Size-exponents for scaling maximal oxygen uptake in over 6500 humans: a systematic review and meta-analysis

Heading title: Allometric scaling of $\dot{V}O_{2max}$

Authors: Lorenzo Lolli[1], Alan M Batterham[1], Kathryn L Weston[1], Greg Atkinson[1]

[1] Health and Social Care Institute, School of Health and Social Care, Teesside University, Middlesbrough, UK

Corresponding Author:

Lorenzo Lolli (L.Lolli@tees.ac.uk, +44 (0) 1642 342934)
Teesside University, Health and Social Care Institute
School of Health and Social Care
Constantine Building, Middlesbrough, TS1 3BA, United Kingdom

Key points

Both maximal oxygen uptake ($\dot{V}O_{2max}$) and body composition are predictive of cardiovascular disease and mortality. $\dot{V}O_{2max}$ is typically scaled to body size but there is disagreement over the value of the size exponent for both of the most commonly-studied variables of body mass and fat-free mass.

In this first formal quantitative synthesis of human studies, the fat-free mass exponent was 0.90 (95% prediction interval: 0.68 to 1.12) and found to be independent of sex.

The probability that the fat-free mass exponent in a future study would be higher than the ⅔- and ¾-power laws was estimated to be 0.98 (very likely) and 0.92 (likely), respectively.

Abstract

Background: Maximal oxygen uptake ($\dot{V}O_{2max}$) is conventionally normalized to body size as a simple ratio or using an allometric exponent < 1. Nevertheless, the most appropriate body size variable to use for scaling and the value of the exponent are still enigmatic. Studies tend to be based on small samples and can, therefore, lack precision.

Objective: To provide a quantitative synthesis of reported static allometric exponents used for scaling $\dot{V}O_{2max}$ to whole body mass and fat-free mass.

Methods: Eight electronic databases (CINAHL, Cochrane Central Register of Controlled Trials, EMBASE, MEDLINE, PubMed, Scopus, SPORTDiscus and Web of Science) were searched for relevant studies published up to January 2016. Search terms included 'oxygen uptake', 'cardiorespiratory fitness', '$\dot{V}$O2max', '$\dot{V}$O2peak', 'scaling' and all interchangeable terms. Inclusion criteria included human cardiorespiratory fitness data; cross-sectional study designs; an empirical derivation of the exponent; reported precision statistics; and reported information regarding participant sex, age and sports background, $\dot{V}O_{2max}$ protocol, body composition protocol and line-fitting methods. A random-effects model was used to quantify weighted pooled exponents and 95% confidence limits (95% CL). Heterogeneity was quantified with the tau-statistic (τ). Meta-regression was used to quantify the impact of selected moderator variables on the exponent effect size. A 95% prediction interval (95% PI) was calculated to quantify the likely range of true fat-free mass exponents in similar future studies, with this distribution used to estimate the probability that an exponent would be above theorised universal values of ⅔ and ¾.

Results: Thirty-six studies, involving 6,514 participants, met the eligibility criteria. Whole-body mass and fat-free mass was used as the scaling denominator in 27 studies and 15 studies, respectively. The pooled allometric exponent (95% CL) was found to be 0.70 (0.64 to 0.76) for body mass and 0.90 (0.83 to 0.96) for fat-free mass. The between-study heterogeneity was greater for whole-body mass (τ = ± 0.15) vs. fat-free mass (τ = ± 0.11). Participant sex explained 30% of the between-study variability in the whole-body mass exponent, but the influence on the fat-free mass exponent was trivial. The body mass exponent of 0.52 (0.40 to 0.64) for females was substantially lower than the 0.76 (0.70 to 0.83) for males, whereas the fat-free mass exponent was similar for both sexes. The effects of all other moderators were trivial. The 95% PI for fat-free mass ranged from 0.68 to 1.12. The estimated probability of a true fat-free mass exponent in a future study being greater than ⅔- or ¾-power scaling is 0.98 (very likely) and 0.92 (likely), respectively.

Conclusions: In this quantitative synthesis of published studies involving over 6,500 humans, the body mass exponent was found to be spuriously low and prone to substantial heterogeneity. We conclude that the scaling of $\dot{V}O_{2max}$ in humans is consistent with the allometric cascade model with an estimated prediction interval for the fat-free mass exponent not likely to be consistent with the ⅔ and ¾ power laws.

1 Introduction

Maximal oxygen uptake ($\dot{V}O_{2max}$), sometimes referred to as "peak" oxygen uptake ($\dot{V}O_{2peak}$) in young [1] and pathophysiological populations [2], is a well-established indicator of cardio-metabolic health and a strong prognostic factor for all-cause mortality [3]. $\dot{V}O_{2max}$ is also a key indicator of top-class sports performance, especially in cross-country skiers, endurance cyclists, marathon runners and football (soccer) players [4-7]. $\dot{V}O_{2max}$ is typically expressed in absolute values or as a simple ratio to body size. In recent systematic reviews and meta-analyses, the effects of training interventions on $\dot{V}O_{2max}$, either in a healthy, clinical or sports population, were quantified in absolute terms or per kg of body weight [8-11]. This ratiometric index is statistically robust only if there is a direct proportional relationship between numerator and denominator, an underlying assumption that is seldom confirmed in nature [12]. Therefore, the inherent theoretical, mathematical and empirical flaws of ratios render them unsuitable to normalize physiological data [13]. These shortcomings of ratios have led many scientists to consider alternative approaches based on power-law constructs in an attempt to partition out the confounding effects of body size [14-17]. Nevertheless, indices of $\dot{V}O_{2max}$ normalized to power law-based exponents could yet display a residual size correlation [18, 19]. Allometric scaling approaches are accurate only if the correlation between the normalized index and the size variable approaches zero [20].

Whole-body mass is the conventional scaling denominator of $\dot{V}O_{2max}$, particularly in interspecific allometry studies. Interspecific allometry refers to differences in morphological, physiological or ecological traits between different species when measured at the same growth stage, whereas intraspecific allometry defines how the traits of individuals within a single species alter at, for example, different ages [21]. When the allometric exponents are quantified in human studies, values appear to fall into one of two "camps" corresponding to the ⅔ or ¾ power scaling laws [22-25]. The majority of these studies also tend toward small sample sizes ($n < 60$), resulting in relatively imprecise estimates of the exponent due to sampling error [26]. As such, the spectrum of findings for the whole-body mass exponent limits the definition of the exact power function for the scaling of $\dot{V}O_{2max}$ in human samples. The current uncertainty might also reflect the lack of information on relevant covariates in the model. The absence of intra-species variation, quantified as differences in $\dot{V}O_{2max}$, age, sex, body composition, is somewhat unrealistic [27]. While in interspecific allometry the large body size range overwhelms between-individual variability in other prognostic variables, careful model adjustment (e.g. for body composition) is essential in intraspecific studies that inevitably have a curtailed size range [28].

The fundamental validity of scaling human energy metabolism to whole-body mass was itself challenged more than half-century ago [29, 30]. Indeed, fat-free mass, on physiological grounds, is arguably a more appropriate scaling denominator for $\dot{V}O_{2max}$, as over 90% of the oxygen passing through the lungs of an exercising mammal is destined for a single sink in the skeletal muscle mitochondria [23, 31]. Some scientists have reported a substantial difference in the allometric exponents for whole-body mass and fat-free mass, with the latter being observed to be closer to unity [23, 32]. Nevertheless, most of these studies, again, involve relatively small sample sizes, which lead to wide confidence limits for each exponent estimate [33-37].

To date, there is no published quantitative synthesis of all the derived size-exponents to clarify if

the relationship between $\dot{V}O_{2max}$ and body size equates to a universal scaling exponent within a large, heterogeneous human sample of males and females. To improve statistical precision and ultimately help to resolve some inconsistencies in the literature, we aim to provide a quantitative synthesis of derived static allometric exponents for the scaling of $\dot{V}O_{2max}$ to whole-body mass and fat-free mass.

2 Methods

2.1 Study eligibility

2.1.1 Inclusion criteria

The primary outcome measure was the estimated static allometric exponent (*b*) for whole-body mass and fat-free mass to scale $\dot{V}O_{2max}$. To be selected for quantitative synthesis, studies had to meet all of the following criteria: *i)* $\dot{V}O_{2max}$ data were gathered directly in human samples during maximal cardiopulmonary exercise testing; *ii)* the investigation had a cross-sectional design; *iii)* the allometric exponent was originally derived from the examined sample; *iv)* either confidence limits (CL), standard deviation (SD), standard error (SE) or Pearson's correlation coefficient (r) for the allometric model were reported; *v)* details relevant to the identified moderator variables were outlined. The target population had no restriction.

2.1.2 Exclusion criteria

Studies were excluded if: *i)* ontogenetic allometry was the primary analysis*; ii)* static exponents were derived from a longitudinal or mixed-longitudinal study; *iii)* power function scaling was not performed; *iv)* power law-based exponents were adopted a priori rather than being derived empirically; *v*) sample and results matched a previous publication (duplication); *vi)* the full-text manuscript was written in languages other than English.

2.2 Literature search and study selection

Electronic searching was performed across eight databases (CINAHL, Cochrane Central Register of Controlled Trials, EMBASE, MEDLINE, PubMed, Scopus, SPORTDiscus and Web of Science) and conducted by two of the authors (LL, GA) from October 2015 to January 2016. Search terms included "oxygen uptake" OR "oxygen consumption" OR "aerobic power" OR "aerobic fitness" OR "cardiorespiratory fitness" OR "cardio?respiratory fitness" OR "V̇O2peak" OR "peakV̇O2" OR "V̇O2max" OR "maxV̇O2" OR "V̇O2?peak" OR "peak?V̇O2" OR "V̇O2?max" OR "max?V̇O2" OR "cardiorespiratory function" OR "cardio?respiratory function" OR "exercise capacit*" OR "physical fitness" OR "functional capacit*" OR "exercise performance*" AND "allomet*" OR "exponent" OR "exponents" OR "scaling" OR "scaled" either singly or in combination. Additional records were obtained from the reference lists of the retrieved articles or by hand searching. Only investigations conducted in

humans were considered. All titles and abstracts were initially scrutinized to exclude studies that were beyond the aim of this meta-analysis. The remaining full-text-articles that met each of the eligibility criteria were then included in quantitative synthesis. Disagreements were resolved with the aid of a third reviewer (KLW). A complete overview of the process is illustrated in Fig. 1. The systematic literature review was conducted in compliance with the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-analyses) statement [38].

