} and E_c as measures of RE across three velocities. *Post hoc* analysis with Bonferroni adjustment was applied to identify where any significant differences occurred. Magnitude-based inferences proposed by (Hopkins et al., 2009) were utilized to reveal RE differences between males and females, as they provide a more accurate and informative ternary result with probability and magnitude (e.g., possibly substantially positive, possibly trivial, possibly substantially negative), avoiding the ‘all or nothing’ thinking that p-values encourage. Cohen’s D effect sizes (ES) were calculated to reflect the extent of the difference, with the inferences associated with the effects defined as trivial (<0.20), small (0.20-0.59), moderate (0.60-1.19), large (1.2-1.9), very large (2.0-3.9) and extremely large (≥4.0). All the data were processed in SPSS (version 20.0., IBM, Chicago, USA). The significance level was set at 0.05.

Results

Metabolic characteristics under different intensities

Relevant metabolic and physiological information is provided in Table 1. The average VO_{2max} for males and females was 3.78 ± 0.43 and 2.35 ± 0.44 L/min, respectively. The average VO_{2sub} at $55\%VO_{2max}$, $65\%VO_{2max}$, and $75\%VO_{2max}$ were 2.15 ± 0.26, 2.48 ± 0.27, and 2.77 ± 0.33 L/min, respectively, for males, and 1.35 ± 0.25, 1.58 ± 0.29, and 1.85 ± 0.32 L/min, respectively, for females. The corresponding E_c for males were 10.72 ± 1.27, 12.39 ± 1.35, and 13.83 ± 1.62 kcal/min, respectively, while those for females were 6.71 ± 1.21, 7.87 ± 1.42, and 9.23 ± 1.59 kcal/min, respectively.

Moreover, based on equations **1** and **2**, the fat utilization rate (g/min) and the corresponding percentage of energy derived from fat (fat%) were calculated. Specifically, concerning the fat utilization rate in males, it experienced an increase from 0.18 g/min ($55\%VO_{2max}$) to 0.23 g/min ($65\%VO_{2max}$) before decreasing back to 0.18 g/min ($75\%VO_{2max}$). In contrast, this rate consistently decreased in females, experiencing a slight decrease from 0.26 g/min ($55\%VO_{2max}$) to 0.24 g/min ($65\%VO_{2max}$) and ultimately reaching 0.19 g/min at $75\%VO_{2max}$. Although fat% for males was relatively stable when the intensity increased from $55\%VO_{2max}$ to $65\%VO_{2max}$, it still displayed a decreasing trend in both males and females (male: 16%, 18%, and 13% at $55\%VO_{2max}$, $65\%VO_{2max}$, and $75\%VO_{2max}$, respectively; female: 37%, 31%, and 20% at $55\%VO_{2max}$, $65\%VO_{2max}$, and $75\%VO_{2max}$, respectively). Together, these suggested the metabolic shift towards carbohydrates as intensity increased.

Table 1. Metabolic and physiological characteristics. Data are means ± SD.

Intensity		VO_{2sub} (L/min)	VO_{2max} (L/min)	E_c (kcal/min)	Blood lactate (mmol/L)	RER
55%VO_{2max}	Male	2.15±0.26	-	10.72±1.27	2.56±1.12 (range: 1.3-5.3)	0.95±0.08
	Female	1.35±0.25	-	6.71±1.21	2.09±0.93 (range: 0.7-4.5)	0.89±0.06
65%VO_{2max}	Male	2.48±0.27	-	12.39±1.35	2.82±1.11 (range: 1.1-5.8)	0.94±0.06
	Female	1.58±0.29	-	7.87±1.42	2.24±1.18 (range: 0.7-4.9)	0.91±0.06
75%VO_{2max}	Male	2.77±0.33	-	13.83±1.62	3.24±1.01 (range: 0.8-5.1)	0.96±0.07
	Female	1.85±0.32	-	9.23±1.59	2.82±0.98 (range: 1.0-4.9)	0.94±0.07
VO_{2max}	Male	-	3.78±0.43		13.8±4.1	1.11±0.06
	Female	-	2.35±0.44		11.2±3.4	1.16±0.10

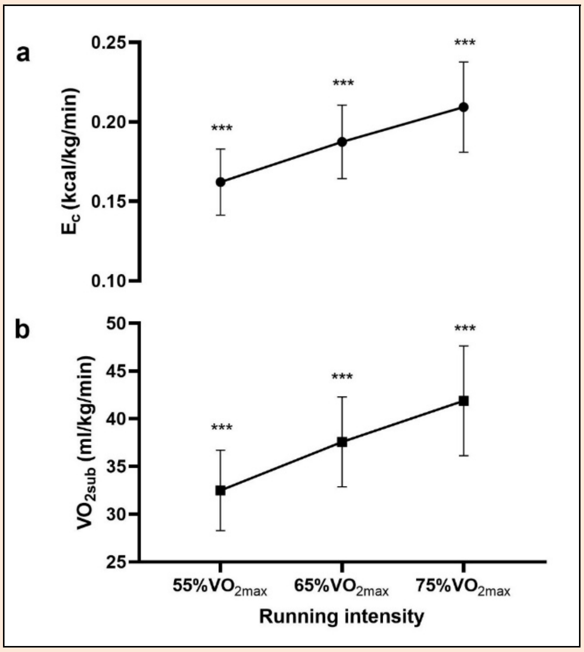

Figure 1. Physiological measures for males at three speeds relative to %VO_{2max}. *** Significant difference from all other intensities (ANOVA with Bonferroni post hoc adjustment; $p < 0.001$.

