



OPEN ACCESS

Children's fitness and health: an epic scandal of poor methodology, inappropriate statistics, questionable editorial practices and a generation of misinformation

Jo Welsman , Neil Armstrong 

10.1136/bmjebm-2019-111232

Children's Health and Exercise Research Centre, University of Exeter, Exeter, UK

Correspondence to:

Dr Jo Welsman, Children's Health and Exercise Research Centre, University of Exeter, Exeter, UK; j.r.welsman2@exeter.ac.uk

A global explosion of research into children and adolescents' health and cardiorespiratory or aerobic fitness has resulted in a flurry of papers and subsequently systematic reviews revealing apparently worrying but fallacious assumptions such as: (1) aerobic fitness is declining¹; (2) aerobic fitness expressed in ratio with body mass reflects present² and predicts future³ cardiovascular and metabolic health risk; (3) a single sex-specific 'cut-point' of aerobic fitness expressed in ratio with body mass identifies children and adolescents who 'may benefit from primary and secondary cardiovascular prevention programming', (Ruiz *et al* p1451)⁴ - the so-called 'clinical red flags'.

Our serious concerns with these conclusions, despite their basis in large data sets and publication in internationally respected journals, is that they are not founded on rigorous science but on flawed methodology, namely predicting aerobic fitness from the 20metre shuttle run test (20mSRT)⁵ and interpreting paediatric fitness data expressed in ratio with body mass.

Problem 1: the 20mSRT is not a valid measure of children's aerobic fitness

Over 30 years ago⁶ we demonstrated the poor criterion validity of the 20mSRT or 'bleep' test.⁵ We discounted the test as a research tool not only because of poor statistical validity but because of its dependence on participant motivation and body size, particularly fatness. The 20mSRT was never originally validated against laboratory-determined peak oxygen uptake ($\dot{V}O_2$) (the internationally recognised gold-standard measure of paediatric aerobic fitness). Subsequent validation studies with children are sparse and statistically inadequate being based in correlation and regression not agreement. A recent review, although not specifying the underlying statistics, reported that peak $\dot{V}O_2$ can be estimated within ± 10 mL kg^{-1} min^{-1} from the 20mSRT,⁷ but as this represents around 20%–25% of typical values this is hardly a test we would want to see underpinning recommendations for international public health policy.⁸

Problem 2: the expression of aerobic fitness in simple ratio with body mass (ie as VO_2 in mL kg^{-1} min^{-1}) is not a valid method for controlling for body size differences

Over 30 years ago, our attention was drawn to a paper published by Tanner⁹ which detailed the

fallacy of simple division by body mass to control for body size in describing physiological functions. As an assumed, rather than fitted mathematical relationship, per-body-mass ratios typically overestimate values of fitness for light individuals, and artefactually penalise heavier people. Thus, in subsequent correlation analyses, or through subdivision into high vs low fitness groups, for example, to examine relationships with cardio-metabolic risk factors,⁴ spurious conclusions are inevitable and reflect levels of fatness rather than levels of fitness.

Aware of the significance of this paper for our own research, we comprehensively searched the literature but failed to find a published scientific or statistical justification for 'per-body-mass scaling' for youth aerobic fitness.¹⁰ It has become absorbed into accepted practice simply because it is 'traditional',¹¹ 'convenient'¹¹ and 'feasible'¹² and so evades challenge by peer reviewers and editors.

Discussion

The speed at which research studies based on this combination of two fundamentally flawed methodologies have come to dominate the international literature on paediatric aerobic fitness has been alarming. In the decade to 2000, on average two papers reporting 20mSRT data per year were published in journals summarised by PubMed. In the past 9 years, 379 papers have been published. In response to this we have refocused our efforts to raise awareness of the methodological inaccuracies inherent in this body of research and published, with comprehensive commentary and reanalyses, 20 of our published cross-sectional studies¹⁰ and new longitudinal multilevel modelling analyses^{13 14} of ~1400 rigorous determinations of 10–18 years old's aerobic fitness. In all cases the data did not meet the statistical assumptions underpinning ratio scaling of peak $\dot{V}O_2$ with body mass.

Our recent longitudinal studies confirm evidence we first published over 30 years ago: when determined in a laboratory using rigorous assessment procedures, appropriately size-adjusted aerobic fitness *increases* with age and maturity in both girls and boys (eg, 13), that is, does not decline or level off as suggested by per-body-mass international norms.¹⁵ Thus recommendations for single sex-specific 'cut-off' points for 'healthy'



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Welsman J, Armstrong N. *BMJ Evidence-Based Medicine* 2021;26:12–13.

fitness from childhood through adolescence which do not accommodate age or maturational effects⁴ are meaningless.