2.3 Data extraction

DigitizeIt (Braunschweig, Germany), graph digitizer software, was used in the cases where only plots were available for descriptive data. Study characteristics, pooled value of the allometric exponent and respective statistical precision measures were extracted by two authors (LL, GA). If precision measures were not reported, standard errors were derived via the relationship between the standard deviation of the primary outcome (i.e. $\dot{V}O_{2max}$ expressed in $L \cdot min^{-1}$) and Pearson's correlation coefficient for the allometric model. Data relevant to age category (i.e. cut-off point for adults: mean age ≥ 18 years), sports background, sex, physical testing, body composition assessment method for determining fat-free body mass (i.e. greater precision means: dual-energy X-ray absorptiometry; lower precision means: bioelectrical impedance analysis), and line-fitting methods were obtained to evaluate the potential effects of the identified moderator variables on the global allometric exponent. Disagreements over the accuracy and comprehensiveness of the extracted data were resolved with the aid of KLW. Descriptive statistics for studies included in the meta-analysis are shown in Table 1.

Figure 1 about here

2.4 Statistical analysis

A random-effects model was selected on the basis of the methodological diversity across the studies [39]. Weighted raw point estimate and 95% confidence limits (95% CL) were calculated as summary statistics. Among-studies heterogeneity was quantified with the Tau statistic (τ) – a standard deviation describing the typical variability in the mean size-exponent between studies [40]. We derived a 95% prediction interval (95% PI) to quantify the expected range of true exponents for 95% of similar future studies [41]. This dispersion of true effects was then used together with the t-distribution [41] to estimate the probability of the true size-exponent in new studies being above the theoretical values of ⅔- or ¾-power, respectively. The observed probabilities were interpreted according to the following scale: <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 [42]. A meta-regression analysis was performed to assess the impact of each moderator on the estimated exponent. Small study bias was examined through Egger's test [43]. All statistical analyses were conducted using Comprehensive Meta-analysis software, version 3 (Biostat Inc., Englewood, NJ, USA).

Table 1 about here

3 Results

3.1 Study selection

Overall, 36 studies met the eligibility criteria out of 3,487 abstracts. The ultimate dataset also encompassed baseline values from two clinical trials conducted in a group of thirty obese boys and eighty-four adolescent girls, respectively [48, 52]. In addition to this, one time-based cross-sectional analysis of mixed-longitudinal study data, within a small sample of youth football players, was also extracted [36]. In one study [55], allometric exponents were quantified for fat-free mass measured using both dual-energy X-ray absorptiometry (DXA) and sum-of-skinfolds methods on the same samples of males and females. Being aware of the important issue of double counting in meta-analysis [68], we selected the DXA-derived data since it represents a criterion method for measuring body composition [69, 70]. Nevertheless, the overall pooled exponent, confidence limits, and prediction interval were found to be unaffected in the sensitivity analysis we undertook using the skinfolds-derived data from the above-quoted study. Likewise, we extracted distinct body mass [23] and fat-free mass [44] exponents derived from modeling $\dot{V}O_{2max}$ in two different studies with substantially overlapping sample characteristics (Table 1). Conversely, 17 investigations were dismissed because of the missing model precision statistics (i.e. 95% CL; SD; SE; Pearson's r), one due to lack of relevant information for each selected moderator and another study in which healthy and clinical subjects were combined.

The identified cohort of studies summarizes thirty-two years of research published between 1984 and 2016. The derived allometric exponents were dichotomized into two main domains relevant to whole-body mass (n = 27) and fat-free mass (n = 15) respectively (Fig. 2). The whole sample involved 6,514 humans. The observed size range was 57.2 kg (30.1 to 87.3).

Figure 2 about here

3.2 Study outcomes

3.2.1 Body mass

The meta-analysed exponent for whole-body mass (Fig. 2a) was found to be 0.70 (0.64 to 0.76). Substantial heterogeneity was observed, with the tau statistic being ± 0.150. A meta-regression revealed a substantial sex difference in the body mass exponent of 0.24 (0.11 to 0.38), with an observed magnitude of 0.76 in males and 0.52 in females. When sex was included in the full model as a moderator, tau reduced to ± 0.126. Sex accounted for 30% of the between-study variability in exponents (Table 2). None of the other moderators we examined explained a substantial proportion of between-study variance in mass exponent. The positive Egger's regression coefficient (intercept) of 1.35 (-0.34 to 3.05) suggested a small study bias, whereby studies involving small samples tend to have larger exponents (Fig. 3).

3.2.2 Fat-free mass

The pooled allometric exponent for fat-free mass (Fig. 2b) was 0.90 (0.83 to 0.96). The observed degree of heterogeneity between studies ($\tau = \pm 0.11$) was lower than that found for whole-body mass. The 95% PI for the fat-free mass exponent was estimated to be 0.68 to 1.12. The estimated probability in a future study for the true fat-free mass exponent being above ⅔- or ¾-power scaling is 0.98 (very likely) and 0.92 (likely), respectively. The meta-regression results indicated that none of the explored moderators had a substantial influence on the fat-free mass exponent (Table 3).

There was no trend for small study bias in the fat-free mass study cohort (Fig. 4), with Egger's coefficient being 0.38 (-1.11 to 1.87).

Table 2 about here

Table 3 about here

Figure 3 about here

Figure 4 about here

4 Discussion

The substantial proportion of unexplained between-study variance in the mean *b* estimates indicates that the notion of a constant generalizable exponent is untenable. The observed degree of heterogeneity (τ) reflects systematic differences between the empirically-derived size-exponents in the synthesised literature. The pooled estimated exponent for whole-body mass we derived (Fig. 2a) was approximately midway between the hypothesised ⅔ and ¾ power laws for the scaling of $\dot{V}O_{2max}$ to whole-body mass [71, 72]. Conversely, the mean of the distribution of the more physiologically-relevant fat-free mass exponent (Fig. 2b) was consistent with the predictions of the allometric cascade model [73].

According to this observation, we consider it important to emphasise here the correct interpretation of the random-effects meta-analysis. The 95% confidence limits quantify the uncertainty in the pooled mean of systematically different exponents reported in the studies, and do not quantify the heterogeneity [39]. Rather, the 95% PI (0.68 to 1.12) we calculated is now the preferred approach to account for both the variance of the pooled estimate (SE^2) and the between-study heterogeneity (τ^2). The 95% PI essentially quantifies the likely range for the true fat-free mass exponent that might be expected in future settings [41]. If the between-study heterogeneity was observed to be zero, then the 95% PI would equate to the respective 95% CL for the summary effect [41]. Given the between-study heterogeneity we found in this quantitative synthesis of the literature, the probability of the fat-free mass exponent in a future study being higher than the theorised ⅔ or ¾ power laws is 0.98 (very likely to be greater) and 0.92 (likely to be greater), respectively.

Importantly, in our meta-analysis, we have included estimates of size-exponents derived in independent groups (e.g. males and females; athletes and non-athletes; different age categories) both in different studies and also *within* the same study [55, 63, 67]. However, the random-effects model assumes the same variation in true exponents within the study and across studies, which might not apply [74].

Therefore, we performed a sensitivity analysis, for the fat-free mass exponent only, by conducting a random-effects meta-analysis of those studies that derived an exponent in a single group (N = 11). In this analysis, the pooled fat-free mass exponent was 0.90 (0.81 to 0.98), with a 95% PI of 0.66 to 1.14. The probability that the effect in a similar future study will be greater than ⅔- and ¾-power is 0.97 (very likely) and 0.90 (likely), respectively. The sensitivity analysis thus revealed that our original inference was robust, and not materially affected by including multiple independent samples within studies.

The *a priori* assumption of a universal rate-limited process hinders our understanding of the allometry of $\dot{V}O_{2max}$ [75]. The assumption dictates that energy metabolism increases to a fixed rate with whole-body mass, irrespective of an individual's metabolic state [71, 72]. This direct relationship has been supported only by studies on hibernating animals [75]. If a single-rate limiting process regulates aerobic metabolism, there would be no room left for metabolic scope from resting state to higher exercise intensities [76]. Moreover, body size range is an essential determinant for the precision of the exponent estimate [77]. Interspecific allometry studies provided evidence about the variations in maximal aerobic capacity by quantifying the size-induced effects *per se*. The broad size range reported in those studies (e.g. mouse-to-elephant) supports such a scaling regimen, although certain aspects of metabolic scaling may be concealed or even vary substantially at finer phyletic levels [77]. Our large-scale data highlight the importance of adjusting for relevant covariates since the relatively small size range from 30.1 to 87.3 kg is inadequate to overwhelm the sources of between-subject variability in $\dot{V}O_{2max}$ [28], conversely to what is observed in interspecific allometry [78].

This large-scale evidence synthesis confirms that, from a physiological perspective, fat-free mass, and not whole-body mass, should be considered the most appropriate scaling denominator for $\dot{V}O_{2max}$ [79, 80]. The canonical models dealt with the constraints of body size on Euclidean and biological grounds [71, 72]. The fractal geometry model implies a direct proportional relationship between blood volume and whole-body mass [81]. However, this fractal model does not consider the haemodynamic responses during exercise [80]. There is evidence to indicate that the metabolically active tissues dictate the regulation of supply and demand [31]. The redistribution of blood flow from resting state to maximal exercise intensity is well-documented in animal and human studies [82, 83], in which the systemic delivery converges to match the aerobic demands at the regional level, particularly for skeletal muscle [84]. It is thus paramount to consider the body compartments that contribute to the overall metabolic rate at a given exercise intensity.