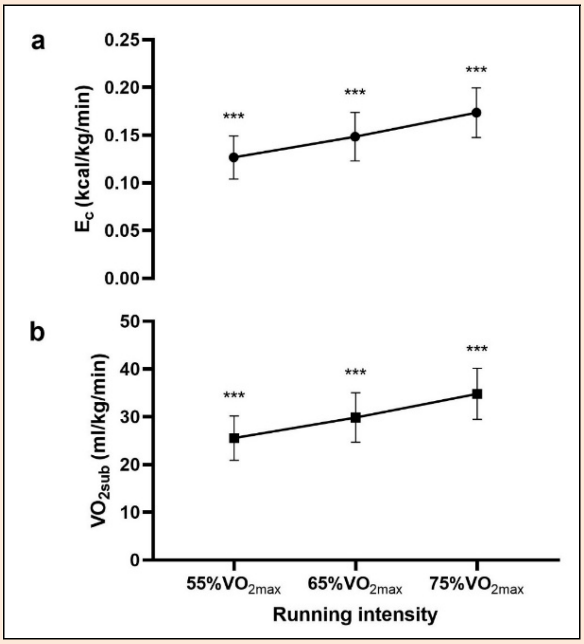

Figure 2. Physiological measures for females at three speeds relative to %VO_{2max}. *** Significant difference from all other intensities (ANOVA with Bonferroni post hoc adjustment; $p < 0.001$).

Increased velocity and characteristics of different RE measures

The one-way repeated-measures ANOVA revealed increases in E_c with increments in running intensity, regardless of sex ($p < 0.001$; Figure 1a and Figure 2a). Increases in VO_{2sub} were also observed with increasing running intensity, regardless of sex (ANOVA, $p < 0.001$; Figure 1b and Figure 2b).

Regression of RE with different measures under different intensities

In this study, "$b_{overall}$" is defined as the exponent b value calculated from the relationship between body weight and VO_{2sub} or E_c averaged across three intensities. The results of regression analysis with allometric scaling showed significant fits between absolute E_c and body weight (males: averaged across three intensities, $b_{overall}$ = 0.48, 95%CI = 0.11-0.85; 55%VO_{2max}, b = 0.53, 95%CI = 0.14-0.92; 65%VO_{2max}, b = 0.48, 95%CI = 0.10-0.87; 75%VO_{2max}, b = 0.44, 95%CI = 0.04-0.85; females: averaged across three intensities, $b_{overall}$ = 0.71, 95%CI = 0.28-1.14; 55%VO_{2max}, b = 0.63, 95%CI = 0.14-1.11; 65%VO_{2max}, b = 0.67, 95%CI = 0.20-1.14; 75%VO_{2max}, b = 0.81, 95%CI = 0.38-1.23) and between VO_{2sub} and body weight (males: averaged across three intensities, $b_{overall}$ = 0.47, 95%CI = 0.10-0.85; 55%VO_{2max}, b = 0.52, 95%CI = 0.13-0.91; 65%VO_{2max}, b = 0.48, 95%CI = 0.09-0.89; 75%VO_{2max}, b = 0.43, 95%CI = 0.03-0.84; females: averaged across three intensities, $b_{overall}$ = 0.71, 95%CI = 0.27-1.15; 55%VO_{2max}, b = 0.63, 95%CI = 0.14-1.13; 65%VO_{2max}, b = 0.68, 95%CI = 0.20-1.15; 75%VO_{2max}, b = 0.79, 95%CI = 0.36-1.22). For future similar studies, the probability of exponent $b_{overall}$ (derived from E_c and body weight) being less than 1 was 99.6% (most likely) in males and 90.7% (likely) in females. Furthermore, the probabilities of exponent b (derived from E_c and body weight) being less than 1 at specific exercise intensities was as follow: in males, 98.9% (very likely) at 55%VO_{2max}, 99.5% (most unlikely) at 65%VO_{2max}, and 99.6% (most likely) at 75%VO_{2max}; in females, 93.7% (likely) at 55%VO_{2max}, 91.9% (likely) at 65%VO_{2max}, and 82% (likely) at 75%VO_{2max}. Likewise, the probabilities of the exponent $b_{overall}$ (calculated from VO_{2sub} and body weight) being less than 1 in future similar studies were 99.6% for males (most likely) and 91.0% for females (likely). Furthermore, the probabilities of exponent b (derived from VO_{2sub} and body weight) being less than at specific exercise intensities was as follow: in males, 99.1% (very likely) at 55%VO_{2max}, 99.5% (very likely) at 65%VO_{2max}, and 99.6% (most likely) at 75%VO_{2max}; in females, and 93.0% (likely) at 55%VO_{2max}, 91.9% (likely) at 65%VO_{2max}, and 82.3% (likely) at 75%VO_{2max}. Similar but marginally higher R^2 values were observed for allometric scaling compared to the linear function, regardless of whether the dependent variable was E_c or VO_{2sub}. Furthermore, R^2 was consistently higher when allometric scaling was applied with E_c as the dependent variable and body weight as the independent variable, compared to when VO_{2sub} was the dependent variable and body weight was the independent variable (Table 2).

Removal effect of body weight

The appropriateness of allometric scaling was confirmed by the absence of any relationship when body weight was replotted against power-scaled E_c (kcal/kg^b/min) and VO_{2sub} (ml/kg^b/min). In contrast, the significant negative correlations between ratio-scaled E_c (kcal/kg/min) or ratio-scaled VO_{2sub} (ml/kg/min) and body weight in males suggested the failure of ratio scaling in removing the influence of body weight. Although the correlations were not significant in females when body weight was replotted against ratio-scaled E_c (kcal/kg/min) and VO_{2sub} (ml/kg/min), there was still a trend towards trivial or even fair negative correlations (Table 3).

Correlations between RE and running performance

The 1000 meters running performance for males was 224.71 ± 17.10 s, while the performance for females was 299.92 ± 33.53 s.

