Section: Public Health Practice

Article Title: 'Fitness Testing' for Children: Let's Mount the Zebra!

Authors: Daniel Cohen[1], Christine Voss[2], and Gavin RH Sandercock[3]

Affiliations: [1]Faculty de Medecina, Universidad de Santander (UDES), Bucaramanga, Santander, Colombia. [2]Centre for Hip Health and Mobility, University of British Columbia, Vancouver, BC, Canada. [3]Department of Biological Sciences, University of Essex, Colchester, Essex, United Kingdom.

Running Head: Let's mount the zebra

Journal: *Journal of Physical Activity & Health*

Acceptance Date: March 7, 2014

DOI: http://dx.doi.org/10.1123/jpah.2013-0345

"'Fitness Testing' for Children: Let's Mount the Zebra!" by Cohen D, Voss C, Sandercock G RH
Journal of Physical Activity & Health

Title: 'Fitness testing' for children: let's mount the zebra!

Running Head: Let's mount the zebra

Article type: Public Health Practice

Manuscript word count: 3729

Abstract

Few paediatric health topics have sparked as much academic and public debate as school-based fitness testing. Since Rowland exclaimed nearly two decades ago that the "horse", referring to school-based fitness testing is dead, opinion has been divided – to test or not to test. Whilst many agreed with Rowland`s criticisms, others suggest that it is not school-based fitness testing per se that is problematic but that we have been riding the wrong animal, and should instead be riding a zebra – signifying a multi-dimensional approach to its implementation. We acknowledge concerns over school-based fitness testing, but argue that the associations between fitness and health, as well as the secular declines in fitness, necessitate such monitoring. More importantly, we highlight several potential opportunities for fitness testing, to not only to map an important aspect of health, but also to improve physical self-concept and challenge the misconception that leanness equates to good health and fitness and its corollary; that fatness is invariably associated with poor fitness and health. We believe that a carefully chosen fitness test battery delivered in an educational context, can transform the horse into the zebra, and it is time for the skilful rider to mount and ride it. (198 words)

Keywords: Pediatrics, Aerobic, Measurement, Physical Fitness, Strength

Introduction

Childhood obesity has increased greatly over the past few decades, whilst children's fitness levels have been declining worldwide.[1-7] Physical fitness, is a multi-dimensional construct that includes skill and health related components of which cardio-respiratory fitness (CRF) and muscular fitness in particular are powerful determinants of health in youth.[8] The association between fitness and health is well documented for CRF, whereby good CRF is protective against cardio-metabolic risk factors across BMI/fatness categories.[9-13] Muscular fitness is also inversely associated with metabolic risk[9,11] and is a valuable part of health monitoring in children.[8] Specifically, poor muscular fitness is associated with elevated cardio-metabolic risk factors in adolescence[9,12] and an increased risk of developing obesity,[14] cardiovascular disease[15] and with cardiovascular and total mortality.[16] This effect appears independent of the associations between metabolic health and low CRF.[9,12]

Given such strong, independent associations between fitness and health, declines in paediatric fitness are worrying from a public health perspective and underpin the former UK Chief Medical Officer's (CMO) recommendation for routine fitness testing in schools.[17] The CMO stated that: *"The introduction of a standardised school-based fitness assessment in England may have multiple benefits that extend beyond the benefits for the individual."* These benefits include: lowering the lifetime risk of six diseases, building a lifelong habit of participation in physical activity, higher educational achievement, maintaining a healthy weight, as well as improving social and mental wellbeing.[16] Field-tests of fitness, such as Leger's 20 m shuttle run test[18] (or "*bleep test*", *"beep test", "PACER", "multistage fitness test"*) used to estimate CRF, as well as hand grip strength, jump performance and/or trunk muscular endurance to assess muscular fitness, particularly lend themselves to school-based assessments due to relatively low space and equipment requirements. More importantly, because they are often already, or can be easily integrated into Physical Education (PE) lessons. In part, the UK CMO's recommendation for school-based fitness testing builds upon a variety of well-reported school-based fitness projects, and the wealth of information that these have yielded. For example, the

