©Journal of Sports Science and Medicine (2020) **19**, 627-629
http://www.jssm.org

**Letter to editor**

# Scientific Rigour in the Assessment and Interpretation of Youth Cardiopulmonary Fitness: A Response to the Paper '*Normative Reference Values and International Comparisons for the 20-Metre Shuttle Run Test: Analysis of 69,960 Test Results among Chinese Children and Youth*'

**Dear Editor-in-chief**

In a data-rich paper, Zhang et al. (2020) reported 20-metre shuttle run test (20mSRT) scores from 69,960 Chinese school students. From the 20m SRT performance data the authors predicted the students' peak oxygen uptake ($\dot{V}O_2$) ratio-scaled with body mass (i.e. in $mL \cdot kg^{-1} \cdot min^{-1}$). They then estimated the percentages of 9-17-year-olds who fell below the cut-points of 35 $mL \cdot kg^{-1} \cdot min^{-1}$ (girls) and 42 $mL \cdot kg^{-1} \cdot min^{-1}$ (boys), recommended by Ruiz et al. (2016, p. 1451) to identify, '*children and adolescents who may benefit from primary and secondary cardiovascular prevention programming'*. We applaud the intention of the authors to promote the health and well-being of Chinese children and adolescents but strongly caution against the use of fallacious methods of estimating and interpreting the peak $\dot{V}O_2$ of children and adolescents.

First, ratio scaling exercise variables with body mass assumes an underlying set of specific statistical assumptions which are seldom (if ever) met in pediatric exercise studies (Welsman and Armstrong, 2019a). In brief, if ratio scaling effectively controlled for body mass then the product-moment correlation coefficient between peak $\dot{V}O_2$ (in $mL \cdot kg^{-1} \cdot min^{-1}$) and body mass (in kg) would not be significantly different from zero. It has been unequivocally demonstrated on numerous occasions that, ratio-scaled values of the peak $\dot{V}O_2$ of children and adolescents remain significantly and negatively correlated with body mass (Armstrong and Welsman, 2020; Welsman and Armstrong, 2019a). Ratio scaling peak $\dot{V}O_2$ with body mass therefore favours lighter (e.g. clinically underweight, younger, or later maturing) and penalizes heavier (e.g. overweight, older, or earlier maturing) youth.

Second, the 20mSRT is not, '*an effective measure of peak oxygen uptake'* (Zang et al., 2020, p. 478). 20mSRT performance is a function of an individual's willingness and capability to transport their body mass between two lines 20 m apart while keeping pace with audio signals which require running speed to increase each minute. The specious logic underpinning 20mSRT performance as a surrogate of youth peak $\dot{V}O_2$ has been extensively documented. Suffice to summarize herein that it has been reported that, i) over 50% of reported correlation coefficients between children's 20mSRT predicted peak $\dot{V}O_2$ and laboratory-determined peak $\dot{V}O_2$ explain less than half the shared variance (Mayorga-Vega et al., 2015); ii) for 9 to 17-year-olds, *'the 95% likely range for a true peak $\dot{V}O_2$ value estimated from the 20mSRT is ~10 $mL \cdot kg^{-1} \cdot min^{-1}$ or ~24%'* (Tomkinson et al., 2019, p.154); iii) the limits of agreement of directly determined peak $\dot{V}O_2$ and 20mSRT predicted peak $\dot{V}O_2$ are only within ~40% (Welsman and Armstrong, 2019b); and, iv) as body fat is metabolically inert, in a 20mSRT excess fat is carried as 'deadweight' which increases the total work done in each 20 m shuttle, negatively affects 20mSRT performance, and lowers the 20mSRT prediction of peak $\dot{V}O_2$ without influencing true peak $\dot{V}O_2$ (Armstrong and Welsman, 2020). Moreover, peak $\dot{V}O_2$ predicted from 20mSRT performance is expressed in $mL \cdot kg^{-1} \cdot min^{-1}$, and therefore subject to all the flaws associated with ratio scaling. In a 20mSRT, overweight youth are doubly penalized by not only having to carry their metabolically inert fat mass as 'deadweight' during a 20mSRT but also having their performance score expressed as peak $\dot{V}O_2$ divided by body mass (including fat mass).