Table 4 displays the correlations between RE expressed by different measures and running performance. Regarding correlations between allometric-scaled E_c and performance in males, significantly positive correlations were observed at 55%VO_{2max} (r = 0.43 [moderate]), 65%VO_{2max} (r = 0.51 [moderate]), 75%VO_{2max} (r = 0.50 [moderate]), and overall intensity (averaged across three intensities, r = 0.51 [moderate]). In females, however, significantly positive correlations were observed at 55%VO_{2max} (r = 0.35 [fair]), 65%VO_{2max} (r = 0.46 [moderate]), and overall intensity (r = 0.38 [fair]). Regarding correlations between ratio-scaled E_c and performance in males, significantly positive correlations were observed at 55%VO_{2max} (r = 0.35 [fair]), 65%VO_{2max} (r = 0.41 [moderate]), 75%VO_{2max} (r = 0.41 [moderate]), and overall intensity (r = 0.41 [moderate]). In females, however, significantly positive correlations were observed at 55%VO_{2max} (r = 0.34 [fair]), 65%VO_{2max} (r = 0.45 [moderate]), and overall intensity (r = 0.37 [fair]). Regarding correlations between allometric-scaled VO_{2sub} and performance in males, significantly positive correlations were observed at 55%VO_{2max} (r = 0.42 [moderate]), 65%VO_{2max} (r = 0.51 [moderate]), 75%VO_{2max} (r = 0.50 [moderate]), and overall intensity (r = 0.51 [moderate]). In females, however, significantly positive correlations were observed at 55%VO_{2max} (r = 0.35 [fair]), 65%VO_{2max} (r = 0.45 [moderate]), and overall intensity (r = 0.37 [fair]). Regarding correlations between ratio-scaled VO_{2sub} and performance in males, significantly positive correlations were observed at 55%VO_{2max} (r = 0.35 [fair]), 65%VO_{2max} (r = 0.41 [moderate]), 75%VO_{2max} (r = 0.40 [moderate]), and overall intensity (r = 0.41 [moderate]). In females, however, significantly positive correlations were observed at 65%VO_{2max} (r = 0.44 [moderate]), and overall intensity (r = 0.36 [fair]). A fair but nonsignificant correlation (r = 0.33) between ratio-scaled VO_{2sub} (at 55%VO_{2max}) and performance was observed in females. Likewise, at 75%VO_{2max}, none of the expressions detected significant correlations in females, although their correlations were fair. Overall, RE measured by E_c consistently showed a marginally higher correlation with running performance and was more sensitive in detecting correlations with performance compared to VO_{2sub}. Moreover, a better correlation was observed when RE was scaled by allometric scaling rather than ratio scaling.

In addition, females consistently displayed better RE than their male counterparts, regardless of whether RE was calculated from E_c or VO_{2sub}, scaled by ratio or allometry. Specifically, magnitude-based inferences revealed large differences when RE was expressed by ratio-scaled E_c (ES = 1.55, RE was averaged across three intensities; ES = 1.66, 1.64, 1.23 for 55%, 65%, and 75%VO_{2max}, respectively) and VO_{2sub} (ES = 1.51, RE was averaged across three intensities; ES = 1.61, 1.60, and 1.2 for 55%, 65%, and 75%VO_{2max}, respectively), all with 100% possibility. Magnitude-based inferences also revealed extremely large differences when RE was expressed by allometric-scaled E_c (ES = 6.63, RE was averaged across three intensities;

Table 2. The coefficient of determination (R^2) from linear and allometric scaling between body weight and both E_c and VO_{2sub}.

Variable	Gender	Intensity	R^2 Linear	R^2 Allometric
E_c (kcal/min)	Male	55%VO_{2max}	0.162	0.192
		65%VO_{2max}	0.153	0.169
		75%VO_{2max}	0.115	0.136
		Overall	0.155	0.178
	Female	55%VO_{2max}	0.159	0.174
		65%VO_{2max}	0.194	0.204
		75%VO_{2max}	0.301	0.314
		Overall	0.242	0.255
VO_{2sub} (ml/min)	Male	55%VO_{2max}	0.155	0.184
		65%VO_{2max}	0.148	0.166
		75%VO_{2max}	0.106	0.128
		Overall	0.147	0.171
	Female	55%VO_{2max}	0.156	0.171
		65%VO_{2max}	0.193	0.202
		75%VO_{2max}	0.284	0.298
		Overall	0.233	0.246

Table 3. Correlations between RE expressions and body weight.

Intensity	Gender	RE expressions	Body weight (kg)	
			r	p
55%VO_{2max}	Male	kcal/kg/min	-0.41	0.02
		ml/kg/min	-0.42	0.01
		kcal/kg$^{0.53}$/min	-0.04	0.81
		ml/kg$^{0.52}$/min	-0.04	0.81
	Female	kcal/kg/min	-0.27	0.11
		ml/kg/min	-0.26	0.13
		kcal/kg$^{0.63}$/min	-0.03	0.86
		ml/kg$^{0.63}$/min	-0.03	0.87
65%VO_{2max}	Male	kcal/kg/min	-0.47	0.01
		ml/kg/min	-0.47	<0.01
		kcal/kg$^{0.48}$/min	-0.05	0.77
		ml/kg$^{0.48}$/min	-0.05	0.77
	Female	kcal/kg/min	-0.24	0.16
		ml/kg/min	-0.23	0.18
		kcal/kg$^{0.67}$/min	-0.02	0.92
		ml/kg$^{0.68}$/min	-0.02	0.93
75%VO_{2max}	Male	kcal/kg/min	-0.46	0.01
		ml/kg/min	-0.47	<0.01
		kcal/kg$^{0.44}$/min	-0.04	0.82
		ml/kg$^{0.43}$/min	-0.04	0.81
	Female	kcal/kg/min	-0.18	0.32
		ml/kg/min	-0.18	0.30
		kcal/kg$^{0.81}$/min	-0.03	0.88
		ml/kg$^{0.79}$/min	-0.03	0.89
Overall	Male	kcal/kg/min	-0.47	<0.01
		ml/kg/min	-0.47	<0.01
		kcal/kg$^{0.48}$/min	-0.05	0.79
		ml/kg$^{0.47}$/min	-0.05	0.79
	Female	kcal/kg/min	-0.24	0.17
		ml/kg/min	-0.24	0.17
		kcal/kg$^{0.71}$/min	-0.03	0.88
		ml/kg$^{0.71}$/min	-0.03	0.88

ES = 4.45, 6.17, and 7 for 55%, 65%, and 75%VO_{2max}, respectively) and VO_{2sub} (ES = 6.61, RE was averaged across three intensities; ES = 4.68, 6.17, and 6.91 for 55%, 65%, and 75%VO_{2max}, respectively), all with 100% possibility.