FITNESSGRAM® fitness testing battery is implemented annually in a number of US states[19] and in New York City public schools,[20] while the ALPHA fitness testing battery based on the pan-European HELENA and Spanish AVENA studies has been successfully piloted in Spanish schools.[21] In the UK, in addition to being implemented sporadically in schools,[22] youth fitness testing is implemented regularly by fitness professionals in conjunction with academics from Liverpool John Moores University, as part of the on-going research and health promotion programme: SportsLinx.[3,23] More recently, our group from the University of Essex launched the East of England Healthy Hearts Study including comprehensive fitness testing for 10-16 year olds during PE classes at schools in London, Essex, Suffolk and Bedfordshire (UK counties), allowing the development of various fitness test norms[24-26] and revealing declining trends in fitness[2,4] and a number of mediating factors.[27-30] Furthermore, the Texas Youth Study based on fitness assessment of over 2.5 million children revealed positive associations between fitness and academic achievement, school attendance, several psychosocial measures and negative associations with indicators of delinquency.[31] While much of these data are not new, one should not ignore the potential value of media coverage of these physical and cognitive or behavioural correlates of fitness and highlighting regional differences[32] as leverage to promote governmental action on physical activity (PA).[31]

While few would argue against the value of monitoring trends in children's health, routine school-based fitness testing is not greeted with approval by all. The Association for Physical Education and parent groups criticised such plans for the UK, echoing concerns about fitness testing in children aired in the academic literature.[21,33] One specific concern is the potential for fitness testing to be an unpleasant and embarrassing experience for overweight children,[34] with possibly reinforcing poor physical self-concept. While empirical examination of children's experience of fitness testing is under-researched,[35] there is indeed some evidence to suggest negative psychosocial consequences, such as embarrassment or teasing by other children.[36,37] Since physical self-concept is a determinant of PA,[38] this is indeed an important concern for school-based fitness testing. In addition, fitness testing is also criticised for not promoting PA *per se*,[21] yet even its strongest advocates do not regard this as an expected outcome of the process.[32] Similarly, the National Child Measurement Programme, which assesses BMI in UK schoolchildren, does not cite weight loss as an objective of the measurement process but rather: *"to inform local planning and delivery of services for children."*[39] Another important point for consideration, and one which has previously been alluded to in attempts to advance the debate around fitness testing for children, is that while assessment of academic progress is almost universally accepted, there remains strong resistance to school-based health and fitness assessments.[40]

Fitness testing as a zebra

Based on the aforementioned concerns, as well as additional concerns regarding fitness tests validity, use of criterion standards[33] or inappropriate use of data,[41] a number of authors have challenged the value of youth fitness testing and likened it to a dead horse,[41] from which "*it is time to dismount*".[21] In keeping with the metaphor of a horse, others have defended youth fitness testing, asserting that its value depends on the skill and training of the riders or the understanding that is not a horse, but a zebra that is being ridden.[40] The zebra's stripes symbolise the potential for fitness testing to have multiple applications and outcomes, including a role in the development of: *"physical literacy, a multidimensional and interactive construct."*[40]

We argue that the principal aim of youth fitness testing, which is the mapping and tracking of trends in fitness, is of substantial value and provides data beyond that of BMI alone. This contributes to the surveillance of youth health and to the evaluation of the effectiveness of PA promotion strategies[16,32,42] and can inform school curricula or policies.[32] While fitness testing itself may not promote increased PA and fitness, improvements in fitness measures could detect increases in the quantity and quality of school PE[43] or PA overall, and provide an non-invasive objective assessment of the physiological impact of interventions. Both cross-sectional and longitudinal fitness data may also provide convincing empirical data and thus impetus for legislative action to support programmes aimed at PA promotion and obesity prevention.[42]

Further, we believe that the delivery and contextualisation of youth fitness testing and provision of feedback, or *"the skill of the rider"*, could transform the horse into the zebra which can achieve improved health, as well as psychosocial and educational outcomes. We argue that youth fitness testing can be a vehicle both for the promotion of PA and the development of physical literacy and it can do so without undermining physical self-concept in overweight children. However, these outcomes are not an inevitable consequence of fitness testing and we urge for diligence when selecting fitness test batteries, as well as when providing feedback.

We now highlight issues which we believe are important for the design and delivery of youth fitness testing, with the aim of achieving objectives beyond surveillance while addressing concerns about negative psychosocial outcomes and lack of contextualisation.