Due to the substantial heterogeneity of body composition, whole-body mass is not a robust scaling denominator as it fails to reflect the real physiological processes during incremental exercise [30, 85, 86]. Indeed, Graves et al. [69] highlighted the superior fit of allometric models that incorporate fat-free mass, rather than whole-body mass, for scaling $\dot{V}O_{2max}$ within a paediatric population. Notwithstanding the substantial support for both leg muscle mass and fat-free mass as denominators of $\dot{V}O_{2max}$ [69], the diversity of physiological profiles within clinical and sports populations requires an independent holistic predictor of aerobic fitness. The concurrent declines of aerobic fitness and fat-free mass with aging and disease are well-documented [87-90]. Leg muscle mass appears to be a suitable body size descriptor in Association Football (soccer) [51], but on the other hand, upper-body musculature is a critical determinant of physical performance in sports like cross-country skiing [91, 92]. Therefore, fat-free mass should be

355 regarded as a more accurate reference standard for normalizing $\dot{V}O_{2max}$, irrespective of the examined
356 population.
357 We found a substantial sex-based difference of 0.24 (0.11 to 0.38) for the body mass exponent
358 (Table 2). Previous studies have pointed out the importance of testing the commonality of size-exponent
359 between sexes to avoid biases in the normalized indices [93, 94]. The main reason for the smaller exponent
360 in females is likely due to their greater amount of adipose tissue compared with males [52]. On the
361 contrary, our pooled exponent for fat-free mass was closer to a constant proportion, in both sexes (Table 3).
362 Notably, the normalization of cardiovascular function in adult subjects [93] shows an identical pattern to
363 that observed for energy metabolism [23, 64]. Since cardiac output is the major determinant of oxygen
364 supply [95], the allometric relationship between left-ventricular mass [93] and both the considered scaling
365 denominators for $\dot{V}O_{2max}$ [23] are in line with the observed magnitudes (Fig. 2). The interconnection
366 between fat-free mass and cardiovascular supply determines an additional line of evidence to account for
367 differences not only in body size but particularly in body composition when comparing maximal aerobic
368 capacity within heterogeneous samples [28, 79].
369 Not only do our results support the use of fat-free mass for routine normalisation of maximal
370 aerobic metabolism in humans [86, 69], they are also consistent with the allometric cascade model [73].
371 For fat-free mass, the pooled *b* value of 0.90 and the high probability that the true exponents in future
372 studies will be greater than ⅔- or ¾-power scaling thus appear to reflect the multi-level matching of energy
373 supply and demand under aerobic conditions [75, 96]. If, contrary to the fact, body fatness were observed
374 to be homogeneous between individuals across the examined cohort of studies, the exponents for whole-
375 body mass and fat-free mass would be identical. We interpret the deflation in the *b* value for whole-body
376 mass as a statistical artefact that results from the concurrent inclusion in the model of non-contributing
377 body proportions to maximal aerobic metabolism in humans. Our line of evidence, thus, substantiates a
378 notion already advanced 60 years ago that the amount of body fat mass could spuriously affect the obtained
379 size-exponent [30].
380
381 **5 Conclusions**
382
383 Our results indicate the adoption of an empirical approach, contrary to the proposed power laws, is
384 paramount to derive size-independent indices of $\dot{V}O_{2max}$. The observed discrepancy in the magnitudes of
385 the allometric exponents between whole-body mass and fat-free mass is likely the result of the lack of
386 statistical adjustment for known confounders in the body mass studies, especially adipose tissue. Since the
387 variation in energy supply appears to match the increasing demand towards maximal aerobic metabolism in
388 the active tissues, it is important to account for body proportions that do not contribute to the work
389 performed at vigorous intensity.
390 Our results provide large-scale empirical evidence against the normalization of $\dot{V}O_{2max}$ to whole-
391 body mass and underline the validity of normalising to fat-free mass. Our large sample of study cohorts
392 enabled a robust estimate of the expected range of true *b* values for fat-free mass in similar future studies,
393 the magnitude of which is consistent with the allometric cascade model.
394

Compliance with ethical standards

Funding

No sources of funding were used to assist in the preparation of this article

Conflicts of Interest

Lorenzo Lolli, Alan Batterham, Kathryn Weston and Greg Atkinson declare that they have no conflicts of interest relevant to the content of this review

References

1. Rowland TW. Does peak VO2 reflect VO2max in children?: evidence from supramaximal testing. Med Sci Sports Exerc. 1993;25(6):689-93.

2. Day JR, Rossiter HB, Coats EM, et al. The maximally attainable VO2 during exercise in humans: the peak vs. maximum issue. J Appl Physiol. 2003;95(5):1901-7.

3. Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. JAMA. 2009;301(19):2024-35.

4. Atkinson G, Davison R, Jeukendrup A, et al. Science and cycling: current knowledge and future directions for research. J Sports Sci. 2003;21(9):767-87.

5. Helgerud J. Maximal oxygen uptake, anaerobic threshold and running economy in women and men with similar performances level in marathons. Eur J Appl Physiol Occup Physiol. 1994;68(2):155-61.

6. Helgerud J, Rodas G, Kemi OJ, et al. Strength and endurance in elite football players. Int J Sports Med. 2011;32(9):677-82.

7. Sandbakk O, Hegge AM, Losnegard T, et al. The physiological capacity of the world's highest ranked female cross-country skiers. Med Sci Sports Exerc. 2016;48(6):1091-100.

8. Bacon AP, Carter RE, Ogle EA, et al. VO2max trainability and high intensity interval training in humans: a meta-analysis. PLoS One. 2013;8(9):e73182.

9. Milanovic Z, Pantelic S, Covic N, et al. Is recreational soccer effective for improving V̇O2max? a systematic review and meta-analysis. Sports Med. 2015;45(9):1339-53.

10. Sloth M, Sloth D, Overgaard K, et al. Effects of sprint interval training on VO2max and aerobic exercise performance: a systematic review and meta-analysis. Scand J Med Sci Sports. 2013;23(6):e341-52.

11. Weston KS, Wisloff U, Coombes JS. High-intensity interval training in patients with lifestyle-induced cardiometabolic disease: a systematic review and meta-analysis. Br J Sports Med. 2014;48(16):1227-34.

12. Curran-Everett D. Explorations in statistics: the analysis of ratios and normalized data. Adv Physiol Educ. 2013;37(3):213-9.

13. Packard GC, Boardman TJ. The use of percentages and size-specific indices to normalize physiological data for variation in body size: wasted time, wasted effort? Comp Biochem Phys A. 1999;122(1):37-44.

435 14. Bjorgen S, Helgerud J, Husby V, et al. Aerobic high intensity one-legged interval cycling improves
436 peak oxygen uptake in chronic obstructive pulmonary disease patients; no additional effect from hyperoxia.
437 Int J Sports Med. 2009;30(12):872-8.
438 15. Hoff J, Kemi OJ, Helgerud J. Strength and endurance differences between elite and junior elite ice
439 hockey players. The importance of allometric scaling. Int J Sports Med. 2005;26(7):537-41.
440 16. Loe H, Rognmo O, Saltin B, et al. Aerobic capacity reference data in 3816 healthy men and women 20-
441 90 years. PLoS One. 2013;8(5):e64319.
442 17. Saltin B, Larsen H, Terrados N, et al. Aerobic exercise capacity at sea level and at altitude in Kenyan
443 boys, junior and senior runners compared with Scandinavian runners. Scand J Med Sci Sports.
444 1995;5(4):209-21.
445 18. Cunha G, Lorenzi T, Sapata K, et al. Effect of biological maturation on maximal oxygen uptake and
446 ventilatory thresholds in soccer players: an allometric approach. J Sports Sci. 2011;29(10):1029-39.
447 19. Viickberg U, Purge P, Jürimäe T, et al. Interpretation of peak oxygen consumption in 10–12-year-old
448 soccer players: effect of biological maturation and body size. Acta Kinesiol Univ Tartu. 2013;19:16-30.
449 20. Tanner JM. Fallacy of per-weight and per-surface area standards, and their relation to spurious
450 correlation. J Appl Physiol. 1949;2(1):1-15.
451 21. Gould SJ. Allometry and size in ontogeny and phylogeny. Biol Rev Camb Philos Soc. 1966;41(4):587-
452 640.
453 22. Armstrong N, Welsman JR, Kirby BJ. Peak oxygen uptake and maturation in 12-yr olds. Med Sci
454 Sports Exerc. 1998;30(1):165-9.
455 23. Batterham AM, Vanderburgh PM, Mahar MT, et al. Modeling the influence of body size on VO2peak:
456 effects of model choice and body composition. J Appl Physiol. 1999;87(4):1317-25.
457 24. Heil DP. Body mass scaling of peak oxygen uptake in 20- to 79-yr-old adults. Med Sci Sports Exerc.
458 1997;29(12):1602-8.
459 25. Pettersen SA, Fredriksen PM, Ingjer F. The correlation between peak oxygen uptake (VO2peak) and
460 running performance in children and adolescents. Aspects of different units. Scand J Med Sci Sports.
461 2001;11(4):223-8.
462 26. Bokma F. Evidence against universal metabolic allometry. Funct Ecol. 2004;18(2):184-7.
463 27. Ives AR, Midford PE, Garland T, Jr. Within-species variation and measurement error in phylogenetic
464 comparative methods. Syst Biol. 2007;56(2):252-70.
465 28. Batterham AM, Jackson AS. Letter to the Editor. Respir Physiol Neurobiol. 2005;146:3-4.
466 29. Miller AT, Jr., Blyth CS. Lean body mass as a metabolic reference standard. J Appl Physiol.
467 1953;5(7):311-6.
468 30. Von Dobeln W. Maximal oxygen intake, body size, and total hemoglobin in normal man. Acta Physiol
469 Scand. 1957;38(2):193-9.
470 31. Weibel ER, Hoppeler H. Exercise-induced maximal metabolic rate scales with muscle aerobic capacity.
471 J Exp Biol. 2005;208(Pt 9):1635-44.
472 32. Amara CE, Koval JJ, Johnson PJ, et al. Modelling the influence of fat-free mass and physical activity
473 on the decline in maximal oxygen uptake with age in older humans. Exp Physiol. 2000;85(6):877-86.