Discussion

Previous studies have mainly focused on the application of ratio-scaled VO_{2sub} (Hung et al., 2019). However, VO_{2sub} cannot avoid the influence of substrate-related issues, and the application of ratio scaling has been met with skepticism. E_c may be a more reliable measure, but previous studies exploring allometric scaling applied to E_c are limited and lack verification of practical application (Shaw et al., 2014; Zakeri et al., 2006). Thus, this study aimed to explore: 1) the validity of VO_{2sub} as a quantification of RE by comparing this variable with the underlying E_c of running; 2) the appropriateness of ratio scaling in normalizing RE measures. No disparity was found between changes in VO_{2sub} and E_c as running intensity increased. A further finding was that allometric scaling, rather than ratio scaling, was more appropriate to account for body weight when using both measures to quantify RE. Moreover, RE expressed by allometric-scaled E_c (kcal/kg^b/min) showed a higher correlation with running performance.

Comparison among different RE measurements

It has been argued that VO_{2sub}, as the quantification of RE, may not accurately reflect the underlying energy cost since it ignores issues related to substrates (di Prampero et al., 2009; Fletcher et al., 2009; Jeukendrup and Wallis, 2005; Shaw et al., 2014). Instead, direct quantification of RE using E_c considers this influence and aligns with the definition of RE, making it a more accurate, meaningful measure. This assertion is, at least partly, supported by the consistently higher R^2 values in the regression models (regardless of allometric scaling or linear function) when E_c was the dependent variable (Table 2).

However, other findings in this study do not entirely reject the validity of VO_{2sub}. As indicated by the one-way repeated-measures ANOVA, E_c (kcal/kg/min) increased progressively with rising running intensity. This trend is consistent with previous studies indicating that energy cost increases as exercise intensity rises (Beneke and Leithäuser, 2017; Fletcher et al., 2009; Shaw et al., 2014). As intensity increases, the demand for adenosine triphosphate (ATP) also rises, prompting a metabolic shift toward greater carbohydrate utilization. Although fat oxidation yields more energy per gram than carbohydrate oxidation (9.75 kcal vs. 4.07 kcal), the high efficiency of carbohydrates in producing "quicker" energy is more crucial in higher-intensity exercise that demands a greater amount of ATP, which in turn contributes to heightened oxygen consumption (Costill, 1988; Jeukendrup and Wallis, 2005).

Table 4. Correlations between different RE expressions and running performance.

Intensity	Gender	RE Expressions	RE Means±SD	r
$55\%VO_{2max}$	Male	kcal/kg/min	0.16±0.02	0.35*
		$kcal/kg^{0.53}/min$	1.16±0.13	0.43*
		ml/kg/min	32.49±4.15	0.35*
		$ml/kg^{0.52}/min$	243.59±27.46	0.42*
	Female	kcal/kg/min	0.13±0.02	0.34*
		$kcal/kg^{0.63}/min$	0.56±0.09	0.35*
		ml/kg/min	25.56±4.59	0.33
		$ml/kg^{0.63}/min$	109.97±18.88	0.35*
$65\%VO_{2max}$	Male	kcal/kg/min	0.19±0.02	0.41*
		$kcal/kg^{0.48}/min$	1.63±0.17	0.51**
		ml/kg/min	37.57±4.63	0.41*
		$ml/kg^{0.48}/min$	330.31±34.67	0.51**
	Female	kcal/kg/min	0.15±0.02	0.45**
		$kcal/kg^{0.67}/min$	0.55±0.09	0.46**
		ml/kg/min	29.87±5.10	0.44**
		$ml/kg^{0.68}/min$	107.99±17.91	0.45**
$75\%VO_{2max}$	Male	kcal/kg/min	0.21±0.03	0.41*
		$kcal/kg^{0.44}/min$	2.16±0.25	0.50**
		ml/kg/min	41.89±5.67	0.40*
		$ml/kg^{0.43}/min$	453.96±52.19	0.50**
	Female	kcal/kg/min	0.17±0.03	0.27
		$kcal/kg^{0.81}/min$	0.37±0.05	0.27
		ml/kg/min	34.83±5.28	0.26
		$ml/kg^{0.79}/min$	79.17±11.81	0.26
Overall	Male	kcal/kg/min	0.19±0.02	0.41*
		$kcal/kg^{0.48}/min$	1.64±0.17	0.51*
		ml/kg/min	37.31±4.62	0.41*
		$ml/kg^{0.47}/min$	340.64±35.77	0.51*
	Female	kcal/kg/min	0.15±0.02	0.37*
		$kcal/kg^{0.71}/min$	0.47±0.07	0.38*
		ml/kg/min	30.09±4.74	0.36*
		$ml/kg^{0.71}/min$	95.04±14.52	0.37*

*, $p < 0.05$; **, $p < 0.01$.