Third, as noted earlier, ratio-scaled peak $\dot{V}O_2$ (i.e. in $mL \cdot kg^{-1} \cdot min^{-1}$) is not a body mass-free variable and remains correlated with body mass, therefore, when used in subsequent correlational analyses with other health-related variables correlated with body mass it will produce spurious correlations. For example, associations of ratio-scaled peak $\dot{V}O_2$ with cardiovascular risk factors in overweight youth are likely to reflect overweight (or over fatness) rather than cardiopulmonary fitness (Loftin et al., 2016).

Fourth, in childhood and adolescence, peak $\dot{V}O_2$ develops in accord with age- and maturity status-driven, concurrent changes in morphological, cardiopulmonary, and intra-muscular covariates, with the timing and tempo of changes governed by individual biological clocks (Armstrong and Welsman, 2020). The use of fixed cut-points based on a single value of peak $\dot{V}O_2$ ratio-scaled with body mass, to classify the cardiometabolic health of 9-17-year-old children and adolescents is, therefore, in direct conflict with current understanding of the development of pediatric cardiopulmonary fitness. The cardiopulmonary fitness of an 9-year-old, pre-pubertal female with a predicted peak $\dot{V}O_2$ of 35 $mL \cdot kg^{-1} \cdot min^{-1}$ is very different from that of an 17-year-old, post-pubertal female with the same predicted ratio-scaled peak $\dot{V}O_2$. To identify pre-pubertal, pubertal, and post-pubertal 9-17-year-olds for '*primary and secondary cardiovascular prevention programming'* on the basis of the same 20mSRT predicted ratio-scaled value of peak $\dot{V}O_2$ is not tenable and if adopted has the potential to adversely affect the health and well-being of some young people.

We congratulate Zhang et al. (2020) for raising the profile of pediatric health and well-being in China. The authors acknowledge the important limitation in their paper

Received: 05 June 2020 / Accepted: 17 June 2020 / Published (online): 01 September 2020

of not considering growth and maturation and note the need for longitudinal studies rather than cross-sectional 'snapshots' of youth cardiopulmonary fitness. We urge researchers in China and elsewhere, to embrace these vital issues and to apply rigorous scientific methodology to the determination, assessment, and interpretation of the cardiopulmonary fitness of children and adolescents.

**Neil Armstrong ✉ and Jo Welsman**
Children's Health and Exercise Research Centre, University of Exeter, Exeter, United Kingdom

## References

Armstrong, N. and Welsman, J. (2020) Traditional and new perspectives on youth cardiorespiratory fitness. *Medicine and Science in Sports and Exercise* **52,** (in press).

Loftin, M., Sothern, M., Abe, T. and Bonis, M. (2016) Expression of $\dot{V}O_{2peak}$ in children and youth with special reference to allometric scaling. *Sports Medicine* **46,** 1451-1460.

Mayorga-Vega, D., Aguila-Solo, P. and Viciana, J. (2015) Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *Journal of Sports Science and Medicine* **14,** 536-547.

Ruiz, J.R., Cavero-Redondo, I., Ortega, F.B., Welk, G.J., Andersen, L.B. and Martinez-Vizcaino, V. (2016) 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. *British Journal of Sports Medicine* **50,** 1451-1458.

Tomkinson, G.R., Lang, J.J., Blanchard, J., Leger, L. and Tremblay, M.S. (2019) The 20-m shuttle run: assessment and interpretation of data in relation to youth aerobic fitness and health. *Pediatric Exercise Science* **31,** 152-163.

Welsman, J. and Armstrong, N. (2019a) Interpreting aerobic fitness in youth: The fallacy of ratio scaling. *Pediatric Exercise Science* **32,** 184-190.

