Developing a new curvilinear allometric model to improve the fit and validity of the 20-m shuttle run test as a predictor of cardiorespiratory fitness in adults and youth

*Alan M Nevill[1], Roger Ramsbottom[2], Gavin Sandercock[3], Carlos Eduardo Bocachica-González[4], Robinson Ramírez-Vélez[5] and Grant Tomkinson[6,7].

*Corresponding Author

1. Faculty of Education, Health and Wellbeing, University of Wolverhampton, Walsall Campus, Walsall, U.K; a.m.nevill@wlv.ac.uk, http://orcid.org/0000-0003-0506-3652
2. Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK; rramsbottom@brookes.ac.uk
3. School Sport, Rehabilitation and Exercise Science, University of Essex, Colchester, U.K; gavins@essex.ac.uk, https://orcid.org/0000-0002-5443-6518
4. La Victoria. Institución Educativa Distrital. Secretaria de Educación del Distrito. Bogotá, Colombia; eduardobg22@gmail.com
5. Department of Health Sciences, Public University of Navarra, Navarrabiomed-IdiSNA, Complejo Hospitalario de Navarra (CHN), 31008 Pamplona, Spain; robin640@hotmail.com, https://orcid.org/0000-0003-3075-6960
6. Department of Education, Health and Behavior Studies, University of North Dakota, Grand Forks, ND, USA; grant.tomkinson@und.edu, https://orcid.org/0000-0001-7601-9670
7. Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health Sciences, University of South Australia, Adelaide, SA, Australia

Abstract

Background and Objectives: Doubts have been raised concerning the validity of the 20m shuttle run test (20mSRT) as a predictor of cardiorespiratory fitness (CRF) in youth based on Léger's equation/model. An alternative allometric model has been published recently that is thought to provide, not only a superior fit (criterion validity) but also a more biologically and physiologically interpretable model (construct validity). The purposes of this study were to explore whether allometry can provide a more valid predictor of CRF using 20mSRT compared with Léger's equation/model.

Methods: We fitted and compared Léger's original model and an alternative allometric model using two cross-sectional datasets (youth, n=306; adult n=105) that contained measurements of CRF ($\dot{V}O_{2peak}$ /$\dot{V}O_{2max}$) and 20mSRT performance. Quality-of-fit was assessed using explained variance (R^2) and Bland and Altman's limits of agreement.

Results: The allometric models provided superior fits for the youth (explained variance R^2=71.9%) and adult (R^2=77.7%) datasets compared with Léger's equation using their original fixed (R^2=35.2%) or re-estimated parameter models (R^2=65.9%), confirming that the allometric models demonstrate acceptable criterion validity. However, the allometric models also identified a non-linear "J-shaped" increase in energy cost ($\dot{V}O_{2peak}$/$\dot{V}O_{2max}$) with faster final shuttle-run speeds, (fitted speed exponent =1.52; 95% CI 1.38 to 1.65).

Conclusion: Not only do allometric models provide more accurate predictions of CRF ($\dot{V}O_{2peak}$/$\dot{V}O_{2max}$; $ml.kg^{-1}.min^{-1}$) for both youth and adults (evidence of criterion validity), the "J-shaped" rise in energy demand with increasing final shuttle-run speed also provides evidence of construct validity, resulting in a more plausible, physiologically sound and interpretable model.

Key words: allometry, 20-m shuttle run, prediction, aerobic fitness, cardiorespiratory fitness, youth, adults

Key Points

- Directly measured cardiorespiratory fitness (CRF) is universally acknowledged as a key index of health as well as a valuable indicator of potential endurance performance. The 20-m shuttle run test (20mSRT) is probably the most widely used field test of CRF, although recent doubts have been raised concerning its validity in youth.
- The current paper demonstrates that these doubts are unfounded. Using allometry, a new biologically and physiologically-sound model is shown to be valid (using criterion and construct validity) in both youth and adults.
- We recommend that the 20mSRT should continue to be used as an indirect measure of CRF, but the “J-shaped” rise in energy demand also provides evidence of construct validity, resulting in a more plausible, physiologically sound and interpretable model.

1. Introduction

Recent doubts have been raised concerning the validity of the 20-m shuttle run test (20mSRT) as a predictor of cardiorespiratory fitness (CRF) in youth [1]. Robust counter arguments to these doubts have also been published by Tomkinson et al. [2]. The debate between the doubters [3] and the supporters [2] of using the 20mSRT as a predictor of CRF is fuelled by two possible misconceptions/questions: 1) how should peak oxygen uptake ($\dot{V}O_{2peak}$) be adjusted or scaled for body mass to best reflect CRF, and 2) is the original Léger equation the most appropriate/valid equation to predict $\dot{V}O_{2peak}$?