Our findings support this statement, as we observed a decrease in fat% and an increase in VO_{2sub} with increasing exercise intensity. Overall, the observed increase in VO_{2sub} (ml/kg/min) across running intensities in this study suggests that VO_{2sub} remains a valid indirect measure of RE, at least within the context of our participant cohort. This finding aligns with the concurrent rise in E_c, supporting the notion that VO_{2sub} can reflect changes in metabolic demand during incremental exercise. However, this result contrasts with previous studies that reported a divergence between E_c and VO_{2sub}, where E_c increased with intensity while VO_{2sub} remained relatively stable VO_{2sub} (Fletcher et al., 2009; Shaw et al., 2014). One plausible explanation is the difference in participant characteristics. Unlike earlier studies that focused on elite athletes, our study involved amateur runners, whose physiological responses and metabolic profiles may differ significantly (e.g., lower fat% in this study). Another possible explanation is that the conversion of the original unit, in which the speed information (km/h) was incorporated and the unit was changed from ml/kg/min to ml/kg/km (oxygen consumption per unit of distance), may have compromised its sensitivity to intensity. As a result, such conversion may lead to a misleading conclusion that RE was independent of speed changes (Fletcher et al., 2009; Shaw et al., 2014).

Although VO_{2sub} is a valid quantification of RE in this study, inter-individual comparisons in RE are valid only when the statistical adjustment is reliable and accurate enough to remove the influence of body weight. According to the results shown in Table 2, the R^2 of allometric scaling (from 0.128 to 0.314) was consistently higher than that of the linear function (ranging from 0.106 to 0.301), regardless of whether E_c or VO_{2sub} was used. Additionally, regardless of E_c or VO_{2sub}, the calculated exponent b ranged from 0.43 to 0.81, aligning closer to the 2/3 theory (Günther, 1975; Lee and Zhang, 2021). As early as 1838, the "surface law", which states that the basal metabolism of animals with different sizes is nearly proportional to their body surfaces, was recommended to describe the relationship between physiological function variables and body dimensions (Lee and Zhang, 2021). According to the law, metabolic rate should be proportional to body weight raised to the power of 2/3. Later, in 1927, Lambert and Teissier also confirmed this theoretical exponent b value for body weight from a new perspective—kinematic or biological similarity (KBS), which combined Newton's dimensional analysis with geometric similarity (Lee and Zhang, 2021). The KBS believes that the relationship between physiological function and body weight should satisfy the equation of "$x = w^{\alpha+\frac{1}{3}\beta+\frac{1}{3}\gamma}$", where x represents a physiological function and w represents body weight (Lee and Zhang, 2021). Given that α=-1, β=4, and γ=1 in the relationship between oxygen consumption and body weight, the same result, "2/3", can be calculated (Lee and

Zhang, 2021). Although the other theoretical b value, “3/4”, has been extrapolated to scale body weight, with some empirical studies on the relationship between VO_{2max} and body weight supporting it, no consensus has yet been reached on which theoretical b value is better (Lee and Zhang, 2021; Lolli et al., 2017; West et al., 1997). However, (Darveau et al., 2002) believed that the increased importance of energy-supply processes and decreased reliance of energy-demand processes due to the shifts of metabolic rate from resting to maximal condition may cause a significant increase in b value. Moreover, some scholars believed that both 2/3 and 3/4 laws are applicable, but the 2/3 law might be more suitable for an intraspecific (within-species) mass exponent (Batterham et al., 1999). Thus, the 2/3 law appears to be more appropriate for application in scaling RE, and our findings, at least partly, support this law. Furthermore, the probabilities of these b values being less than 1 ranged from likely to most likely. These results indicate that the traditional ratio-scaled RE (ml/kg/min) is untenable. Instead, allometric scaling with a higher R^2 should be a more accurate statistical adjustment for RE measurement. Further analysis of the ‘removal effect of body weight’ (Table 3) supports this viewpoint, as no correlations with body weight were observed when allometric scaling, rather than ratio scaling, was applied. This indicates that RE expressed by allometric scaling can effectively eliminate the influence of body weight, whereas the linear function is not powerful enough to exclude that influence. (Chamari et al., 2005) stated that the RE of senior soccer players might be overestimated when expressed in the traditional way (ml/kg/min). Inappropriate statistical adjustments may lead to faulty information, thereby misleading coaches' training schedules.

When comparing allometric-scaled E_c and allometric-scaled VO_{2sub}, the former consistently showed a better R^2 (Table 2). Combined with the results of the ‘removal effect of body weight’, allometric-scaled E_c can better improve the accuracy of RE compared to allometric-scaled $VO_{2sub.}$ Overall, although VO_{2sub} is a valid quantification of RE in this study, consistently higher R^2 values in regression models with E_c as the dependent variable suggest E_c to be a more appropriate quantification of RE. Furthermore, allometric scaling, rather than ratio scaling, should be applied to normalize these quantifications in inter-individual comparisons. For the convenience of calculation, the 2/3 law can be considered for the body weight exponent.

RE and its correlations with performance

Due to the unreliable nature of RE reflected by the ratio-scaled VO_{2sub}, it is recommended to reconsider findings and conclusions to avoid providing misleading feedback or information to coaches or scholars (e.g., overestimating one’s real RE, inappropriate training arrangements).

RE is a controversial topic in terms of gender. Some studies have found that males have better RE than females, while others have argued that females have better or that no gender difference exists (Ariëns et al., 1997; Daniels and Daniels, 1992; Davies et al., 1997; Helgerud, 1994; Janz et al., 1998). A key factor contributing to these inconsistent findings is the measure of RE. In this study, we expressed RE with allometric-scaled E_c and observed similar results to findings by (Helgerud, 1994), suggesting that males’ RE is not better than females’ RE. This is also supported by the relatively higher RER (Table 1) and lower fat% in males compared to females, suggesting that males relied more on carbohydrates as an energy source. Magnitude-based inferences suggest that the likelihood of future similar studies observing these extremely large gender differences in RE is most likely. Although a definitive conclusion cannot be made based on the limited evidence, the uncertainty of RE in gender is unlikely due to the sample’s sports background. Specifically, when RE was scaled by ratio-scaled VO_{2sub}, no gender difference was observed in a study by (Helgerud, 1994) with intermediate national standard elite athletes, or by (G et al., 2020) with amateur runners similar to our study. Therefore, while ratio-scaled VO_{2sub} in the current study also successfully detected large differences in RE between females and males, the mathematical limitations of ratio scaling raise concerns about the validity of such findings. Based on these considerations, we speculated that statistical defects of ratio scaling in normalizing RE are the critical reason for controversy concerning gender differences in RE, contributing to misleading information. More studies based on allometric scaling are needed before making a potent conclusion regarding the gender difference in RE.