Fitness testing and physical self-concept

We acknowledge that fitness testing has the potential to negatively impact on the physical self-concept of the overweight (or unfit) child. Although there is little empirical evidence to support this notion,[35] there are limited reports of negative psychosocial consequences of fitness testing, such as embarrassment or teasing by other children.[36] More specifically, a survey of over 2500 PE teachers, who administered the FITNESSGRAM® in Texas, revealed that approximately 25% reported observing some negative consequences of fitness testing,[37] while in another report, few PE teachers reported such incidents.[36] Since physical self-concept is a determinant of PA,[38] we emphasize that it is essential to minimise such negative experiences and ensure that for less active or overweight children – those for whom increased levels of PA and fitness are most important – fitness testing does not undermine physical-concept. We further believe that this outcome is not an unavoidable consequence of the process and we argue that fitness testing could, on the contrary, be a means to *enhance* the physical self-concept in such children.

Fitness testing can also be a vehicle used to challenge pervasive and potentially harmful assumptions about the relationship between body size, health and fitness. In particular, it challenges the notion that being overweight is incompatible with being fit and healthy, while thinness equates to good health and fitness. For example, Bacon and Aphramor[44] recently reported that 51% of overweight adults had 'normal' cardio-metabolic health, whereas 24% of normal weight adults had 'abnormal' cardio-metabolic health, illustrating the limitations associated with using BMI alone as a proxy for health. Similarly, evidence from the Aerobics Longitudinal Study[45] long ago introduced the 'fat-fit' phenomenon, whereby overweight but fit individuals had a lower risk of cardiovascular mortality than normal weight but unfit individuals. This phenomenon is also observed in youth, where higher fitness is associated with lower cardio-metabolic risk factors in children classified as normal and overweight by BMI.[9-11] Furthermore, Parrett et al found that in children with body composition assessed by DEXA, higher % fat children with high aerobic fitness had significantly lower metabolic risk score than their low fitness, higher % fat counterparts.

Indeed, based on the work of Blair and others[45-48] it is now widely agreed that differences in fitness and PA contribute greatly to the variability in cardio-metabolic health observed at any given BMI. To further illustrate this point, we cross-tabulated the prevalence of children with varying levels of CRF according to BMI weight status from our East of England Healthy Hearts Study dataset (Figure 1, Panel A). As expected, low CRF was typically higher in overweight and obese children compared with normal-weight children; yet, our data also revealed that ~ 50% of overweight and a smaller proportion of obese children were in the middle and top tertile of CRF. By not testing them, we deny 'fat-fit' children an opportunity to demonstrate their adequate levels of CRF and thus positively reinforce the healthy PA behaviours that presumably underlie this trait. In addition, the notion that fitness testing does not provide additional information to BMI reinforces the false assumption that normal weight or thinness equates to good fitness, or indeed health. According to our findings, the proportion of unfit normal weight children is lower than the proportion of unfit overweight children (Figure 1, Panel A), but overall, there was a greater *number* of normal weight unfit than overweight unfit children (n=855 vs. n=615). This further highlights the importance of communicating to normal weight children, and perhaps more importantly their parents, that their levels of PA or fitness are inadequate. To add to this, we previously reported significant secular declines in CRF[2] and muscular fitness[4] in English 10-yr olds over a 10-year period, even though BMI remained stable; these data further call into question the usefulness of measuring BMI alone to monitor youth health.

The importance of muscular fitness tests

It appears that in the debate in the UK over youth fitness and fitness testing, 'fitness' has become synonymous with cardio-respiratory fitness (CRF), with critics referring to the *ad hoc* way in which the 20 m shuttle run is implemented in British schools.[33] In this context, the multidimensional nature of fitness cannot be emphasised enough and an opportunity to give overweight children fitness-related positive feedback is potentially missed if only CRF and weight-bearing tests of muscular fitness are administered. Overweight children tend to outperform normal weight children in non-weight bearing tests of strength and power, such as hand grip strength[49] (Figure 1, Panel B), medicine ball throw[50] or in weight bearing strength and power tests when performance is expressed with relative values (Figure 1, Panel C).