The answer to the first question is relatively simple. Given the majority of physical activities are performed against resistances under normal gravitational forces, participants must carry their entire body mass (unlike some activities, e.g., cycling or swimming, that are "weight supported"). Certainly there is strong evidence that gross $\dot{V}O_{2peak}$ ($l.min^{-1}$) must be divided by the entire body mass to best reflect distance running performance in both youth [4] and adults [5]. It is the entire body mass ($ml.kg^{-1}.min^{-1}$), NOT a reduced mass such as ($ml.kg^{-0.66}.min^{-1}$) that fully reflects the detrimental effect of carrying too much body mass under normal gravitational resistance forces. The denominator of $\dot{V}O_{2peak}$ ($l.min^{-1}$) must remain the *entire* body mass, NOT an adjusted or scaled level of body mass. Welsman and Armstrong [3] may well have identified the *entire* body mass as the appropriate scaling parameter when allometrically scaling or adjusting $\dot{V}O_{2peak}$ for differences in body mass, but as far as we can determine, the authors did not report their fitted mass exponents in the manuscript.

The second question is not so simple to answer. Welsman and Armstrong [3] chose Léger's equation [6] to predict $\dot{V}O_{2peak}$ but then reported only moderate agreement (explained variance $R^2 = 0.32$) and wide Bland and Altman's [7] limits of agreement (LoA). These results might have triggered "alarm bells" that Léger's original prediction equation [6] may be the *cause* of the lack of fit, rather than the 20mSRT itself and the reason why Armstrong and Welsman [1] might have justifiably felt the need to question the validity of the 20mSRT. The results of the 20mSRT may be entirely adequate as a valid predictor of $\dot{V}O_{2peak}$, the limitation might simply be that Léger's original equation (fitted parameters) may not be the most appropriate. Indeed recently Nevill et al. [8] demonstrated that alternative allometric models were superior to linear, additive models when predicting CRF ($\dot{V}O_{2max}$; ml.kg^{-1}.min^{-1}).

Hence the purpose of the present study was to compare the validity of Léger's original equation [6] with an alternative allometric model [8], using the traditional criterion validity (i.e., cross-validation). A further aim was to assess whether the allometric model was sensitive enough to detect a non-linear, J-shaped rise in energy demand as participants perform the 20mSRT to exhaustion. This should be anticipated as shuttle running requires additional 180 degree turns involving decelerations and accelerations as well as an increase in straight-line running speed with each additional level of the 20mSRT. If detected, this nonlinear rise in energy demand ($\dot{V}O_{2max}$; ml.kg^{-1}.min^{-1}) with faster final 20mSRT speeds will provide a new insight of biological (construct) validity.

2. Methods

2.1 Design and setting

The current report is part of two cross-sectional studies. One of the studies was "The FUPRECOL Study, a school-based prospective cohort study from Colombia in schoolchildren from 9 to 18 years old" [9]. The second study is retrospective study in recreationally active adults from UK [5].

2.2 Participants (youth data)

Youth (boys, *n*= 158; girls, *n*= 148) included in the study were recruited from three Colombian schools (see Table 1 for descriptive details)

--Table 1 about here ---

Exclusion criteria were physical inability or health problems that might limit the levels of CRF. The study complied with the Declaration of Helsinki and was approved by the Ethics Committee of the Rosario University Board (Code DVO005-1-383-CEI874). Parents/guardians and school supervisors were informed by letter about the purpose, benefits, and potential risks of the study, and written informed consent was provided.

2.3 Assessments for the youth data

Anthropometric assessment included standing height (cm), which was measured with a portable stadiometer (Seca® 206, Hamburg, Germany) with a precision of 0.1 cm. Body mass (kg) was measured with a standard digital scale (Model Tanita® BC-418®, Tokyo, Japan) with an accuracy of 0.1 kg, recorded in light clothing and without shoes. All anthropometric measurements (Table 1) were taken by a single individual, who was accredited at level 2 by the International Society for the Advancement of Kinanthropometry [10].

All participants wore a portable heart rate monitor (Polar Electro Oy, Kempele, Finland). $\dot{V}O2_{peak}$ ($ml \cdot kg^{-1} \cdot min^{-1}$) was measured directly during the 20mSRT [6] using a portable oxygen and carbon dioxide analyzer (COSMED K5®, Rome, Italy), which was calibrated before each test. Respiratory parameters were recorded breath-by-breath, and averaged over a 15-s period. Measurements were taken on a normal school day (07:00–15:00 h), with participants running in comfortable clothes and shoes.

During the 20mSRT, participants ran in a straight line defined by two parallel lines on a 20-m outdoor slip-resistant court. During the test, participants were verbally motivated at each change of stage with the test terminated either volitionally, or when the participant failed to maintain the required running speed on two-consecutive laps. The running speed (km.h^{-1}) at the last completed stage was then used to predict $\dot{V}O_{2peak}$ using Léger's equation.