To further evaluate the ability of allometric-scaled E_c in explaining performance, we applied different expressions of RE in practice. Correlation analyses revealed that RE scaled by allometric E_c consistently displayed a better correlation with running performance and was more sensitive in detecting correlations at increasing intensities. Notably, these correlations remained evident in male participants but disappeared in females at 75%VO_{2max}. This finding raises important considerations regarding the appropriateness of exercise intensity when assessing RE. Normally, submaximal intensity in RE should not exceed 85% VO_{2max} (Williams et al., 1991) to ensure a stable physiological state. However, this recommendation was primarily based on data from male athletes with relatively high aerobic capacities and may not be applicable to amateur populations, especially female runners. In our study, elevated blood lactate levels and RER observed in females at 75% VO_{2max} (see Table 1) suggest a deviation from steady-state conditions. The elevated intensity may induce greater neuromuscular demand and activation of type II fibers (Hargreaves and Spriet, 2020), which may help to cope with this high exercise intensity anaerobically and result in an unstable state (i.e., a small increase in oxygen uptake). Therefore, the sensitivity and accuracy of RE scaled by allometric E_c may be compromised, thereby weakening the observed correlations with performance. However, the anaerobic contribution to ATP production should be limited, as only approximately 15% of ATP is attributed to anaerobic contribution during a 6-minute maximal exercise (Gastin, 2001). Overall, the recommended intensity for the RE test limited at around 65%VO_{2max} seems to be more appropriate, as the highest correlations were observed around this intensity.

Limitations

The present study has several limitations. First, additional activities may influence athletic performance. Although participants' exercise frequency and intensity were self-reported, we did not strictly monitor their daily physical activities using validated pedometers. Second, this study only included recreationally active male and female college students. Therefore, the findings should be limited to this population. Meanwhile, the sample size of this study is still small. To ensure the applicability, reliability, and accuracy of RE measured by allometric-scaled E_c, further studies should increase the sample size and focus on subjects with different characteristics or backgrounds (e.g., athletes). Third, it was challenging to assess whether participants exerted their best to complete the 1000-meter field test, despite arranging a group of participants in the test to simulate the tournaments and providing them with strong encouragement. Moreover, since it was impossible for participants to complete all the laboratory tests on the same day, the 1000-meter outdoor performance took place on different days. This may influence the validity of the performance because multiple confounding factors, including wind speed, surface moisture and friction, drafting of other participants, and pacing familiarity, may influence the performance. Future studies may consider an indoor performance test to better control these factors. Fourth, the menstrual cycle was not strictly controlled at the same phase (e.g., luteal phase) for female participants, although tests were rescheduled to avoid periods when female participants were menstruating. Finally, some individuals experienced relatively higher blood lactate (i.e., blood lactate greater than the commonly used reference point for lactate threshold at 4 mmol/L) and RER during the RE tests, which may indicate the 'false' steady state. Although the small proportion of them (~10%) may not greatly influence the power of our findings, future studies should consider improving the protocols, including extending the break time before each test, monitoring baseline blood lactate, and evaluating the specific lactate threshold for each participant.

Conclusion

For the convenience of application, VO_{2sub} can be a valid quantification of RE. However, its application should be restricted to individuals who share characteristics similar to those in this study (i.e., amateur runners). In contrast, E_c, as the quantification of RE, is recommended for studies seeking accuracy, as it aligns with the definition of RE and can reflect the underlying energy cost during submaximal intensity running. 65%VO_{2max} intensity as the submaximal testing intensity in the RE test is recommended. Furthermore, allometric scaling, rather than ratio scaling, is more appropriate to normalize RE quantifications, as it can effectively remove the influence of body weight in inter-individual comparisons. For the convenience of allometric calculation, directly applying the 2/3 law in allometric scaling seems acceptable.

Acknowledgements
The experiments comply with the current laws of the country in which they were performed. The authors have no conflict of interest to declare. The datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author who was an organizer of the study.

References

Abad, C. C., Barros, R. V., Bertuzzi, R., Gagliardi, J. F., Lima-Silva, A. E., Lambert, M. I. and Pires, F. O. (2016) 10 km running performance predicted by a multiple linear regression model with allometrically adjusted variables. *Journal of Human Kinetics* **51**, 193-200. https://doi.org/10.1515/hukin-2015-0182

Ariëns, G. A., van Mechelen, W., Kemper, H. C. and Twisk, J. W. (1997) The longitudinal development of running economy in males and females aged between 13 and 27 years: the Amsterdam Growth and Health Study. *European Journal of Applied Physiology and Occupational Physiology* **76(3)**, 214-220. https://doi.org/10.1007/s004210050239

Armstrong, N. and Welsman, J. (2019) Clarity and confusion in the development of youth aerobic fitness. *Frontiers in Physiology* **10**, 979. https://doi.org/10.3389/fphys.2019.00979

Armstrong, N., Welsman, J. R. and Kirby, B. J. (1999) Submaximal exercise and maturation in 12-year-olds. *Journal of Sports Sciences* **17(2)**, 107-114. https://doi.org/10.1080/026404199366208

Batterham, A. M., Vanderburgh, P. M., Mahar, M. T. and Jackson, A. S. (1999) Modeling the influence of body size on VO_2peak: Effects of model choice and body composition. *Journal of Applied Physiology* **87(4)**, 1317-1325. https://doi.org/10.1152/jappl.1999.87.4.1317

Beneke, R. and Leithäuser, R. (2017) Energy cost of running related to running intensity and peak oxygen uptake. *Deutsche Zeitschrift für Sportmedizin* **68**, 196-202. https://doi.org/10.5960/dzsm.2017.296