Welsman, J. and Armstrong, N. (2019b) The 20-metre shuttle run is not a valid test of cardiorespiratory fitness in boys aged 11-14 years. *British Medical Journal Open Sports and Exercise Medicine* **5,** e000627.

Zhang, F., Yin, X., Bi, C., Li, Y., Sun, Y., Zhang, T., Yang, X., Li, M., Lui, Y., Cao, J., Yang, Y. and Song, Ge. (2020) Normative reference values and international comparisons for the 20-metre shuttle run test: analysis of 69,960 test results among Chinese children and youth. *Journal of Sports Science and Medicine* **19,** 478-488.

✉ **Prof. Neil Armstrong, PhD, DSc**
Children's Health and Exercise Research Centre, University of Exeter, Exeter, EX1 2LU, United Kingdom
**E-mail:** N.Armstrong@exeter.ac.uk

---

# Authors' response

**Dear Editor-in-chief**

As mentioned by Dr. Armstrong and Welsman, we used rich and nationally representative data intended to promote the health and well-being of Chinese children and adolescents. While as with any study, there were some limitations in our paper. In their editorial, Dr. Armstrong and Welsman expressed concerns on our paper and highlighted the need for scientific rigor in the assessment and interpretation of youth cardiopulmonary fitness. We appreciate their interest and are grateful for the opportunity to clarify each of the topics they raised as highlighted below:

1. ***Ratio scaling $VO_{2peak}$ with body mass favors lighter and penalizes heavier youth***. We learn from Armstrong and Welsmans' paper (2019) that allometric scaling can help correct the limitations of traditional ratio-scaled $VO_{2peak}$ which fails independent of body mass. While body mass may not be the best variable to descript body size since fat-free mass, is arguably a more appropriate scaling denominator for $VO_{2peak}$ on physiological grounds (Weibel et al., 2005). Moreover, the value of the size exponent for both body mass and fat-free mass were also controversial (Lolli et al., 2017). In general, there is a lack of a universally applicable alternative for ratio scaling $VO_{2peak}$ with body mass (Welsman and Armstrong, 2019). Despite this, we have also reported 20mSRT results in measuring units such as the number of laps, stages, peak running speed, which can also provide references for Chinese children and adolescents.

2. ***20m shuttle run test (20m SRT) is not an effective measure of peak oxygen uptake.*** Indeed, the laboratory-based maximal exercise test by exercising until voluntary exhaustion with direct measurement of the $VO_{2peak}$ is the most reliable assessment. Population health research requires large samples to ensure representativeness, while the necessity of sophisticated and costly instrumentation, qualified technicians precluding the directly measured $VO_{2peak}$ in schools and large-scale research studies (Pescatello et al., 2014). In contrast, the 20mSRT has good feasibility, utility, and scalability for population health surveillance to monitor trends. Although its criterion validity as moderate for estimating peak $VO_{2peak}$ in youth (Mayorga-Vega et al., 2015), it is routinely used as a preferred option to estimate CRF in large-scale research studies because of the low cost of equipment, and its ability to test large groups of students, simultaneously (Tomkinson et al., 2017).

3. ***Ratio-scaled* $VO_{2peak}$ *(i.e. in mL·kg$^{-1}$·min$^{-1}$) is not a body mass-free variable and the will produce spurious correlations when used in subsequent correlational analyses with other health-related variables correlated with body mass.*** Thank you for Dr. Armstrong and Welsman's insightful comments and remanding. Since we didn't involve correlation analysis in the present study, we will be cautious when conducting subsequent correlational analyses.