2.4 Participants (adult data)

Included data were collected, but not reported in a previously published article [5]. Over a 7-year period, 308 recreationally active participants (men, *n*= 179; women, *n*= 129) provided informed consent and volunteered in studies at Loughborough University which required the measurement of $\dot{V}O_{2max}$ using expired gas analysis. Experiments were conducted according to the Declaration of Helsinki for research involving human participants.

2.5 Assessments for the adult data

Upon entry to the laboratory, participants had their body mass measured using a mechanical beam balance (CMS, London) accurate to the nearest 0.1 kg, with standing height measured using a wall-mounted stadiometer accurate to the nearest 0.1 cm (Holtain Ltd., Crymych, Wales). These measures were taken with minimal clothing, typically running shorts and athletics vest or t-shirt and without shoes. Of the 308 participants, 105 (men, *n*= 56; women, *n*= 49) completed both the $\dot{V}O_{2max}$ test and 20mSRT. The 20mSRT was performed in a sports hall according to the Eurofit protocol [11], with further details available in Ramsbottom et al. [12]

2.6 Data Analyses

The model originally proposed by Léger et al. [6] to predict $\dot{V}O_{2max}$, is given by

$$\dot{V}O_{2max}\ (\text{ml.kg}^{-1}.\text{min}^{-1}) = 31.025 + 3.238 \cdot S - 3.248 \cdot A + 0.1536 \cdot A \cdot S, \qquad \text{(Eq. 1)}$$

where S=20mSRT running speed (km h^{-1}) and A=age (in years). The authors claim the model is appropriate for both boys and girls and the same equation could be used for adults keeping age constant at 18.

An alternative multiplicative model with allometric body-size components originally proposed by Nevill and Holder [13] and recently used by Nevill et al. [8] can be adapted to incorporate shuttle-run speed as follows,

$$\dot{V}O_{2max}\ (ml.kg^{-1}.min^{-1}) = M^{k_1} \cdot H^{k_2} \cdot S^{k_3} \cdot \exp(a + b_1 \cdot Age), \qquad \text{(Eq. 2)}$$

where M=Mass, H=Height and S=20mSRT running speed (km h^{-1}). The model (Eq. 2) can be linearized with a log transformation (Ln=$\log_e$). A linear regression analysis or ANCOVA on Ln($\dot{V}O_{2max}$) can then be used to estimate the unknown parameters in the log-transformed model i.e., the transformed model (Eq. 3) is now additive and conforms with the assumptions associated with ordinary least-squares and ANCOVA:

$$Ln(\dot{V}O_{2max}) = k_1 \cdot Ln(M) + k_2 \cdot Ln(H) + k_3 \cdot Ln(S) + a + b_1.Age, \qquad \text{(Eq. 3)}$$

where the intercept “a” is allowed to vary with categorical variables, e.g., sex. This model can be refined (to obtain a parsimonious solution) using backward elimination [14].

2.6.1 Cross-validation using the youth data

To assess the validity or success of the allometric model, we adopted the same cross-validation adopted by Nevill et al. [8] by splitting the data into two independent groups using a random split (80:20). We used the 80% sample to predict $\dot{V}O_{2max}$ (ml.kg^{-1}.min^{-1}) using Eq. 2 (parameters estimated using Eq. 3) and the 20% sample to test/validate the model. This was achieved by predicting the $\dot{V}O_{2max}$ of the 20% sample using the model derived from the 80% prediction model. The success of cross-validation process was then assessed by comparing the *measured* $\dot{V}O_{2max}$ of 20% validation sample with the *predicted* $\dot{V}O_{2max}$ scores using R^2, correlations and Bland and Altman's LoA.

2.6.2 Cross-validation using the adult data

We further cross-validated the multiplicative, allometric model by fitting the parsimonious model to the adult data. The success of the allometric model was assessed by computing the level of the explained variance (R^2) and by comparing the parameters with 95% confidence intervals (CI) obtained in the adult model with those obtained in the youth model.

3 Results

The agreement between youth predicted $\dot{V}O_{2max}$, based on Léger's equation (Eq. 1), and the measured $\dot{V}O_{2max}$ resulted in R^2= 35.2% (r= 0.593), CV=17.0% and a 95% LoA= -3.57 ± 14.05 ($ml.kg^{-1}.min^{-1}$).

When we allowed the parameters of Léger's equation 1 to be estimated (rather than being fixed by Léger's original equation) using the youth data, the fitted model was

$$\dot{V}O_{2max}\ (ml.kg^{-1}.min^{-1}) = -15.49 + 6.958 \cdot S - 1.812 \cdot A + 0.114 \cdot A \cdot S, \qquad \text{(Eq. 1a)}$$

that explained R^2= 65.9% (r=0.81 with a SEE=5.2 $ml.kg^{-1}.min^{-1}$) of the variance. Although the equation provides a statistically/mathematically reasonable fit, it is not a real/biologically sound model. For example the equation would predict negative $\dot{V}O_{2max}$ scores for small speeds and ages. Also, because this model was fitted using the youth measured $\dot{V}O_{2max}$, reporting the agreement with this predicted model (Eq. 1a) and the measured $\dot{V}O_{2max}$ using the CV and 95% LoA would NOT be independent and hence misleading.