Morano et al.[50] reported a bolstering of physical self-esteem in overweight children when tested with the medicine ball throw; a measure of upper body power. Anecdotally we can confirm observing a similar effect in the East of England Healthy Hearts Study for hand grip strength testing. Indeed, muscular strength is one of the only sub-components of physical self-concept scales on which overweight children rate themselves highly.[50] Highlighting these pupils' capabilities in this specific aspect of fitness through testing could be a way to positively reinforce a specific "*athletic identity*". This concept (athletic identity) predicts sports participation and PA levels independently of gender, body mass, ethnicity[51] and perceived task competence.[52] The inclusion of non-weight bearing strength and power tests increases the likelihood of good physical performances and is critical to ensuring that children across the range of BMI values receive positive feedback on physical performance. This approach also more accurately reflects the multi-dimensional nature of fitness and athletic performance, enshrined in such sporting events as the Olympic Games where the athlete comes in a range of body shapes and sizes.

To conceptualise this point we have created an arbitrary 'athletic index' score; a composite score of CRF (20 m shuttle run test z-score),[53] strength (absolute hand grip z-score),[24] vertical jump height (countermovement jump z-score)[26] and peak power (z-score, derived from vertical jump height and body mass)[26] for participants in the East of England Healthy Hearts Study. There was no significant difference in the prevalence of upper tertile athletic index between normal and overweight children (Figure 1, Panel D), although the contribution of each athletic index component differed by IOTF weight category (Figure 1, Panel A-C). On this basis, the inclusion of muscular fitness tests not only improves the social and educational experience for all children - and especially those who are overweight - but importantly, it also provides a more complete evaluation of health-related fitness, given the accumulating evidence associating muscular fitness with cardio-metabolic health.[9,14,15,54]

Feedback and expression of data

In addition to the importance of carefully selecting the fitness testing battery, we also argue that careful consideration must be given to how the results are communicated to the child. Domangue and Solmon[35] highlighted this when examining the association between 5th grade children's motivational constructs and attaining awards based on achieving norm-referenced fitness standards. Achievement goals, intrinsic motivation and future intention to participate were significantly lower in those who did not receive the awards. Domangue and Solmon[35] suggested that this would serve to accentuate gaps in fitness between high- and low-performing children by influencing motivation to participate in sports. Implicitly, these findings also demonstrate the potential for positive fitness test feedback as a tool to enhance intrinsic motivation for PA. Positive feedback and fitness awards based on personal improvement rather than attainment of criterion referenced standards, might provide an alternative and more achievable source of PA related positive feedback. This may be particularly important in promoting physical self-competence in those children with negative experiences in team sports, which has been identified as a predictor of reduced PA levels.[55]

While it is evident that fewer overweight children are high performers in weight-bearing fitness tests,[49,56] results can be expressed in a number of ways. For example, in the assessment of CRF using the 20 m shuttle-run test, performance may be reported as levels or shuttles achieved, or as relative $\dot{V}O_{2peak}$ predicted based on the number of shuttles completed (or corresponding running speed).[17,49] Figure 2 compares a variety of performance values of one normal weight and one overweight 13 year-old boy who participated in the East of England Healthy Hearts Study. When their CRF test results were expressed as shuttle count, the normal weight boy completed more than twice as many shuttle laps as the overweight boy (74 shuttles (or level 8.9) vs. 34 shuttles (or level 5.2), respectively). Unsurprisingly, predicted relative $\dot{V}O_{2peak}$[57] was also higher in the normal weight boy (45 $ml \cdot kg^{-1} \cdot min^{-1}$) compared with the overweight boy (39 $ml \cdot kg^{-1} \cdot min^{-1}$); however, when relative $\dot{V}O_{2peak}$ values were converted to absolute $\dot{V}O_{2peak}$ ($l \cdot min^{-1}$) based on the individuals' body mass (kg), the normal and overweight boys achieved $\dot{V}O_{2peak}$ values of 2.91 $l \cdot min^{-1}$ and 2.82 $l \cdot min^{-1}$, respectively. These comparably small differences between the normal weight and overweight child are in agreement with previous work showing that after adjusting for fat mass, there were non-significant differences in 20 m shuttle-run test performance between overweight and normal weight girls and higher performances in overweight boys [49]. It also indicates that the overweight child is not necessarily characterised by a lower capacity for aerobic energy production, but their greater body mass limits their ability to express this in weight-bearing tests such as the 20 m shuttle-run test. Similarly, Figure 2 also compares jump height, as well as peak power, of the same normal weight and overweight 13 year-old boys who participated in the East of England Healthy Hearts Study. Peak power is calculated from vertical jump height, but is relative to body mass; [26] although the overweight boy's vertical jump height was lower than that of the normal weight boy (38 cm vs. 33 cm, respectively), his absolute peak power was greater (2540 W vs. 2615 W, respectively). Offering feedback in absolute as well as relative values, thus offers another potential avenue for providing more positive feedback.