4. ***The use of fixed cut-points based on a single value of* $VO_{2peak}$ *ratio-scaled with body mass in direct conflict with the current understanding of the development of pediatric cardiopulmonary fitness***. Indeed, $VO_{2peak}$ develops in accord with age- and maturity status-driven. The use of fixed cut-points to assess CRE of 9-17 years old children and adolescents has its limitation (Armstrong and Welsman, 2020). It is perfect that if there was age- and sex-specific cut-point to estimate healthy CRE. Unfortunately, to our best knowledge, we don't find this type of cut-point. As the purpose of our study was to estimate the prevalence of healthy CRE from an epidemiological perspective. The use of cut-points of 35

mL·kg$^{-1}$·min$^{-1}$ (girls) and 42 mL·kg$^{-1}$·min$^{-1}$ (boys), recommended by Ruiz et al. (2016) was also an alternative to estimate the prevalence of healthy CRE for Chinese children and adolescents, which has also been used to estimate the prevalence of healthy CRE for 1 .1 million international children and youth (Lang et al., 2019).

We appreciate the insightful comments and observations by Dr. Armstrong and Welsman regarding our paper. We have taken into account the strengths of the study, as well as further explaining the limitations in our responses. The suggestions and comments raised are important aspects, and we will be cautious and rigor in the assessment and interpretation of cardiopulmonary fitness for Chinese children and adolescents.

**Feng Zhang [1,2], Xiaojian Yin [1,2,3]✉, Cunjian Bi [1,2], Yuqiang Li [1,2], Yi Sun [1,2], Ting Zhang [1,2], Xiaofang Yang [1,2], Ming Li [1,2], Yuan Liu [1,2], Junfang Cao [1,2], Ting Yang [1,2], Yaru Guo [1,2] and Ge Song [1,2]**

[1] Key Laboratory of Adolescent Health Assessment and Exercise Intervention of Ministry of Education, East China Normal University, Shanghai, China; [2] College of Physical Education and Health, East China Normal University, Shanghai, China; [3] College of Economics and Management, Shanghai Institute of Technology, Shanghai, China

✉ **Xiaojian Yin**

Key Laboratory of Adolescent Health Assessment and Exercise Intervention of Ministry of Education, East China Normal University, Shanghai, China

**E-mail:** xjyin1965@163.com

## Reference

Armstrong, N. and Welsman, J. (2020) Traditional and new perspectives on youth cardiorespiratory fitness. *Medicine and Science in Sports and Exercise* **52,** (in press).

Lang, J.J.; Tremblay, M.S.; Ortega, F.B.; Ruiz, J.R.; Tomkinson, G.R. (2019) Review of criterion-referenced standards for cardiorespiratory fitness: what percentage of 1 142 026 international children and youth are apparently healthy? *British Journal of Sports Medicine* **53,** 953-958. Lolli, L.; Batterham, A.M.; Weston, K.L.; Atkinson, G. (2017) Size Exponents for Scaling Maximal Oxygen Uptake in Over 6500 Humans: A Systematic Review and Meta-Analysis. *Sports Medicine* **47,** 1405-1419.

Mayorga-Vega, D., Aguila-Solo, P. and Viciana, J. (2015) Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *Journal of Sports Science and Medicine* **14,** 536-547.

Pescatello, L.S., Arena, R., Riebe, D. and Thompson, P.D. (2014) *ACSM's guidelines for exercise testing and prescription*. 9th edition. Wolters Kluwer/ Lippincott Williams & Wilkins, Philadelphia.

Ruiz, J.R., Cavero-Redondo, I., Ortega, F.B., Welk, G.J., Andersen, L.B. and Martinez-Vizcaino, V. (2016) 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. *British Journal of Sports Medicine* **50,** 1451-1458.

Tomkinson, G.R., Lang, J.J., Tremblay, M.S., Dale, M.; LeBlanc, A.G., Belanger, K., Ortega, F.B. and Leger, L. (2017) International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. *British Journal of Sports Medicine* **51,** 1545-1554,

Welsman, J.; Armstrong, N. (2019) Interpreting Aerobic Fitness in Youth: The Fallacy of Ratio Scaling. *Pediatric Exercise Science* **31,** 184-190,

Weibel, E.R. and Hoppeler, H. (2005) Exercise-induced maximal metabolic rate scales with muscle aerobic capacity. *Journal of Experimental Biology* **208(Pt 9),**1635-1644.