The estimated parameters β (± standard error SE) and 95% CIs obtained from fitting the alternative log-transformed multiplicative, allometric model (Eq. 3) to the youth data, are given in Eq. 4 and Table 2

$$\text{Ln}(\dot{V}O_{2peak}) = -0.126 \cdot \text{Ln}(M) + 1.53 \cdot \text{Ln}(S) + 0.808 - 0.116 \cdot \text{girls} \qquad \text{(Eq. 4)}$$

where M=Mass and S=20mSRT running speed (km h^{-1}) and “girls” is entered as a [0,1] indicator variable (boys=0 and girls=1). Note that H=height and A=age failed to make a significant contribution to the final, parsimonious model but the girls were predicted to have significantly lower Ln($\dot{V}O_{2peak}$) (β= -0.116, SE= 0.014; t= 8.33, P<0.001) compared to the boys.

--Table 2 about here ---

The log-transformed model explained R^2= 69.6% (r= 0.83) of the variance with the standard error of estimate SEE= 0.11 (or 11.6%).

Taking antilogs, the model becomes

$$\dot{V}O_{2peak}\ (\text{ml.kg}^{-1}.\text{min}^{-1}) = M^{-0.126} \cdot S^{1.531} \cdot \exp(0.808 - 0.116 \cdot \text{girls}) \qquad \text{(Eq. 5)}$$

3.1 Cross-validation of the model

A 20% sample of *n*= 73 (i.e., 73/306) youth were randomly selected (using the SPSS software) to validate the model derived using the remaining 80% (*n*= 233) of youth. The results from cross-validation (using the 20% validation sample) indicated that the correlation between predicted $\dot{V}O_{2max}$ (using Eq. 5 fitted to the 80% data) and the measured $\dot{V}O_{2peak}$ was r= 0.85. No significant bias (measured $\dot{V}O_{2peak}$ mean= 44.8 – predicted $\dot{V}O_{2peak}$ mean= 44.1 [ml.kg^{-1}.min^{-1}]) was found using a paired samples t-test (t_{72}= 1.46, P= 0.15) and the explained variance was R^2= 71.9%, with a CV= 9.6% and a 95% LoA 0.73 ± 8.3 (ml.kg^{-1}.min^{-1}). When we plotted measured $\dot{V}O_{2peak}$ vs. predicted $\dot{V}O_{2peak}$ (ml.kg^{-1}.min^{-1}) in Figure 1, we found evidence of heteroscedasticity confirming the need to report the agreement as the ratio limits of agreement and expressing the standard deviation of differences as a ratio or as a coefficient of variation (CV) [15].

--Figure 1 about here ---

3.2 Adult data

The log-transformed multiplicative, allometric model (Eq. 5) was fitted to the adult data resulting in the following estimated parameters β (±SE) and 95% CI in Table 3. The model can be written as follows;

$$\mathrm{Ln}(\dot{V}O_{2peak}) = -0.178 \cdot \mathrm{Ln}(M) + 1.562 \cdot \mathrm{Ln}(S) + 0.718 - 0.071 \cdot \mathrm{female},$$

or $$\dot{V}O_{2peak}\ (\mathrm{ml.kg^{-1}.min^{-1}}) = M^{-0.178} \cdot S^{1.562} \cdot \exp(0.718 - 0.071 \cdot \mathrm{female})$$

after taken antilogs.

--Table 3 about here ---

The log-transformed model explained R^2= 77.7% of the variance with SEE= 0.082 (or 8.5%). Comparing the log-transformed parameters (with the help of 95% CI) from the youth and adult models in Tables 2 and 3, we get very similar body-weight and shuttle-run speed exponents. In particular the speed exponents were (1.53 [95% CI 1.37 to 1.69]) for youth and (1.56 [95% CI 1.25 to 1.87]) for adults. Note that the 95% CIs of both speed exponents exclude unity, confirming that both models follow a “J” shape curve that precludes a straight-line fit. This curvature can be seen in Figure 2.