474 33. Chamari K, Moussa-Chamari I, Boussaïdi L, et al. Appropriate interpretation of aerobic capacity:
475 allometric scaling in adult and young soccer players. Br J Sports Med. 2005;39(2):97-101.
476 34. Davies MJ, Dalsky GP, Vanderburgh PM. Allometric scaling of VO2max by body mass and lean body
477 mass in older men. J Aging Phys Act. 1995;3(4):324-31.
478 35. Goosey-Tolfrey VL, Batterham AM, Tolfrey K. Scaling behavior of V̇O2peak in trained wheelchair
479 athletes. Med Sci Sports Exerc. 2003;35(12):2106-11.
480 36. Segers V, De Clercq D, Janssens M, et al. Running economy in early and late maturing youth soccer
481 players does not differ. Br J Sports Med. 2008;42(4):289-94.
482 37. Tolfrey K, Barker A, Thom JM, et al. Scaling of maximal oxygen uptake by lower leg muscle volume
483 in boys and men. J Appl Physiol. 2006;100(6):1851-6.
484 38. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and
485 meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ.
486 2009;339:b2700.
487 39. Borenstein M, Hedges LV, Higgins JPT, et al. A basic introduction to fixed-effect and random-effects
488 models for meta-analysis. Res Synth Methods. 2010;1(2):97-111.
489 40. Higgins JPT. Commentary: heterogeneity in meta-analysis should be expected and appropriately
490 quantified. Int J Epidemiol. 2008;37(5):1158-60.
491 41. IntHout J, Ioannidis JP, Rovers MM, et al. Plea for routinely presenting prediction intervals in meta-
492 analysis. BMJ Open. 2016;6(7):e010247.
493 42. Hopkins WG, Batterham AM. Error rates, decisive outcomes and publication bias with several
494 inferential methods. Sports Med. 2016;46(10):1563-73.
495 43. Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test.
496 BMJ. 1997;315(7109):629-34.
497 44. Batterham AM, Jackson AS. Validity of the allometric cascade model at submaximal and maximal
498 metabolic rates in exercising men. Respir Physiol Neurobiol. 2003;135(1):103-6.
499 45. Batterham AM, Tolfrey K, George KP. Nevill's explanation of Kleiber's 0.75 mass exponent: An
500 artifact of collinearity problems in least squares models? J Appl Physiol. 1997;82(2):693-7.
501 46. Bloxham SR, Welsman JR, Armstrong N. Ergometer-specific relationships between peak oxygen
502 uptake and short-term power output in children. Pediatr Exerc Sci. 2005;17(2):136-48.
503 47. Carvalho HM, Coelho-e-Silva MJ, Eisenmann JC, et al. Aerobic fitness, maturation, and training
504 experience in youth basketball. Int J Sports Physiol Perform. 2013;8(4):428-34.
505 48. Carvalho HM, Milano GE, Lopes WA, et al. Peak oxygen uptake responses to training in obese
506 adolescents: a multilevel allometric framework to partition the influence of body size and maturity status.
507 Biomed Res Int. 2013;2013:618595.
508 49. Chia M, Aziz AR. Modelling maximal oxygen uptake in athletes: allometric scaling versus ratio-
509 scaling in relation to body mass. Ann Acad Med Singapore. 2008;37(4):300-6.
510 50. Cooper DM, Weiler-Ravell D, Whipp BJ, et al. Aerobic parameters of exercise as a function of body
511 size during growth in children. J Appl Physiol Respir Environ Exerc Physiol. 1984;56(3):628-34.
512 51. Cunha G, Vaz MA, Geremia JM, et al. Maturity status does not exert effects on aerobic fitness in
513 soccer players after appropriate normalization for body size. Pediatr Exerc Sci. 2016;28(3):456-65.

514 52. Eliakim A, Scheett T, Allmendinger N, et al. Training, muscle volume, and energy expenditure in
515 nonobese American girls. J Appl Physiol. 2001;90(1):35-44.
516 53. Jullien H, Ahmaidi S, Doutrellot PL, et al. Relationship between oxygen consumption and body mass
517 during treadmill and cycle ergometry respectively. Sports Med Training Rehab. 1999;9(2):89-99.
518 54. Markovic G, Vucetic V, Nevill AM. Scaling behaviour of V̇O2 in athletes and untrained individuals.
519 Ann Hum Biol. 2007;34(3):315-28.
520 55. Neder JA, Lerario MC, Castro ML, et al. Peak V̇O2 correction for fat-free mass estimated by
521 anthropometry and DEXA. Med Sci Sports Exerc. 2001;33(11):1968-75.
522 56. Nes BM, Osthus IB, Welde B, et al. Peak oxygen uptake and physical activity in 13- to 18-year-olds:
523 the Young-HUNT study. Med Sci Sports Exerc. 2013;45(2):304-13.
524 57. Nevill A, Rowland T, Goff D, et al. Scaling or normalising maximum oxygen uptake to predict 1-mile
525 run time in boys. Eur J Appl Physiol. 2004;92(3):285-8.
526 58. Nevill AM, Brown D, Godfrey R, et al. Modeling maximum oxygen uptake of elite endurance athletes.
527 Med Sci Sports Exerc. 2003;35(3):488-94.
528 59. Nevill AM, Markovic G, Vucetic V, et al. Can greater muscularity in larger individuals resolve the 3/4
529 power-law controversy when modelling maximum oxygen uptake? Ann Hum Biol. 2004;31(4):436-45.
530 60. Rogers DM, Turley KR, Kujawa KI, et al. Allometric scaling factors for oxygen uptake during exercise
531 in children. Pediatr Exerc Sci. 1995;7(1):12-25.
532 61. Rowland T, Miller K, Vanderburgh P, et al. Cardiovascular fitness in premenarcheal girls and young
533 women. Int J Sports Med. 2000;21(2):117-21.
534 62. Tartaruga MP, De Medeiros MH, Alberton CL, et al. Application of the allometric scale for the
535 submaximal oxygen uptake in runners and rowers. Biol Sport. 2010;27(4):297-300.
536 63. Valente-Dos-Santos J, Coelho-E-Silva MJ, Tavares ÓM, et al. Allometric modelling of peak oxygen
537 uptake in male soccer players of 8-18 years of age. Ann Hum Biol. 2015;42(2):125-33.
538 64. Vanderburgh PM, Katch FI. Ratio scaling of VO2max penalizes women with larger percent body fat,
539 not lean body mass. Med Sci Sports Exerc. 1996;28(9):1204-8.
540 65. Welsman JR, Armstrong N, Kirby BJ, et al. Exercise performance and magnetic resonance imaging-
541 determined thigh muscle volume in children. Eur J Appl Physiol Occup Physiol. 1997;76(1):92-7.
542 66. Welsman JR, Armstrong N, Nevill AM, et al. Scaling peak VO2 for differences in body size. Med Sci
543 Sports Exerc. 1996;28(2):259-65.
544 67. Wijndaele K, Duvigneaud N, Matton L, et al. Muscular strength, aerobic fitness, and metabolic
545 syndrome risk in Flemish adults. Med Sci Sports Exerc. 2007;39(2):233-40.
546 68. Senn SJ. Overstating the evidence: double counting in meta-analysis and related problems. BMC Med
547 Res Methodol. 2009;9:10.
548 69. Graves LE, Batterham AM, Foweather L, et al. Scaling of peak oxygen uptake in children: a
549 comparison of three body size index models. Med Sci Sports Exerc. 2013;45(12):2341-5.
550 70. Jensky-Squires NE, Dieli-Conwright CM, Rossuello A, et al. Validity and reliability of body
551 composition analysers in children and adults. Br J Nutr. 2008;100(4):859-65.
552 71. Schmidt-Nielsen K. Energy metabolism, body size, and problems of scaling. Fed Proc.
553 1970;29(4):1524-32.

554 72. West GB, Brown JH, Enquist BJ. The fourth dimension of life: fractal geometry and allometric scaling
555 of organisms. Science. 1999;284(5420):1677-9.
556 73. Darveau CA, Suarez RK, Andrews RD, et al. Allometric cascade as a unifying principle of body mass
557 effects on metabolism. Nature. 2002;417(6885):166-70.
558 74. Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version
559 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from
560 http://www.handbook.cochrane.org/.
561 75. Suarez RK, Darveau CA. Multi-level regulation and metabolic scaling. J Exp Biol. 2005;208(Pt
562 9):1627-34.
563 76. Hochachka PW, Darveau CA, Andrews RD, et al. Allometric cascade: a model for resolving body mass
564 effects on metabolism. Comp Biochem Phys A. 2003;134(4):675-91.
565 77. Calder WA. Scaling energetics of homeothermic vertebrates - an operational allometry. Annu Rev
566 Physiol. 1987;49:107-20.
567 78. White CR, Kearney MR. Metabolic scaling in animals: methods, empirical results, and theoretical
568 explanations. Compr Physiol. 2014;4(1):231-56.
569 79. Hunt BE, Davy KP, Jones PP, et al. Role of central circulatory factors in the fat-free mass-maximal
570 aerobic capacity relation across age. Am J Physiol. 1998;275(4 Pt 2):H1178-82.
571 80. Saltin B. Exercise hyperaemia: magnitude and aspects on regulation in humans. J Physiol. 2007;583(Pt
572 3):819-23.
573 81. Savage VM, Deeds EJ, Fontana W. Sizing up allometric scaling theory. PLoS Comput Biol.
574 2008;4(9):e1000171.
575 82. Mueller PJ, O'Hagan KP, Skogg KA, et al. Renal hemodynamic responses to dynamic exercise in
576 rabbits. J Appl Physiol. 1998;85(5):1605-14.
577 83. Parks CM, Manohar M. Distribution of blood flow during moderate and strenuous exercise in ponies
578 (Equus caballus). Am J Vet Res. 1983;44(10):1861-6.
579 84. Andersen P, Saltin B. Maximal perfusion of skeletal muscle in man. J Physiol. 1985;366:233-49.
580 85. Goran M, Fields DA, Hunter GR, et al. Total body fat does not influence maximal aerobic capacity. Int
581 J Obes. 2000;24(7):841-8.
582 86. Toth MJ, Goran MI, Ades PA, et al. Examination of data normalization procedures for expressing peak
583 VO2 data. J Appl Physiol. 1993;75(5):2288-92.
584 87. Carrick-Ranson G, Hastings JL, Bhella PS, et al. The effect of age-related differences in body size and
585 composition on cardiovascular determinants of VO2max. J Gerontol A Biol Sci Med Sci. 2013;68(5):608-
586 16.
587 88. Leikis MJ, McKenna MJ, Petersen AC, et al. Exercise performance falls over time in patients with
588 chronic kidney disease despite maintenance of hemoglobin concentration. Clin J Am Soc Nephrol.
589 2006;1(3):488-95.
590 89. Toth MJ, Gardner AW, Ades PA, et al. Contribution of body composition and physical activity to age-
591 related decline in peak VO2 in men and women. J Appl Physiol. 1994;77(2):647-52.
592 90. Workeneh BT, Mitch WE. Review of muscle wasting associated with chronic kidney disease. Am J
593 Clin Nutr. 2010;91(4):1128-32.