What is more, the underlying calculations to determine relative as well as total performance outcomes could be performed by the children themselves within PE, maths, science and/or personal social education class and would enhance the educational value of fitness testing. Involving pupils in this process and developing their understanding of the interaction between different fitness measures, body size and physical capacity, as well as the underlying biology, serves several purposes. Apart from addressing the criticism of lack of context associated with the administration of fitness testing in schools ,[21] increased scientific content may also increase pupils' engagement in PE. Trudeau and Shephard[55] suggested that a declining interest in PE in adolescence, particularly in girls, is associated more with the low academic prestige of the subject than with a lack of interest in PE *per se*. Developing the scientific content around fitness testing to develop pupils' understanding of how these measures relate to underlying biology, health and physical performance could enhance the status and interest in PE and be used to increase engagement in science. Indeed, under their remit for engagement with science, the Wellcome Trust[58] launched the *"in the zone"* initiative; coinciding with the 2012 Olympic and Paralympic games, this initiative offers age-appropriate curricula materials for schools that combine physical performance and physiological measures. Similarly, the Active Science Scheme[59], which has been piloted in some US schools, also involves pupils measuring, calculating and understanding their PA levels by downloading and interpreting their own pedometer data.

Why not simply measure Physical Activity?

It has previously been argued that instead of testing fitness, efforts should be redirected to the measurement of PA.[33] Yet the methodological difficulties of measuring PA in youth are well documented; self-reported PA is subject to questionable recall ability in youth,[60] whereas the objective assessment of PA via accelerometry is still relatively expensive and time consuming, and equally not free from measurement limitations (i.e. cycling). Anecdotally, we have observed that adherence to the accelerometer wear-time protocol in school-based studies is comparably poor, thereby not only potentially introducing bias, but also calling into question the usefulness of accelerometry as a tool to monitor youth health at the population level. More problematic still is the ongoing scientific debate over various technical issues related to accelerometry, such as appropriate choice of manufacturer and model, epoch length, inclusion criteria for what constitutes a valid file, and ultimately what accelerometry count cut-points offer meaningful interpretation and information on PA.[61] In contrast, the feasibility and reliability of standardised large scale and school-based fitness assessment by teachers or other trained individuals is well documented in the UK and internationally.[3,24,62,63] Importantly, while fitness and PA are independently associated with health,[12,64] studies in adults[65] and children[54] suggest that fitness is more strongly associated with cardio-metabolic risk than PA is. Given these health-associations, we argue that any cross-sectional, longitudinal or secular assessments of youth health should include a measure of CRF as well as at least one component of muscular fitness.

Conclusion

In the UK and many other countries, youth health monitoring, if carried out at all, relies on assessments of BMI. Yet, secular trends in BMI may mask a concurrent loss of lean mass and gain in fat mass,[4,66,67] and furthermore, may not accurately describe population health[8,10,49] or secular trends thereof.[2,4] Irrespectively, an exclusive emphasis on weight (or BMI) in children is of concern for various obvious reasons. We argue that the troublesome decline in health-related fitness[1-7] is a major public health concern, and urge to give the UK's former CMO's recommendation for routine fitness testing in schools due consideration.[16]

All the same, we agree that valid concerns have been raised for the vast potential for negative experiences of fitness testing, particularly for the overweight and/or less fit individuals; poor performance, in the wrong context, can be detrimental to physical self-concept,[36,37] and reinforce low PA.[21,35] However, we believe that this is not an inevitable consequence of fitness testing and depends on the skills and knowledge of the test administrators, as well as the context in which fitness testing is carried out. Indeed, it is thanks to the ongoing debate - academic and public - that we can suggest ways by which to overcome concerns, and furthermore, recognise the opportunities for effective and engaging health education, as well as health- and PA promotion that comes with school-based fitness testing.