--Figure 2 about here ---

3.3 Combined youth and adult data

The estimated parameters β (±SE) and 95% CI obtained from fitting the log-transformed multiplicative, allometric model (Eq. 3) to the combined adults and youth data, are given in Table 4 and Eq. 6

--Table 4 about here ---

$$\mathrm{Ln}(\dot{V}O_{2max}) = -0.133 \cdot \mathrm{Ln}(M) + 1.519 \cdot \mathrm{Ln}(S) + 0.854 - 0.104 \cdot female - 0.16 \cdot adult \quad \text{(Eq. 6)}$$

where “female” and “adult” are entered as [0,1] indicator variables (male=0 and female=1 and child=0 and adult=1 respectively). Note that the female parameter estimating $\mathrm{Ln}(\dot{V}O_{2max})$ was significantly lower (β=-0.104, SE=0.012; t=8.66, P<0.001) than the male parameter. The log-transformed model explained R^2=77.0% of the variance with SEE= 0.104 (or 11%).

Taking antilogs, the model becomes

$$\dot{V}O_{2max}\ (ml.kg^{-1}.min^{-1}) = M^{-0.133} \cdot S^{1.519} \cdot \exp(0.854 - 0.104 \cdot female - 0.16 \cdot adult) \quad \text{(Eq. 7)}$$

with the baseline group being male youth (boys).

4. Discussion

4.1 Criterion validity

Our initial analysis revealed the association between youth predicted $\dot{V}O_{2peak}$ based on Léger’s equation [6], and the measured $\dot{V}O_{2peak}$ resulted in moderate agreement with R^2= 35.2%, CV=17.0% and a 95% LoA -3.57±14.05 ($ml.kg^{-1}.min^{-1}$). The explained variance was a little higher, but the limits of agreement were wider than those reported by Welsman and Armstrong [3] (R^2= 0.32 with 95% LoA of −9.1 to +11.9 $ml.kg^{-1}.min^{-1}$). Nevertheless, both studies confirm that Léger’s equation [6] demonstrates only moderate validity when predicting $\dot{V}O_{2peak}$. When we allowed the parameters of Léger’s equation [6] to be re-estimated (rather than fixed from their original model), the model explained R^2= 65.9% of the variance, suggesting that the model itself may be suitable, but it is their original parameters of the Léger’s equation [6] that may be the cause the “moderate validity”. This might explain why Armstrong and Welsman [1] felt justified in reporting that the 20mSRT might be an invalid predictor of CRF, but the lack of validity appears to be based on Léger's original equation NOT the 20mSRT itself.

When we fitted the alternative log-transformed multiplicative, allometric model (Eq. 3) to the youth data, the model explained more of the variance R^2= 69.6%. The results from the cross-validation assessment (using the 20% sample) found that the explained variance between the predicted and measured $\dot{V}O_{2peak}$ was very high R^2= 71.9%, with a CV= 9.55% and a 95% LoA 0.73 ± 8.3 ($ml.kg^{-1}.min^{-1}$) (see Figure 1). Further evidence that the allometric model is a valid predictor of CRF comes from fitting the model (Eq. 4) to the adult data reported in Table 3 (further cross validation). This model explained a very high percentage of the variance R^2= 77.7% and reassuringly the speed and mass exponent parameters were remarkably similar to those reported for youth in Table 2. The only discrepancy appears to be that the adult's intercept term appears lower than the youth parameter, confirmed in Table 4, with the adults' intercept being -0.16 (95% CI -.20 to -0.12) lower that the youth intercept term. This suggests that for the same shuttle-running speed and mass, the youth are running at a higher energy cost, possibly due to youth having a higher stride frequency, as well as inconsistent stride-to-stride patterning which may reflect an immature neuromotor control [16].

All three allometric models (youth, adult and combined) suggest that for the same final 20mSRT running speed and body mass, the female energy cost is significantly lower than males (Youth β= -0.116, SE= 0.014; t=-8.33, P<0.001; and combined sample β= -0.104, SE= 0.012; t= -8.66, P<0.001). It is inconceivable that these sex differences (eight standard deviations below the mean, see t-scores) could have happen by chance. There must be a physiological or biological mechanism to explain such differences. Likely explanations might include the fact that girls tend to have lower levels of lean muscle mass and higher levels of fat mass compared with boys [17]. The pubertal increase in fat mass might also help explain why body mass has been shown to increase proportionally more than $\dot{V}O_{2peak}$ in girls but not boys [18]. It has also been suggested that during intense exercise (and possibly submaximal intensities), women demonstrate pulmonary limitations (e.g., greater expiratory flow limitation, an increased work of breathing, and perhaps greater exercise induced arterial hypoxemia) compared to men [19]. This is thought to be due to the influence of the reproductive hormones (progesterone and estrogen) and pulmonary structural differences leading to reduced pulmonary capacity.

Based on the cross-validation assessment using the youth data reinforced with the adult data, the allometric models incorporating shuttle speed and mass (Eq. 5 and Eq. 7) appear to be valid models to predict $\dot{V}O_{2peak}$/$\dot{V}O_{2max}$ in both youth and adults respectively (criterion validity).