91. Sandbakk O, Ettema G, Holmberg HC. Gender differences in endurance performance by elite cross-country skiers are influenced by the contribution from poling. Scand J Med Sci Sports. 2014;24(1):28-33.
92. Stoggl T, Enqvist J, Muller E, et al. Relationships between body composition, body dimensions, and peak speed in cross-country sprint skiing. J Sports Sci. 2010;28(2):161-9.
93. Batterham AM, George KP, Mullineaux DR. Allometric scaling of left ventricular mass by body dimensions in males and females. Med Sci Sports Exerc. 1997;29(2):181-6.
94. Vanderburgh PM. Two important cautions in the use of allometric scaling: the common exponent and group difference principles. Meas Phys Educ Exerc Sci. 1998;2(3):153-63.
95. Wagner PD. Algebraic analysis of the determinants of VO2max. Respir Physiol. 1993;93(2):221-37.
96. Zakeri I, Puyau MR, Adolph AL, et al. Normalization of energy expenditure data for differences in body mass or composition in children and adolescents. J Nutr. 2006;136(5):1371-6.

Table Legends

Table 1 Summary of characteristics of the thirty-six studies meeting the eligibility criteria

Table 2 Allometric exponents for whole-body mass and the various potential moderator variables

Table 3 Allometric exponents for fat-free mass and the various potential moderator variables

Figure Legends

Fig. 1 Flow diagram of the systematic review process; $\dot{V}O_{2max}$: maximal oxygen uptake

Fig. 2 Forest plots illustrating the included studies relevant to whole-body mass (a) fat-free mass (b)

Fig. 3 Funnel plot of the mean point estimate for the whole-body mass exponent versus standard error

Fig. 4 Funnel plot of the mean point estimate for the fat-free mass exponent versus standard error

Table 1 Summary of characteristics of the thirty-six studies meeting the eligibility criteria

Study	n	Population	Age (years) Mean	Body size (kg) Mean	Body composition assessment	$\dot{V}O_{2max}$ ($L \cdot min^{-1}$) Mean	Exercise testing mode	Scaling denominator(s)	Line-fitting method
Amara et al. [32]	152	Healthy men	68.7	78.2	SSK	1.78	Treadmill	Fat-free mass	Log-linear
	146	Healthy women	70.0	63.8		1.21			
Armstrong et al. [22]	106	Middle school boys	12.2	41.2	SSK	2.10	Treadmill	Body mass	Log-linear
	106	Middle school girls	12.2	43.9		1.92			
Batterham and Jackson [44]	1629	Healthy men	45.0	78.8	SSK	3.09	Treadmill	Fat-free mass	Non-linear
Batterham et al. [45]	26	Pre-pubertal boys	11.0	34.3	-	1.82	Treadmill	Body mass	Log-linear
	26	Teenage boys	14.1	49.5		2.60			
	23	Healthy men	22.4	76.7		4.18			
Batterham et al. [23]	1314	Healthy men employed at NASA	44.6	79.2	SSK	3.08	Treadmill	Body mass / fat-free mass	Non-linear
Bloxham et al. [46]	28	Middle school boys	11.7	40.1	SSK	2.09	Treadmill	Body mass	Log-linear
						1.84	Cycle ergometer		
	28	Middle school girls	11.7	42.2		1.94	Treadmill		
						1.69	Cycle ergometer		
Carvalho et al. [47]	37	Youth male basketball players	15.3	73.3	ADP	4.65	Treadmill	Body mass / fat-free mass	Log-linear
Carvalho et al. [48]	30	Obese boys	13.0[a]	72.4	BIA	2.41	Treadmill	Body mass / fat-free mass	Log-linear
Chamari et al. [33]	24	Senior male football players	24.0	75.7	SSK	4.45	Treadmill	Fat-free mass	Log-linear
	21	Youth male football players	14.0	60.2		3.60			
Chia and Aziz [49]	158	Male athletes	21.7	64.8	-	3.73	Treadmill	Body mass	Log-linear
	28	Female athletes	21.9	53.0		2.53			
Cooper et al. [50]	21	Healthy male adolescents	16.0	65.0	-	3.25[a]	Cycle ergometer	Body mass	Log-linear
	37	Healthy boys	10.0	34.0		1.43[a]			
	27	Healthy female adolescents	15.0	52.0		1.77[a]			
	24	Healthy girls	9.0	33.0		1.25[a]			
Cunha et al. [18]	52	Pubescent youth male football players	13.4	62.5	-	3.72	Treadmill	Body mass	Log-linear
	58	Post-pubescent youth male football players	17.0	73.9		4.45			
Cunha et al. [51]	14	Pre-pubescent youth male football players	13.3	55.0	-	3.45[a]	Treadmill	Body mass	Log-linear
	38	Pubescent youth male football players	15.3	68.0		4.16[a]			
	27	Post-pubescent youth male football players	16.4	74.7		4.35[a]			
Davies et al. [34]	73	Healthy ambulatory men	69.7	80.2	DXA	2.20	Treadmill	Body mass / fat-free mass	Log-linear
Eliakim et al. [52]	40	Pre- and early pubertal girls	9.2	33.8	-	1.44[a]	Cycle ergometer	Body mass	Non-linear
	22	High-school girls	16.0[a]	56.4		1.57			
	22			60.3		1.48			

Table 1 Continued

Study	n	Population	Age (years) Mean	Body size (kg) Mean	Body composition assessment	$\dot{V}O_{2max}$ ($L \cdot min^{-1}$) Mean	Exercise testing mode	Scaling denominator(s)	Line-fitting method
Goosey-Tolfrey et al. [35]	20	Male wheelchair athletes	27.0	67.9	SSK	2.35	Wheelchair ergometer	Fat-free mass	Non-linear
	25		28.0	70.1		2.67			
Heil [24]	210	Healthy men	45.6	81.8	SSK	3.54	Treadmill	Fat-free mass	Log-linear
	230	Healthy women	47.5	64.7		2.14			
Jullien et al. [53]	13	Youth male football players	17.5[a]	65.0	-	4.13[a]	Treadmill	Body mass	Log-linear
						4.29[a]	Cycle ergometer		
Markovic et al. [54]	270	Male athletes	22.2	80.1	-	4.50	Treadmill	Body mass	Log-linear
	43	Untrained men	21.7	78.0		3.90			
Neder et al. [55]	9	Healthy men (university staff)	29.5[a]	73.8	SSK / DXA	2.64	Cycle ergometer	Fat-free mass	Log-linear
	13		49.5[a]	79.6		2.21			
	14		70.0[a]	74.9		1.58			
	9	Healthy women (university staff)	29.5[a]	66.4		1.67			
	14		49.5[a]	64.3		1.31			
	18		70.0[a]	62.5		1.06			
Nes et al. [56]	281	Healthy male adolescents	15.7	66.7	-	3.92	Treadmill	Body mass	Log-linear
	289	Healthy female adolescents	15.6	58.0		2.78			
Nevill et al. [58]	98	International-standard male athletes	24.2	73.9	SSK	5.19[a]	Treadmill / rowing ergometer	Fat-free mass	Log-linear
	76	International-standard female athletes	24.5	61.6		3.56[a]			
Nevill et al. [57]	36	Healthy circumpubertal boys	12.2	45.6	SSK	2.14[a]	Cycle ergometer	Body mass	Log-linear
Nevill et al. [59]	4	Professional male football players	23.5	87.3	SSK	3.99[a]	Treadmill	Body mass	Log-linear
	33		22.1	78.9		4.34[a]			
	64		22.0	75.3		4.37[a]			
	18		22.1	77.3		4.58[a]			
Pettersen et al. [25]	11	Healthy boys / adolescents	8.5[a]	31.7	-	1.79	Treadmill	Body mass	Log-linear
	26		10.5[a]	35.4		2.00			
	33		12.5[a]	46.8		2.56			
	24		14.5[a]	57.3		3.47			
	13		16.5[a]	65.8		3.84			
Rogers et al. [60]	21	Healthy boys	9.0	30.4	SSK	1.66	Treadmill	Body mass	Log-linear
	21	Healthy girls	8.8	30.1		1.50			
Rowland et al. [61]	24	Healthy girls	11.7	46.9	SSK	1.89[a]	Cycle ergometer	Body mass	Log-linear
	17	Healthy young adult women	27.4	62.5		2.17[a]			