For effective school-based testing, we suggest a multi-disciplinary approach involving educators, health professionals and academics from sport science and psychology, all of which maintain an up to date understanding of the association between various components of fitness, body composition and health. The large scale school-based fitness testing conducted by UK Universities of Essex and Liverpool John Moores, which involve sports science (or similar) under- and graduate students, as well as fitness industry trainees, as part of the vocational/research elements of their training, provide an economical model of large-scale implementation, that also minimises additional demands on schools and PE teachers. More importantly, perhaps, is that establishing such university-school relationships then provides a framework for the development of health- and PA promotion programmes, for which schools are an ideal setting.[68] There needs to be effective communication of the documented associations between health, CRF and muscular fitness to physical educators and health professionals working with children as part of initial or continuing training, an understanding which in turn would be passed on to parents and children. Such knowledge is essential for successful implementation of fitness testing and health-promotion through physical activity. There is clear need for a greater emphasis on routes to health that are not directed solely by measurements of body weight, where an over-emphasis on weight must be considered alongside the risk of promoting the development of eating disorders.

In summary, we strongly believe that a carefully chosen fitness test battery, which is delivered in an educational context, can transform the horse into the zebra, and it is time for the skilful rider to mount and ride it.

Acknowledgments:

We wish to thank all of the pupils and PE teachers who participated in the East of England Healthy Heart Study.

Funding source:

None

References

1. Boddy LM, Hackett AF, Stratton G. Changes in fitness, body mass index and obesity in 9-10 year olds. *J Hum Nutr Diet.* Jun 2010;23(3):254-259.

2. Sandercock G, Voss C, McConnell D, Rayner P. Ten year secular declines in the cardiorespiratory fitness of affluent English children are largely independent of changes in body mass index. *Arch Dis Child.* Jan 2010;95(1):46-47.

3. Stratton G, Canoy D, Boddy LM, et al. Cardiorespiratory fitness and body mass index of 9-11-year-old English children: a serial cross-sectional study from 1998 to 2004. *Int J Obes (London).* Jul 2007;31(7):1172-1178.

4. Cohen DD, Voss C, Taylor MJD, et al. Ten-year secular changes in muscular fitness in English children. *Acta Paediatr.* 2011;100(10):e175-e177.

5. Tomkinson GR, Leger LA, Olds TS, Cazorla G. Secular trends in the performance of children and adolescents (1980-2000): an analysis of 55 studies of the 20m shuttle run test in 11 countries. *Sports Med.* 2003;33(4):285-300.

6. Tremblay MS, Barnes JD, Copeland JL, Esliger DW. Conquering Childhood Inactivity: Is the Answer in the Past? *Med Sci Sports Exerc.* 2005/07// 2005;37(7):1187-1194.

7. Tremblay MS, Shields M, Laviolette M, et al. Fitness of Canadian children and youth: results from the 2007-2009 Canadian Health Measures Survey. *Health Rep.* Mar 2010;21(1):7-20.

8. Ortega FB, Ruiz JR, Castillo MJ, Sjostrom M. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes (Lond).* Jan 2008;32(1):1-11.

9. Artero EG, Ruiz JR, Ortega FB, et al. Muscular and cardiorespiratory fitness are independently associated with metabolic risk in adolescents: the HELENA study. *Pediatr Diabetes.* Apr 6 2011;12(8):704-712.

10. Eisenmann JC. Aerobic fitness, fatness and the metabolic syndrome in children and adolescents. *Acta Paediatr.* Dec 2007;96(12):1723-1729.

11. Hurtig-Wennlof A, Ruiz JR, Harro M, Sjostrom M. Cardiorespiratory fitness relates more strongly than physical activity to cardiovascular disease risk factors in healthy children and adolescents: the European Youth Heart Study. *Eur J Cardiovasc Prev Rehabil.* Aug 2007;14(4):575-581.