4.2 Construct validity

However, probably the most valuable new insight obtained from adopting these allometric models comes from their biological interpretation, in particular by observing how $\dot{V}O_{2peak}$/$\dot{V}O_{2max}$ increases with increasing 20mSRT speeds (Figure 2). Table 4 reveals the common shuttle-running speed exponent was 1.52 (95% CI; 1.38 to 1.65). This curvilinear response is not entirely unexpected as, albeit at lower walking speeds, a similar association was observed during the incremental 10m shuttle-walking test [20]. This “J” shaped curve confirms that a non-linear rise (precluding a straight line) in energy demand is associated with participants’ final speed at the final stages of the 20mSRT. Given that shuttle running, unlike treadmill exercise or straight-line running, requires additional 180 degree turns (involving decelerations and accelerations) as well as an increase in speed with each additional level of the 20mSRT, this requirement will elicit greater additional and accumulative energy demands as the test progresses to exhaustion. These increases in energy demand will almost certainly explain much of the J-shaped curve identified by the allometric models (Eq. 5 and 7). This insight/interpretation seems both biologically plausible and physiologically sound, providing additional evidence of “construct validity” for the allometric models (Eqs. 5 and 7).

Further evidence of “construct validity” comes from the negative mass exponents reported in the youth (-0.126), adult (-0.178) and the combined allometric models (-0.133) (see Tables 2, 3 and 4). Åstrand and Rodahl [21] in their Figure 9-4 page 400, reported a strong negative association between $\dot{V}O_{2max}$ (ml.kg^{-1}.min^{-1}) and body mass (r= -0.69, P<0.001). Nevill et al. [5] report similar negative associations r= -0.39 (P<0.01; n= 179) and r= -0.35 (P<0.01; n= 129) for men and women respectively. This is because absolute $\dot{V}O_{2max}$ (l.min^{-1}) scales or is associated with body mass ($M^{0.67}$). Thus, when we calculate $\dot{V}O_{2max}$ (ml.kg^{-1}.min^{-1}) by dividing $\dot{V}O_{2max}$ (l.min^{-1}) by body mass (M), the resulting ratio “over-scales”, leaving $\dot{V}O_{2max}$ (ml.kg^{-1}.min^{-1}) proportional to $M^{-0.33}$. It would appear that the introduction of shuttle-run speed into the model explains some, but not all of this negative resistance or drag associated with carrying excess mass.

A limitation of our study is the smaller 20% (i.e.: cross-validation) sample, which means that the results could be confirmed in a larger population (despite the adult data providing further independent “cross-validation” support). Additionally, participants may have arrived at volitional fatigue before achieving their “true” maximal capacity during the field/lab-based measurement. Lastly, all research in which a single measurement is used as a “gold standard” is susceptible to random error, and in this case, it is impossible to know how that error influenced the estimates of bias.

5. Implications

Probably the most valuable new insight obtained from adopting these allometric models comes from their biological interpretation, in particular by observing how $\dot{V}O_{2peak}$ / $\dot{V}O_{2max}$ increases with increasing 20mSRT running speed. The shuttle-running speed exponent was 1.52 (95% CI; 1.38 to 1.65), confirming a “J” shaped, non-linear rise (precluding a straight line) in energy demand as participants perform the test to exhaustion. Given that shuttle running, unlike treadmill exercise or straight-line running, requires additional 180 degree turns (involving decelerations and accelerations) as well as an increase in speed with each additional level of the 20mSRT, this requirement will elicit greater additional and accumulative energy demands as the test progresses to exhaustion. This interpretation seems both biologically plausible and physiologically sound, providing additional evidence of “construct validity” for the allometric model.

6. Conclusion

The allometric models (Eq. 5 and 7) provide more accurate and valid predictions of $\dot{V}O_{2peak}$ and $\dot{V}O_{2max}$ (ml.kg^{-1}.min^{-1}) for both youth and adults respectively. The explained variance (R^2) and LoA were superior (the former being greater and the latter being narrower) to those found using Léger’s prediction models (using either the original or re-estimated parameters). Further support for the allometric model was obtained based on construct validity, where the fitted speed parameter (s) were able to detect a curvilinear “J-shaped” change in energy demand with faster final 20mSRT speeds, providing a more plausible, biologically sound and interpretable model.

Ethics declarations

Acknowledgments: The authors gratefully acknowledge the contributions of postgraduate students (Jose Nelson Gil, Daniel Prieto and Alejandra Hernandez) that carried out the implementation and data collection, and Jorge Enrique Correa-Bautista, PhD from CEMA-UR for excellent technical support for facilitating this study (Colombia Study). This article is part of the master thesis submitted by CEB-G, from the University of Rosario, Colombia.

Contributions: AMN analyzed the data and drafted the manuscript. RR, CEB-G, RR-V and GS designed the study, directed implementation and data collection. AMN, RR and GT edited the manuscript for intellectual content and provided critical comments on the manuscript.