Table 1 Continued

Study	n	Population	Age (years) Mean	Body size (kg) Mean	Body composition assessment	$\dot{V}O_{2max}$ ($L \cdot min^{-1}$) Mean	Exercise testing mode	Scaling denominator(s)	Line-fitting method
Segers et al. [36]	6	Early mature youth male football players	14.3	65.2	SSK	3.84	Treadmill	Fat-free mass	Log-linear
	7	Late mature youth male football players	14.4	43.3		2.57			
Tartaruga et al. [62]	11	Male long distance runners	22.3	61.7	SSK	3.40	Treadmill	Body mass	Log-linear
	15	Elite male rowers	24.0	83.5		5.10	Rowing ergometer		
Tolfrey et al. [37]	15	Healthy boys	12.3	43.6	SSK	2.23	Treadmill	Body mass / fat-free mass	Non-linear
	14	Healthy men	25.4	78.3		4.10			
Valente-Dos-Santos et al. [63]	20	Youth male football players	10.5	40.5	PE / DXA	2.00	Treadmill	Body mass / fat-free mass	Log-linear
	31		14.6	61.8		3.33			
	30		17.4	69.7		3.83			
Vanderburgh and Katch [64]	94	Healthy women	27.4	60.3	UW	2.70	Treadmill	Body mass / fat-free mass	Log-linear
Viickberg et al. [19]	29	Late mature youth male football players	11.4	36.6	-	1.89	Cycle ergometer	Body mass	Log-linear
	26	Average mature youth male football players	11.2	39.0		2.02			
	9	Early mature youth male football players	10.8	42.7		2.14			
Welsman et al. [66]	29	Prepubertal boys	10.7	34.9	-	1.76	Treadmill	Body mass	Log-linear
	26	Circumpubertal male adolescents	14.1	49.5		2.60			
	18	Male adults	22.8	78.6		4.18			
	33	Prepubertal girls	10.7	32.7		1.48			
	34	Circumpubertal female adolescents	13.0	46.5		2.14			
	16	Female adults	21.7	60.5		2.58			
Welsman et al. [65]	16	Middle school boys	9.9	32.0	-	1.95	Treadmill	Body mass	Log-linear
	16	Middle school girls	9.9	35.5		1.81			
Wijndaele et al. [67]	571	Men at risk of metabolic syndrome	46.7	79.5[b]	BIA	2.90	Cycle ergometer	Fat-free mass	Log-linear
	448	Women at risk of metabolic syndrome	45.8	65.3[b]		1.80			

[a] : values derived from the available data; [b] : personal communication; ADP: Air-displacement plethysmography; BIA: Bioelectrical impedance analysis; DXA: Dual-energy X-ray absorptiometry; NASA: National Aeronautics and Space Adminstration Johnson Center; PE: Predictive equation; SSK: Sum-of-skinfolds; UW: Underwater weighing; $\dot{V}O_{2max}$: Maximal oxygen uptake

Table 2 Allometric exponents for whole-body mass and the various potential moderator variables

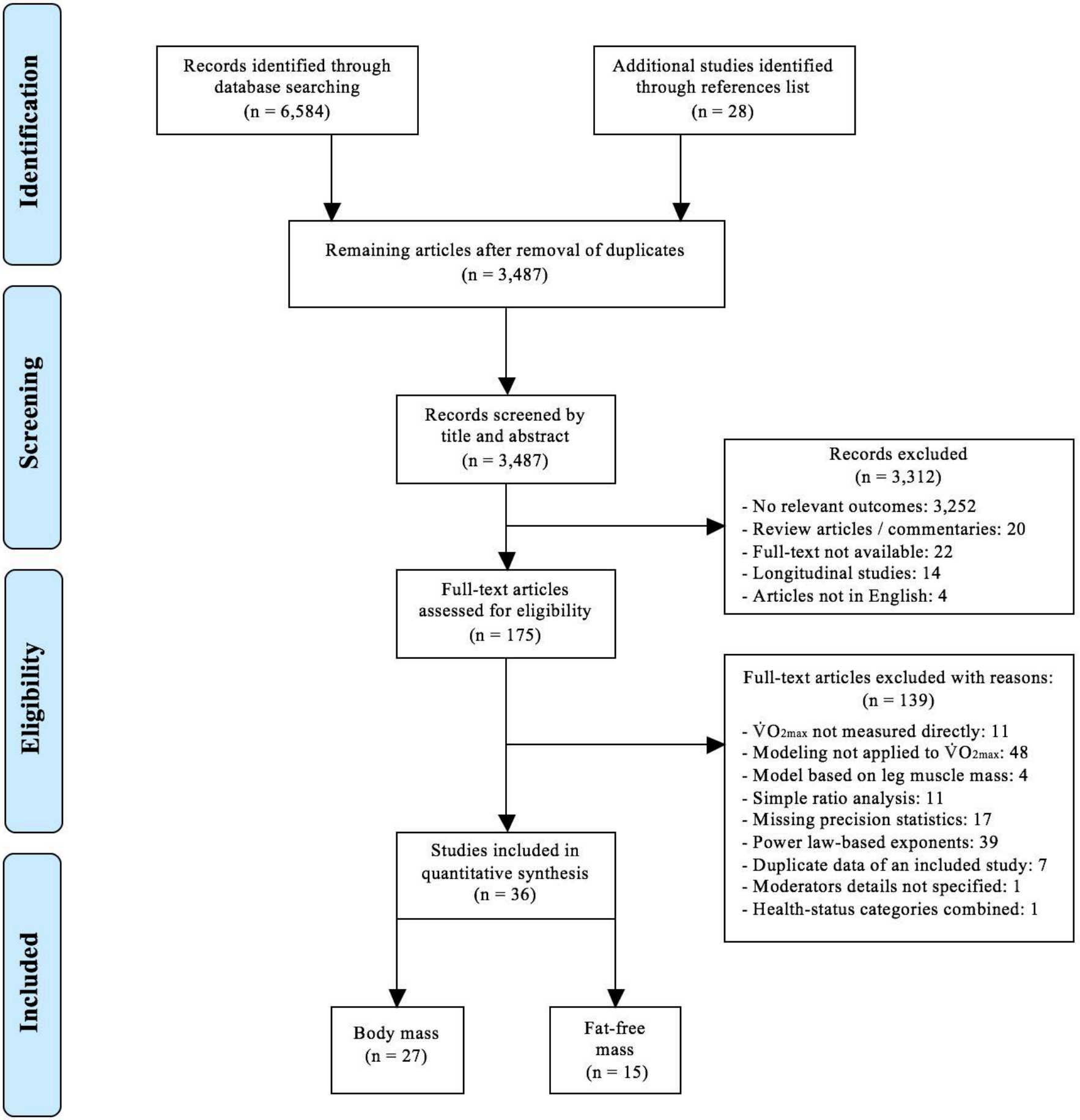
Moderator	Subgroup	n	*b*	95% CL	R^2
Age category	Young	19	0.71	0.62 - 0.79	0.00
	Adult	11	0.67	0.55 - 0.78	
	Mixed	4	0.72	0.54 - 0.89	
Sporting background	Athletes	14	0.71	0.61 - 0.81	0.00
	Non-athletes	20	0.69	0.61 - 0.76	
Sex	Male	23	0.76	0.70 - 0.83	0.30
	Female	6	0.52	0.40 - 0.64	
	Mixed	5	0.64	0.52 - 0.76	
$\dot{V}O_{2max}$ testing					
Criteria	None	1	0.76	0.10 - 1.42	0.02
	Stated but unverified	18	0.71	0.63 - 0.79	
	Stated and verified	15	0.68	0.59 - 0.76	
Mode	Treadmill	26	0.70	0.63 - 0.77	0.00
	Cycle ergometer	6	0.68	0.54 - 0.81	
	Other	2	0.72	0.34 - 1.10	
Line-fitting method	Log-linear	31	0.71	0.65 - 0.78	0.00
	Non-linear	3	0.55	0.35 - 0.75	
Model adjustment	Univariate	22	0.67	0.59 - 0.75	0.02
	Multivariable	12	0.73	0.64 - 0.82	

b : allometric exponent; CL : confidence limits; n: number of effects; R^2 : proportion of between-study variance explained by the moderator; $\dot{V}O_{2max}$: maximal oxygen uptake

Table 3 Allometric exponents for fat-free mass and the various potential moderator variables