Blagrove, R. C., Howatson, G. and Hayes, P. R. (2017) Test-retest reliability of physiological parameters in elite junior distance runners following allometric scaling. *European Journal of Sport Science* **17(10)**, 1231-1240. https://doi.org/10.1080/17461391.2017.1364301

Chamari, K., Moussa-Chamari, I., Boussaïdi, L., Hachana, Y., Kaouech, F. and Wisløff, U. (2005) Appropriate interpretation of aerobic capacity: Allometric scaling in adult and young soccer players. *British Journal of Sports Medicine* **39(2)**, 97-101. https://doi.org/10.1136/bjsm.2003.010215

Costill, D. L. (1988) Carbohydrates for exercise: Dietary demands for optimal performance. *International Journal of Sports Medicine* **9(1)**, 1-18. https://doi.org/10.1055/s-2007-1024971

Daniels, J. and Daniels, N. (1992) Running economy of elite male and elite female runners. *Medicine and Science in Sports and Exercise* **24(4)**, 483-489. https://doi.org/10.1249/00005768-199204000-00015

Darveau, C. A., Suarez, R. K., Andrews, R. D. and Hochachka, P. W. (2002) Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature* **417(6885)**, 166-170. https://doi.org/10.1038/417166a

Davies, M. J., Mahar, M. T. and Cunningham, L. N. (1997) Running economy: Comparison of body mass adjustment methods. *Research Quarterly for Exercise and Sport* **68(2)**, 177-181. https://doi.org/10.1080/02701367.1997.10607995

di Prampero, P. E., Salvadego, D., Fusi, S. and Grassi, B. (2009) A simple method for assessing the energy cost of running during incremental tests. *Journal of Applied Physiology* **107(4)**, 1068-1075. https://doi.org/10.1152/japplphysiol.00063.2009

Edvardsen, E., Hem, E. and Anderssen, S. A. (2014) End criteria for reaching maximal oxygen uptake must be strict and adjusted to sex and age: A cross-sectional study. *Plos One* **9(1)**, e85276. https://doi.org/10.1371/journal.pone.0085276

Fletcher, J. R., Esau, S. P. and Macintosh, B. R. (2009) Economy of running: Beyond the measurement of oxygen uptake. *Journal of Applied Physiology* **107(6)**, 1918-1922. https://doi.org/10.1152/japplphysiol.00307.2009

G, V. M., Matos, P. and Correia, J. M. (2020) Running economy in recreational male and female runners with similar levels of cardiovascular fitness. *Journal of Applied Physiology* **129(3)**, 508-515. https://doi.org/10.1152/japplphysiol.00349.2020

Gastin, P. B. (2001) Energy system interaction and relative contribution during maximal exercise. *Sports Medicine* **31(10)**, 725-741. https://doi.org/10.2165/00007256-200131100-00003

Günther, B. (1975) Dimensional analysis and theory of biological similarity. *Physiological Reviews* **55(4)**, 659-699.