Funding: No financial support was received for the conduct of this study or preparation of this article.

Competing interests: Alan Nevill, Roger Ramsbottom, Gavin Sandercock, Carlos Eduardo Bocachica-González, Robinson Ramírez-Vélez, and Grant Tomkinson have no conflicts of interest that are directly relevant to the content of this article.

Ethics approval: The study was approved by the Ethics Committee of the Rosario University Board (Code DVO005-1-383-CEI874).

Informed consent: Informed consent was obtained from all participating teams in the study.

Patient consent for publication: Not required.

Data availability statement: Data are available on request

References

1. Armstrong N, Welsman J Twenty-metre shuttle run: (mis)representation, (mis)interpretation and (mis)use. Br J Sports Med. 2019;53:1199. doi:10.1136/bjsports-2018-100082
2. Tomkinson GR, Lang JJ, Léger LA, Olds TS, Ortega FB, Ruiz JR, et al. Response to criticisms of the 20 m shuttle run test: deflections, distortions and distractions. Br J Sports Med. 2019;53:1200–01. doi:10.1136/bjsports-2018-100348
3. Welsman J, Armstrong N. The 20 m shuttle run is not a valid test of cardiorespiratory fitness in boys aged 11-14 years. BMJ Open Sport Exerc Med. 2019;5:e000627. doi:10.1136/bmjsem-2019-000627
4. Nevill A, Rowland T, Goff D, Martel L, Ferrone L. Scaling or normalising maximum oxygen uptake to predict 1-mile run time in boys. Eur J Appl Physiol. 2004;92:285–8. doi:10.1007/s00421-004-1071-z
5. 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
6. Léger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. J Sports Sci. 1988;6:93–101. doi:10.1080/02640418808729800
7. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8:135–60. doi:10.1177/096228029900800204
8. Nevill AM, Myers J, Kaminsky LA, Arena R. Improving reference equations for cardiorespiratory fitness using multiplicative allometric rather than additive linear models: Data from the Fitness Registry and the Importance of Exercise National Database Registry. Prog Cardiovasc Dis. 2019;62:515–21. doi:10.1016/j.pcad.2019.11.011

9. Ramírez-Vélez R, Rodrigues-Bezerra D, Correa-Bautista JE, Izquierdo M, Lobelo F. Reliability of Health-Related Physical Fitness Tests among Colombian Children and Adolescents: The FUPRECOL Study. PLoS ONE. 2015;10(10): e0140875. doi: 10.1371/journal.pone.0140875
10. Marfell-Jones M, Olds T, Stewart A. International Standards for Anthropometric Assessment; International Society for the Advancement of Kinanthropometry: Potchefstroom, South Africa, 2011; ISBN 0620362073.
11. Council of Europe. Eurofit: handbook for the Eurofit tests of physical fitness. Rome: Council of Europe, 1988
12. Ramsbottom R, Brewer J, Williams C. A progressive shuttle run test to estimate maximal oxygen uptake. Br J Sports Med. 1988;22:141–4. doi:10.1136/bjsm.22.4.141
13. Nevill A, Holder R. Modelling Maximum Oxygen Uptake, A Case-Study in Non-Linear Regression Model Formulation and Comparison. J R Stat Soc Ser C Appl Stat. 1994;43:653-66. doi:10.2307/2986263
14. Draper N, Smith H. Applied Regression Analysis (Wiley Series in Probability and Statistics). Hardcover. John Wiley. New York, 1981. ISBN 10: 0471170828
15. Nevill AM, Atkinson G. Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science. Br J Sports Med. 1997;31:314–8. doi:10.1136/bjsm.31.4.314
16. Frost G, Bar-Or O, Dowling J, Dyson K. Explaining differences in the metabolic cost and efficiency of treadmill locomotion in children. J Sports Sci. 2002;20:451–461. doi:10.1080/02640410252925125
17. Wells JC. Sexual dimorphism of body composition. Best Pract Res Clin Endocrinol Metab. 2007;21(3):415-430. doi:10.1016/j.beem.2007.04.007

18. Pettersen SA, Fredriksen PM, Ingjer E. The correlation between peak oxygen uptake (VO2peak) and running performance in children and adolescents. aspects of different units. Scand J Med Sci Sports. 2001;11(4):223-228. doi:10.1034/j.1600-0838.2001.110405.x
19. Harms CA. Does gender affect pulmonary function and exercise capacity?. Respir Physiol Neurobiol. 2006;151(2-3):124-131. doi:10.1016/j.resp.2005.10.010
20. Buckley JP, Cardoso FMF, Birkett ST, Sandercock GRH. Oxygen Costs of the Incremental Shuttle Walk Test in Cardiac Rehabilitation Participants: An Historical and Contemporary Analysis. Sports Med. 2016;46:1953-1962. doi:10.1007/s40279-016-0521-1
21. Åstrand P-O, Rodahl K. Textbook of work physiology, 3rd ed. New York: McGraw-Hill. 1986. doi:10.2310/6640.2004.00030