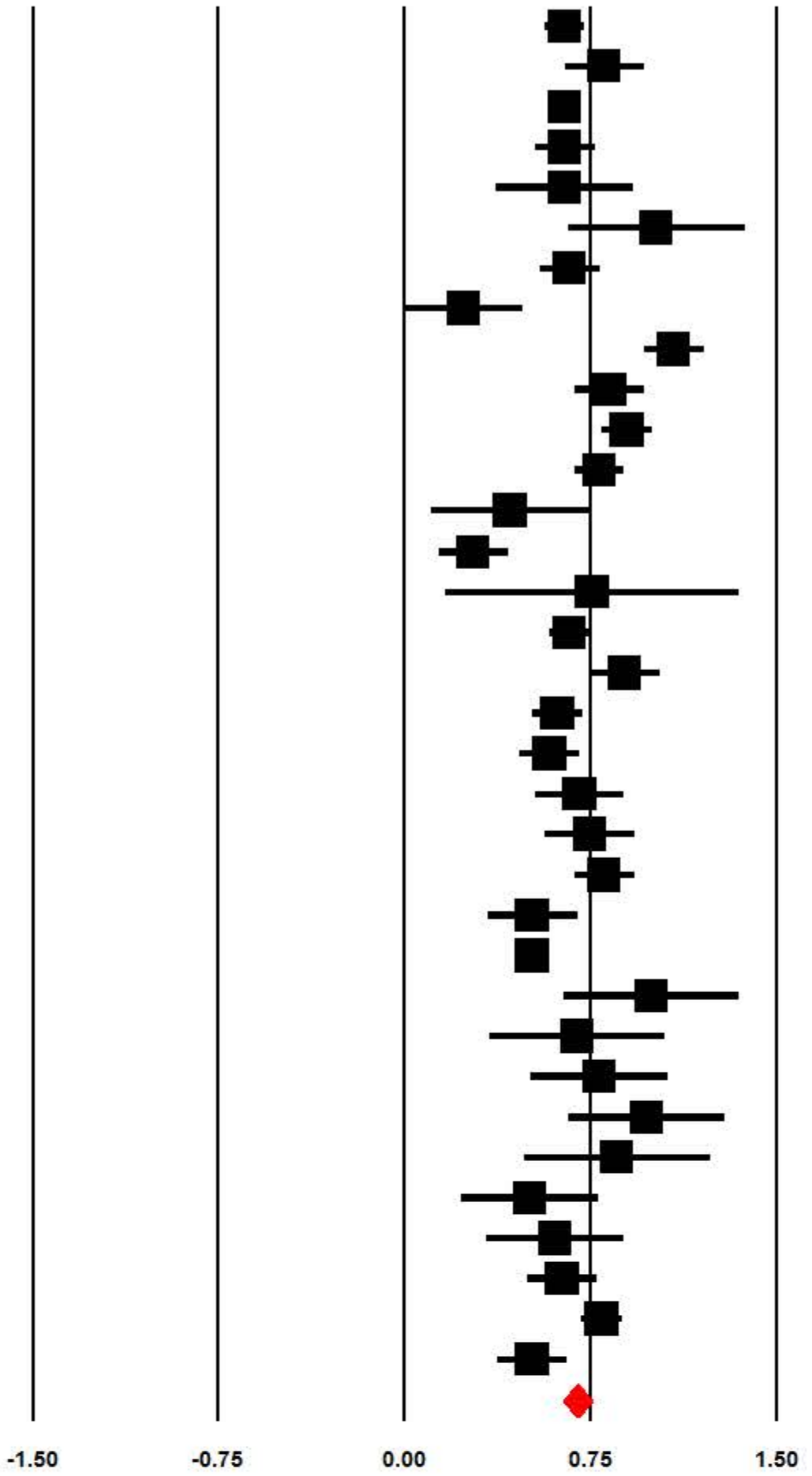
Moderator	Subgroup	n	*b*	95% CL	R^2
Age category	Young	7	0.92	0.79 - 1.05	0.17
	Adult	11	0.90	0.82 - 0.98	
	Mixed	2	0.81	0.65 - 0.98	
Sporting background	Athletes	9	0.85	0.75 - 0.95	0.11
	Non-athletes	11	0.92	0.84 - 1.01	
Sex	Male	14	0.89	0.80 - 0.98	0.00
	Female	3	0.94	0.75 - 1.13	
	Mixed	3	0.88	0.73 - 1.03	
$\dot{V}O_{2max}$ testing					
Criteria	None	2	0.90	0.71 - 1.09	0.00
	Stated but unverified	12	0.88	0.79 - 0.97	
	Stated and verified	6	0.93	0.80 - 1.06	
Mode	Treadmill	14	0.90	0.82 - 0.99	0.00
	Cycle ergometer	4	0.87	0.71 - 1.04	
	Other	2	0.90	0.68 - 1.11	
Body composition assessment	Greater precision (e.g. MRI)	4	0.95	0.76 - 1.14	0.00
	Lower precision (e.g. SSK)	16	0.89	0.82 - 0.96	
Line-fitting method	Log-linear	17	0.89	0.81 - 0.97	0.00
	Non-linear	3	0.93	0.77 - 1.09	
Model adjustment	Univariate	11	0.93	0.82 - 1.04	0.00
	Multivariable	9	0.88	0.79 - 0.96	

b : allometric exponent; CL : confidence limits; n: number of effects; R^2 : proportion of between-study variance explained by the moderator; MRI : magnetic resonance imaging; SSK : sum of skinfolds; $\dot{V}O_{2max}$: maximal oxygen uptake


Identification
Screening
Eligibility
Included
Records identified through database searching (n = 6,584)
Additional studies identified through references list (n = 28)
Remaining articles after removal of duplicates (n = 3,487)
Records screened by title and abstract (n = 3,487)
Records excluded (n = 3,312)
- No relevant outcomes: 3,252
- Review articles / commentaries: 20
- Full-text not available: 22
- Longitudinal studies: 14
- Articles not in English: 4
Full-text articles assessed for eligibility (n = 175)
Full-text articles excluded with reasons: (n = 139)
- V̇O2max not measured directly: 11
- Modeling not applied to V̇O2max: 48
- Model based on leg muscle mass: 4
- Simple ratio analysis: 11
- Missing precision statistics: 17
- Power law-based exponents: 39
- Duplicate data of an included study: 7
- Moderators details not specified: 1
- Health-status categories combined: 1
Studies included in quantitative synthesis (n = 36)
Body mass (n = 27)
Fat-free mass (n = 15)