https://doi.org/10.1152/physrev.1975.55.4.659
Hargreaves, M. and Spriet, L. L. (2020) Skeletal muscle energy metabolism during exercise. *Nature Metabolism* **2(9)**, 817-828. https://doi.org/10.1038/s42255-020-0251-4
Hawley, J. A. and Leckey, J. J. (2015) Carbohydrate dependence during prolonged, intense endurance exercise. *Sports Medicine* **45(Suppl 1)**, 5-12. https://doi.org/10.1007/s40279-015-0400-1
Helgerud, J. (1994) Maximal oxygen uptake, anaerobic threshold and running economy in women and men with similar performances level in marathons. *European Journal of Applied Physiology and Occupational Physiology* **68(2)**, 155-161. https://doi.org/10.1007/BF00244029
Hopkins, W. G. and Batterham, A. M. (2016) Error rates, decisive outcomes and publication bias with several inferential methods. *Sports Medicine* **46(10)**, 1563-1573. https://doi.org/10.1007/s40279-016-0517-x
Hopkins, W. G., Marshall, S. W., Batterham, A. M. and Hanin, J. (2009) Progressive statistics for studies in sports medicine and exercise science. *Medicine and Science in Sports and Exercise* **41(1)**, 3-13. https://doi.org/10.1249/MSS.0b013e31818cb278
Hung, K. C., Chung, H. W., Yu, C. C., Lai, H. C. and Sun, F. H. (2019) Effects of 8-week core training on core endurance and running economy. *Plos One* **14(3)**, e0213158. https://doi.org/10.1371/journal.pone.0213158
Janz, K. F., Burns, T. L., Witt, J. D. and Mahoney, L. T. (1998) Longitudinal analysis of scaling VO_2 for differences in body size during puberty: The Muscatine Study. *Medicine and Science in Sports and Exercise* **30(9)**, 1436-1444. https://doi.org/10.1249/00005768-199809000-00014
Jeukendrup, A. E. and Wallis, G. A. (2005) Measurement of substrate oxidation during exercise by means of gas exchange measurements. *International Journal of Sports Medicine* **26(Suppl 1)**, 28-37. https://doi.org/10.1055/s-2004-830512
Landis, J. R. and Koch, G. G. (1977) The measurement of observer agreement for categorical data. *Biometrics* **33(1)**, 159-174. https://doi.org/10.2307/2529310
Lee, J., Wang, Z., Chen, M., Liu, S., Yu, Q., Hu, M., Kong, Z. and Nie, J. (2024) Allometric exponents for scaling running economy in human samples: A systematic review and meta-analysis. *Heliyon* **10(10)**, e31211. https://doi.org/10.1016/j.heliyon.2024.e31211
Lee, J. and Zhang, X. L. (2021) Physiological determinants of VO_2max and the methods to evaluate it: A critical review. *Science & Sports* **36(4)**, 259-271. https://doi.org/10.1016/j.scispo.2020.11.006
Lolli, L., Batterham, A. M., Weston, K. L. and Atkinson, G. (2017) Size exponents for scaling maximal oxygen uptake in over 6500 humans: A systematic review and meta-analysis. *Sports Medicine* **47(7)**, 1405-1419. https://doi.org/10.1007/s40279-016-0655-1
Mann, T., Lamberts, R. P. and Lambert, M. I. (2013) Methods of prescribing relative exercise intensity: Physiological and practical considerations. *Sports Medicine* **43**, 613-625. https://doi.org/10.1007/s40279-013-0045-x
McMiken, D. F. and Daniels, J. T. (1976) Aerobic requirements and maximum aerobic power in treadmill and track running. *Medicine and Science in Sports* **8(1)**, 14-17. https://doi.org/10.1249/00005768-197621000-00015
Midgley, A. W., McNaughton, L. R. and Jones, A. M. (2007) Training to enhance the physiological determinants of long-distance running performance: Can valid recommendations be given to runners and coaches based on current scientific knowledge? *Sports Medicine* **37(10)**, 857-880. https://doi.org/10.2165/00007256-200737100-00003
Morgan, D. W. and Daniels, J. T. (1994) Relationship between VO_2max and the aerobic demand of running in elite distance runners. *International Journal of Sports Medicine* **15(7)**, 426-429. https://doi.org/10.1055/s-2007-1021082
Rapoport, B. I. (2010) Metabolic factors limiting performance in marathon runners. *Plos Computational Biology* **6(10)**, e1000960. https://doi.org/10.1371/journal.pcbi.1000960
Rogers, D. M., Olson, B. L. and Wilmore, J. H. (1995) Scaling for the VO_2-to-body size relationship among children and adults. *Journal of Applied Physiology* **79(3)**, 958-967. https://doi.org/10.1152/jappl.1995.79.3.958
Segers, V., De Clercq, D., Janssens, M., Bourgois, J. and Philippaerts, R. (2008) Running economy in early and late maturing youth soccer players does not differ. *British Journal of Sports Medicine* **42(4)**, 289-294. https://doi.org/10.1136/bjsm.2007.035915
Shaw, A. J., Ingham, S. A. and Folland, J. P. (2014) The valid measurement of running economy in runners. *Medicine and Science in Sports and Exercise* **46(10)**, 1968-1973. https://doi.org/10.1249/MSS.0000000000000311
Shingleton, A. W. and Frankino, W. A. (2018) The (ongoing) problem of relative growth. *Current Opinion in Insect Science* **25**, 9-19. https://doi.org/10.1016/j.cois.2017.10.001
Welsman, J. R. and Armstrong, N. (2000) Longitudinal changes in submaximal oxygen uptake in 11- to 13-year-olds. *Journal of Sports Sciences* **18(3)**, 183-189. https://doi.org/10.1080/026404100365081
Werneck, A. O., Conde, J., Coelho, E. S. M. J., Pereira, A., Costa, D. C., Martinho, D., Duarte, J. P., Valente-Dos-Santos, J., Fernandes, R. A., Batista, M. B., Ohara, D., Cyrino, E. S. and Ronque, E. R. V. (2019) Allometric scaling of aerobic fitness outputs in school-aged pubertal girls. *BMC Pediatrics* **19(1)**, 96. https://doi.org/10.1186/s12887-019-1462-2
West, G. B., Brown, J. H. and Enquist, B. J. (1997) A general model for the origin of allometric scaling laws in biology. *Science* **276(5309)**, 122-126. https://doi.org/10.1126/science.276.5309.122
Williams, T. J., Krahenbuhl, G. S. and Morgan, D. W. (1991) Mood state and running economy in moderately trained male runners. *Medicine and Science in Sports and Exercise* **23(6)**, 727-731. https://doi.org/10.1249/00005768-199106000-00012
Zakeri, I., Puyau, M. R., Adolph, A. L., Vohra, F. A. and Butte, N. F. (2006) Normalization of energy expenditure data for differences in body mass or composition in children and adolescents. *Journal of Nutrition* **136(5)**, 1371-1376. https://doi.org/10.1093/jn/136.5.1371

Key points

- While the widely used submaximal oxygen uptake (VO_{2sub}) does not directly capture metabolic consumption during exercise, it is a valid alternative to energy cost (E_c) used to quantify running economy (RE).
- Allometric scaling, rather than ratio scaling, is a more appropriate statistical adjustment for normalizing RE quantifications in interindividual comparisons.
- For studies seeking high precision, allometric-scaled E_c, which provides a more accurate and practical measure of RE, is recommended to quantify RE.
- For the convenience of allometric calculation, directly applying the 2/3 law in allometric scaling seems acceptable.

AUTHOR BIOGRAPHY

Jay LEE
Employment
UM PhD teaching research assistant.
Degree
MSc
Research interests
Sports biomechanics, Mathematics in Sports, Biomechanical Principles of Sports Training, Hypoxic exercise & training, Obesity & Cognitive Health
E-mail: 405630519@qq.com

	**Xiuli ZHANG** **Employment** Lecturer & Master's Supervisor **Degree** PhD **Research interests** Human Movement Analysis and Injury Prevention, Biomechanical Principles of Physical Training, Physical and Motor Function Training for Children and Adolescents, Exercise-Based Drug Rehabilitation and Recovery Training. **E-mail:** zhangxl@scnu.edu.cn
	**Zhaowei KONG** **Employment** Director of Centre for Arts and Physical Education **Degree** PhD **Research interests** Physiological response & adaptation to exercise, Hypoxic exercise & training, Health & fitness **E-mail:** zwkong@um.edu.mo

✉ **Prof. Xiuli Zhang, Ph.D.**
School of Physical Education & Sports Science, South China Normal University, Guangzhou, China

✉ **Accoc. Prof, Zhaowei Kong, Ph.D.**
Faculty of Education, University of Macau, Macao, China