Table 1. Anthropometric and performance data for the youth recruited from three Colombian schools

	Age	n	Height (cm)	sd	Mass (kg)	sd	BMI	sd	$\dot{V}O_{2peak}$ (ml.kg.min^{-1})	sd	Speed (km.h^{-1})	sd
Boys	9	14	129.93	6.50	31.26	6.65	18.39	3.07	43.22	5.54	9.04	0.37
	10	16	135.94	3.51	34.11	6.62	18.39	3.16	44.06	5.49	9.28	0.55
	11	22	140.50	6.12	36.89	7.28	18.56	2.61	46.18	8.98	9.89	0.80
	12	14	149.64	8.18	42.94	9.78	18.95	2.92	46.87	6.57	9.93	0.47
	13	15	155.73	9.85	48.39	10.21	19.81	2.81	45.52	7.25	9.77	0.84
	14	27	161.22	6.29	52.11	9.86	19.94	2.67	50.08	7.49	10.46	0.66
	15	16	164.81	5.11	59.43	7.38	21.91	2.73	47.90	5.93	10.38	0.83
	16	21	167.90	6.33	56.34	7.81	19.91	1.84	59.53	4.68	11.55	0.77
	17	13	168.46	7.25	58.94	4.17	20.78	1.35	54.01	8.95	11.00	0.94
	Total	158	153.30	14.77	47.02	12.70	19.62	2.77	48.96	8.45	10.20	1.02
Girls	9	20	132.20	6.08	32.13	6.45	18.25	2.43	35.08	4.63	8.88	0.43
	10	15	137.33	6.34	32.11	3.86	17.07	2.01	41.25	5.00	9.10	0.47
	11	21	141.90	5.87	36.69	5.03	18.20	2.00	39.70	4.91	9.33	0.48
	12	13	148.15	5.35	43.36	7.23	19.75	2.98	42.01	6.21	9.50	0.58
	13	18	151.39	5.55	48.56	7.86	21.24	3.62	36.60	4.95	9.17	0.42
	14	15	152.93	7.40	49.34	8.73	21.11	3.39	40.16	6.03	9.77	0.59
	15	16	158.31	5.76	53.52	7.11	21.34	2.38	40.35	6.50	9.84	0.70
	16	15	155.53	5.84	52.99	5.11	21.91	1.95	39.29	5.30	9.77	0.59
	17	15	157.80	4.87	55.27	7.44	22.13	2.30	36.88	6.56	9.63	0.52
	Total	148	147.72	10.73	44.27	10.93	20.01	3.10	38.85	5.82	9.42	0.61

Table 2. Estimated parameters β (±SE) and 95% CI from fitting the allometric model (Eq. 3) to the youth data

Parameter	β	SE	P	95% Confidence Interval	
				Lower Bound	Upper Bound
Intercept (boys)	0.808	0.177	<0.001	0.46	1.156
LnS	1.531	0.081	<0.001	1.372	1.689
lnMass	-0.126	0.025	<0.001	-0.175	-0.077
Girls (Δ)	-0.116	0.014	<0.001	-0.143	-0.088

SE=standard error. The baseline group were taken as the boys, from which the girl's intercept parameter was estimated (Δ).

Table 3. Estimated parameters β (±SE) and 95% confidence intervals from fitting the allometric model (Eq. 5) to the adult data

Parameter	β	SE	P	95% Confidence Interval	
				Lower Bound	Upper Bound
Intercept (male)	0.718	0.547	0.193	-0.368	1.804
LnS	1.562	0.157	<0.001	1.251	1.874
lnMass	-0.178	0.063	0.006	-0.304	-0.052
Female (Δ)	-0.071	0.028	0.013	-0.127	-0.015

SE=standard error. The baseline group was taken as the male adults, from which the female adults intercept parameter was estimated (Δ).

Table 4. Estimated parameters β (±SE) and 95% confidence intervals from fitting the allometric model (Eq. 3) to the combined youth and adult data

Parameter	β	SE	P	95% Confidence Interval	
				Lower Bound	Upper Bound
Intercept (male)	0.854	0.157	<0.001	0.545	1.163
Female (Δ)	-0.104	0.012	<0.001	-0.127	-0.081
Adult (Δ)	-0.160	0.022	<0.001	-0.204	-0.116
LnS	1.519	0.069	<0.001	1.383	1.655
lnMass	-0.133	0.022	<0.001	-0.176	-0.089

SE=standard error. The baseline group was taken as the male youth, from which the female and adult intercept parameters were estimated (Δ).