Rigorously determined laboratory data¹⁶ do not show the declines over time in children's fitness indicated from 20mSRT data. The latter is an artefact due to increased fatness constituting 'dead weight' which increases the work done per shuttle and adversely affects 20mSRT predictions but does not affect true aerobic fitness. This is further confounded by body fat being included in the denominator when simple per body mass ratios are computed. In fact, when body size and fatness differences are appropriately accounted for using allometric multilevel modeling, there are minimal differences in the fitness of overweight versus healthy weight children and adolescents.¹³

But how do we shift an entire discipline rooted in poor methodology? Not surprisingly young researchers and those in resource poor countries are quick to join the international 20mSRT bandwagon which enables the collection of large volumes of data quickly, cheaply and supports publication in internationally respected journals. Publishing appropriately analysed papers,^{13 14} writing tutorial¹⁷ and commentary style pieces is not enough. We are dismayed by apparent editorial resistance to challenges to the status quo. In the face of demonstrably weak methodology and inappropriate statistics we urgently need those with editorial power, including peer reviewers, to challenge authors to defend their work and for that defence to be based in appropriate statistics. We need better mechanisms and mentoring to support researchers in developing economies to discourage 'quick wins' and guide them towards better quality research. We need to ensure that the next generation of researchers are grounded in appropriate methodologies and have the critical ability and confidence to challenge traditional, but unjustified, practices.

We have an ethical and moral duty with minors to ensure that our research methodologies are rigorous and defensible. Only then will we accurately understand the role of fitness in children's current and future health enabling public health recommendations to be meaningful and evidence-based.

Twitter Jo Welsman @jowelsman

Contributors JW and NA jointly conceived the paper. JW drafted the paper and both authors revised the manuscript and approved submission for review.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given,

any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Jo Welsman <http://orcid.org/0000-0002-8877-3926>

Neil Armstrong <http://orcid.org/0000-0002-3086-629X>

References

- Tomkinson GR, Lang JJ, Tremblay MS. Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. *Br J Sports Med* 2019;53:478–86.
- Lang JJ, Belanger K, Poitras V, et al. Systematic review of the relationship between 20m shuttle run performance and health indicators among children and youth. *J Sci Med Sport* 2018;21:383–97.
- Ruiz JR, Castro-Piñero J, Artero EG, et al. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009;43:909–23.
- Ruiz JR, Caverro-Redondo I, Ortega FB, et al. Cardiorespiratory fitness cut points to avoid cardiovascular disease risk in children and adolescents; what level of fitness should raise a red flag? A systematic review and meta-analysis. *Br J Sports Med* 2016;50:1451–8.
- Léger LA, Mercier D, Gadoury C, et al. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci* 1988;6:93–101.
- Armstrong N, Ringham D, Welsman J. Peak oxygen uptake and progressive shuttle run performance in boys aged 11–14 years. *Br J Phys Educ* 1988;19:10–11.
- Tomkinson GR, Lang JJ, Blanchard J, et al. The 20-m shuttle run: assessment and interpretation of data in relation to youth aerobic fitness and health. *Pediatr Exerc Sci* 2019;31:152–63.
- Lang JJ, Wolfe Phillips E, Orpana HM, et al. Field-based measurement of cardiorespiratory fitness to evaluate physical activity interventions. *Bull World Health Organ* 2018;96:794–6.
- Tanner JM. Fallacy of per-weight and per-surface area standards, and their relation to spurious correlation. *J Appl Physiol* 1949;2:1–15.
- Welsman J, Armstrong N. Interpreting aerobic fitness in youth: the fallacy of ratio scaling. *Pediatr Exerc Sci* 2019;31.
- Bar-Or O, Rowland TW. *Pediatric exercise medicine*. Champaign, IL: Human Kinetics, 2004.
- Mintjens S, Menting MD, Daams JG, et al. Cardiorespiratory fitness in childhood and adolescence affects future cardiovascular risk factors: a systematic review of longitudinal studies. *Sports Med* 2018;48:2577–605.
- Armstrong N, Welsman J. Sex-specific longitudinal modeling of youth peak oxygen uptake. *Pediatr Exerc Sci* 2019;31:204–12.
- Armstrong N, Welsman J. Development of peak oxygen uptake from 11–16 years determined using both treadmill and cycle ergometry. *Eur J Appl Physiol* 2019;119:801–12.
- Tomkinson GR, Lang JJ, Tremblay MS, et al. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. *Br J Sports Med* 2017;51:1545–54.
- Mountjoy M, Andersen LB, Armstrong N, et al. International Olympic Committee consensus statement on the health and fitness of young people through physical activity and sport. *Br J Sports Med* 2011;45:839–48.
- Welsman JR, Armstrong N. Interpreting exercise performance data in relation to body size. In: Armstrong N, van Mechelen W, eds. *Paediatric exercise science and medicine*. 2nd ed. Oxford: Oxford University Press, 2008: 13–21.