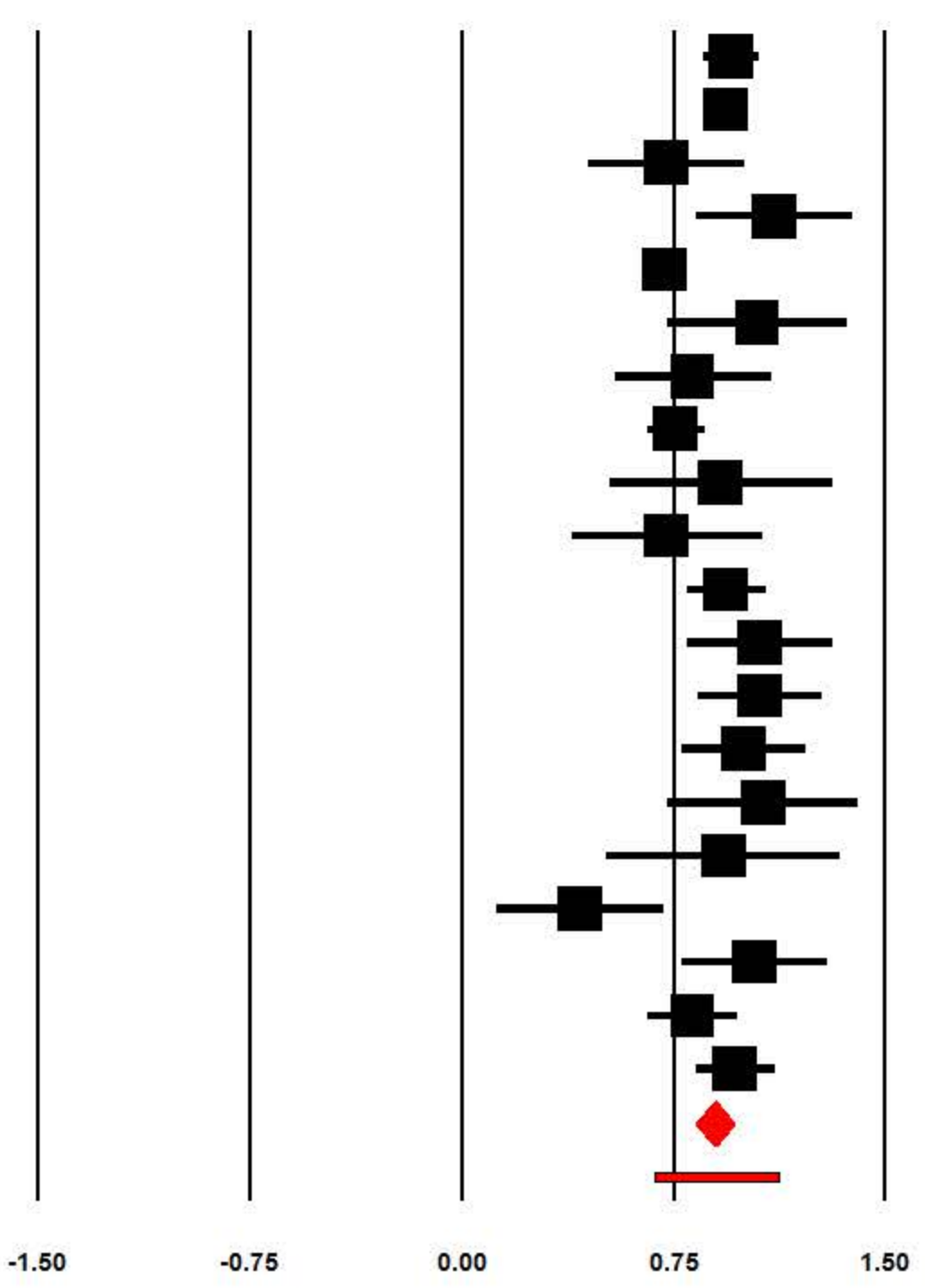
a

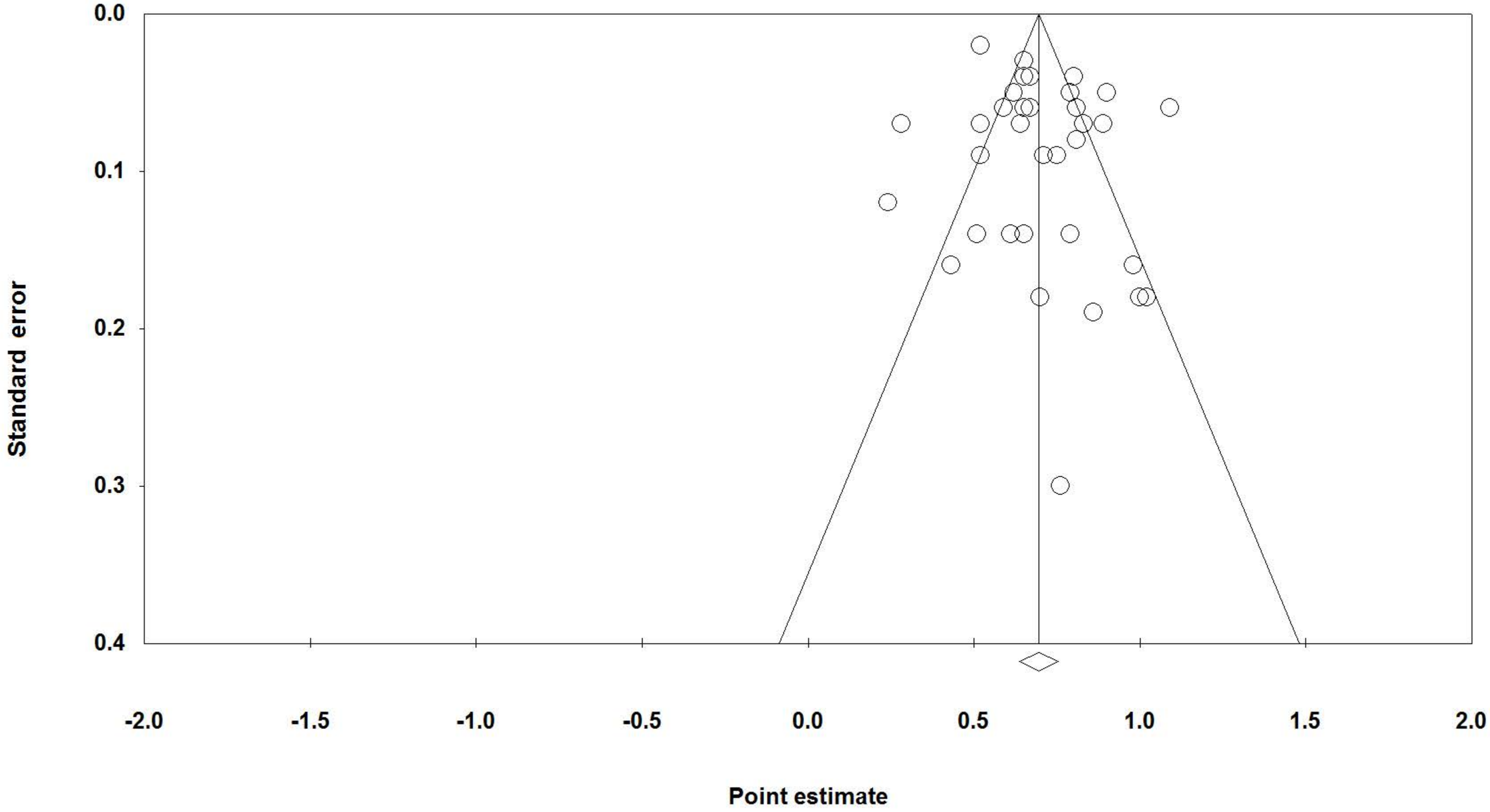
Study	Point estimate	Standard error	Lower limit	Upper limit
Armstrong et al. [22]	0.65	0.04	0.57	0.73
Batterham et al. [45]	0.81	0.08	0.65	0.97
Batterham et al. [23]	0.65	0.03	0.59	0.71
Bloxham et al. [46]	0.65	0.06	0.53	0.77
Carvalho et al. [47]	0.65	0.14	0.38	0.92
Carvalho et al. [48]	1.02	0.18	0.67	1.37
Chia and Aziz [49]	0.67	0.06	0.55	0.79
Chia and Aziz [49]	0.24	0.12	0.00	0.48
Cooper et al. [50]	1.09	0.06	0.97	1.21
Cooper et al. [50]	0.83	0.07	0.69	0.97
Cunha et al. [18]	0.90	0.05	0.80	1.00
Cunha et al. [51]	0.79	0.05	0.69	0.89
Davies et al. [34]	0.43	0.16	0.12	0.74
Eliakim et al. [52]	0.28	0.07	0.14	0.42
Jullien et al. [53]	0.76	0.30	0.17	1.35
Markovic et al. [54]	0.67	0.04	0.59	0.75
Markovic et al. [54]	0.89	0.07	0.75	1.03
Nes et al. [56]	0.62	0.05	0.52	0.72
Nes et al. [56]	0.59	0.06	0.47	0.71
Nevill et al. [57]	0.71	0.09	0.53	0.89
Nevill et al. [59]	0.75	0.09	0.57	0.93
Pettersen et al. [25]	0.81	0.06	0.69	0.93
Rogers et al. [60]	0.52	0.09	0.34	0.70
Rowland et al. [61]	0.52	0.02	0.48	0.56
Tartaruga et al. [62]	1.00	0.18	0.65	1.35
Tartaruga et al. [62]	0.70	0.18	0.35	1.05
Tolfrey et al. [37]	0.79	0.14	0.52	1.06
Valente-Dos-Santos et al. [63]	0.98	0.16	0.67	1.29
Valente-Dos-Santos et al. [63]	0.86	0.19	0.49	1.23
Valente-Dos-Santos et al. [63]	0.51	0.14	0.24	0.78
Vanderburgh and Katch [64]	0.61	0.14	0.34	0.88
Viickberg et al. [19]	0.64	0.07	0.50	0.78
Welsman et al. [66]	0.80	0.04	0.72	0.88
Welsman et al. [65]	0.52	0.07	0.38	0.66
Pooled	**0.70**	**0.03**	**0.64**	**0.76**

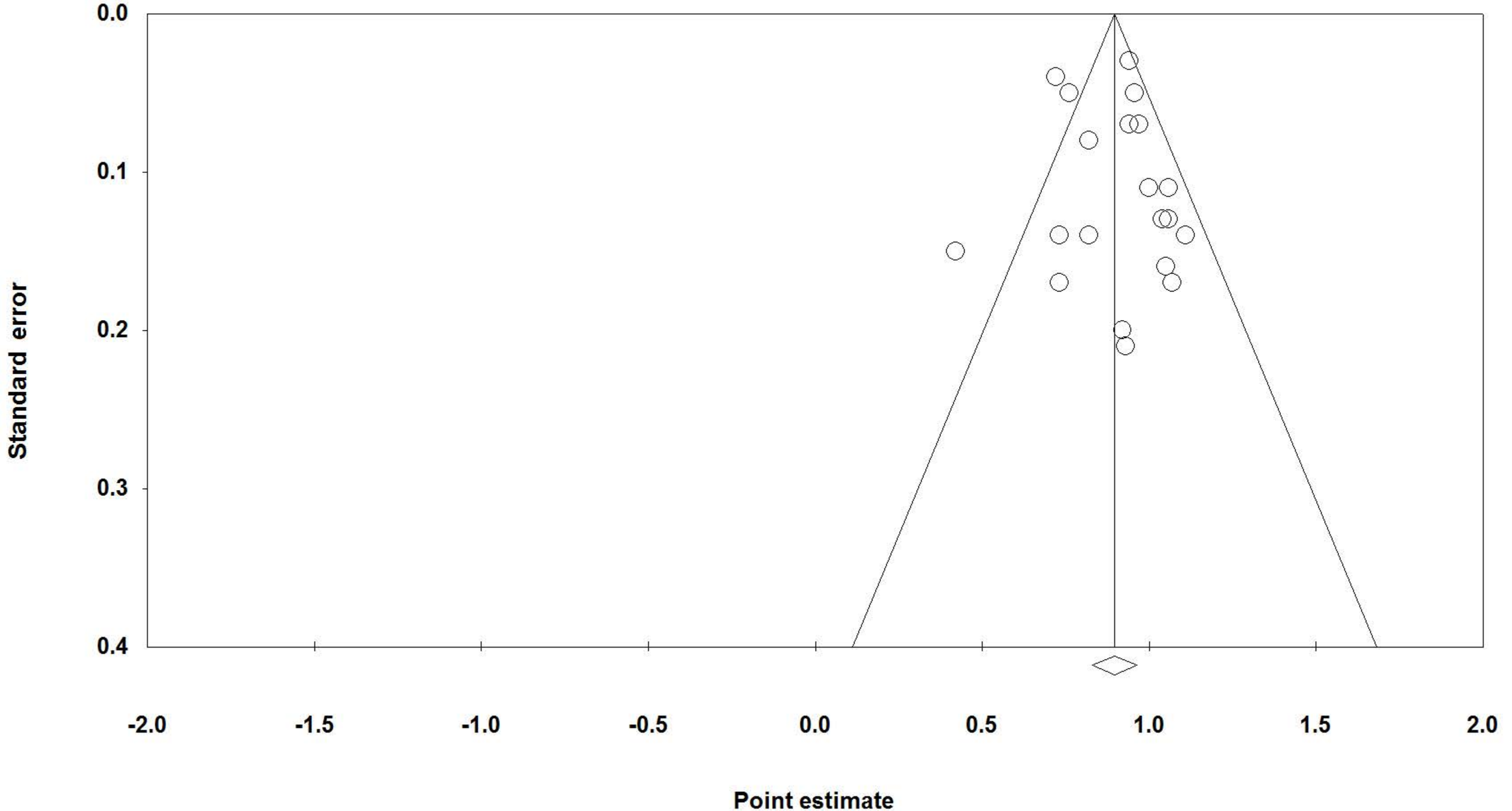


b

Study	Point estimate	Standard error	Lower limit	Upper limit
Amara et al. [32]	0.96	0.05	0.86	1.05
Batterham and Jackson [44]	0.94	0.03	0.88	1.00
Carvalho et al. [47]	0.73	0.14	0.46	1.00
Carvalho et al. [48]	1.11	0.14	0.84	1.38
Chamari et al. [33]	0.72	0.04	0.64	0.80
Davies et al. [34]	1.05	0.16	0.74	1.36
Goosey-Tolfrey et al. [35]	0.82	0.14	0.55	1.09
Heil [24]	0.76	0.05	0.66	0.86
Neder et al. [55]	0.92	0.20	0.53	1.31
Neder et al. [55]	0.73	0.17	0.40	1.06
Nevill et al. [58]	0.94	0.07	0.80	1.08
Segers et al. [36]	1.06	0.13	0.80	1.31
Segers et al. [36]	1.06	0.11	0.84	1.27
Tolfrey et al. [37]	1.00	0.11	0.78	1.22
Valente-Dos-Santos et al. [63]	1.07	0.17	0.74	1.40
Valente-Dos-Santos et al. [63]	0.93	0.21	0.52	1.34
Valente-Dos-Santos et al. [63]	0.42	0.15	0.13	0.71
Vanderburgh and Katch [64]	1.04	0.13	0.79	1.29
Wijndaele et al. [67]	0.82	0.08	0.66	0.98
Wijndaele et al. [67]	0.97	0.07	0.83	1.11
Pooled	**0.90**	**0.03**	**0.83**	**0.96**
95% Prediction interval			**0.68**	**1.12**




0.0
0.1
0.2
0.3
0.4
Standard error
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Point estimate


0.0
0.1
0.2
0.3
0.4
Standard error
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Point estimate