12. Steene-Johannessen J, Anderssen SA, Kolle E, et al. Low muscle fitness is associated with metabolic risk in youth. *Med Sci Sports Exerc.* 2009 Jul;41(7):1361-7.

13. Parrett AL, Valentine RJ, Arngrímsson SA, et al.Adiposity and aerobic fitness are associated with metabolic disease risk in children. *Appl Physiol Nutr Metab.* 2011 Feb;36(1):72-9.

14. Jackson AW, Lee DC, Sui X, et al. Muscular strength is inversely related to prevalence and incidence of obesity in adult men. *Obesity (Silver Spring).* Oct 2010;18(10):1988-1995.

15. Silventoinen K, Magnusson PK, Tynelius P, et al. Association of body size and muscle strength with incidence of coronary heart disease and cerebrovascular diseases: a population-based cohort study of one million Swedish men. *Int J Epidemiol.* Feb 2009;38(1):110-118.

16. Ortega FB, Silventoinen K, Tynelius P, et al. Muscular strength in male adolescents and premature death: cohort study of one million participants.BMJ. 2012 Nov 20;345:e7279.

17. Donaldson L. The 2009 Annual Report of the Chief Medical Officer. 2010; http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_114012.pdf. Accessed 27 March 2012.

18. Leger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci.* Summer 1988;6(2):93-101.

19. Kelder SH, Springer AS, Barroso CS, et al. Implementation of Texas Senate Bill 19 to Increase Physical Activity in Elementary Schools. *J Public Health Pol.* 2009;30:S221-S247.

20. The New York City Department of Education. NYC FITNESSGRAM. 2012; http://schools.nyc.gov/Academics/FitnessandHealth/NycFitnessgram/NYCFITNESSGRAM.htm. Accessed 27 March 2012.

21. Moliner-Urdiales D, Ruiz JR, Ortega FB, et al. Secular trends in health-related physical fitness in Spanish adolescents: the AVENA and HELENA studies. *J Sci Med Sport.* Nov 2010;13(6):584-588.

22. Cale L, Harris J, Chen MH, et al. More than 10 years after "The horse is dead...": Surely it must be time to "dismount"?! *Pediatr Exerc Sci.* May 2007;19(2):115-131.

23. Fairclough SJ, Boddy LM, Hackett AF, Stratton G. Associations between children's socioeconomic status, weight status, and sex, with screen-based sedentary behaviours and sport participation. *Int J Pediatr Obes.* 2009;4(4):299-305.

24. Cohen DD, Voss C, Taylor MJD, Stasinopoulos DM, Delextrat A, Sandercock GRH. Handgrip strength in English schoolchildren. *Acta Paediatr.* Jul 2010;99(7):1065-1072.

25. Sandercock GRH, Voss C, Cohen D, et al. Centile curves and normative values for the twenty metre shuttle-run test in English schoolchildren. *J Sports Sci.* 2012;30(7):679-687.

26. Taylor MJD, Cohen D, Voss C, Sandercock GRH. Vertical jumping and leg power normative data for English school children aged 10-15 years. *J Sports Sci.* 2010;28(8):867-872.

27. Sandercock G, Voss C, Gladwell V. Twenty-metre shuttle run test performance of English children aged 11-15 years in 2007: comparisons with international standards. *J Sports Sci.* Jul 2008;26(9):953-957.

28. Sandercock GR, Voss C, Dye L. Associations between habitual school-day breakfast consumption, body mass index, physical activity and cardiorespiratory fitness in English schoolchildren. *Eur J Clin Nutr.* Oct 2010;64(10):1086-1092.

29. Sandercock GRH, Ogunleye A, Voss C. Comparison of cardiorespiratory fitness and body mass index between rural and urban youth: Findings from the East of England Healthy Hearts Study. *Pediatrics International.* Oct 2011;53(5):718-724.

30. Voss C, Sandercock G. Aerobic fitness and mode of travel to school in English schoolchildren. *Med Sci Sports Exerc.* Feb 2010;42(2):281-287.

31. Corbin CB. Texas Youth Fitness Study: a commentary. *Res Q Exerc Sport.* Sep 2010;81(3 Suppl):S75-78.

32. Cooper KH. Reflections on the Texas Youth Evaluation Project and implications for the future. *Res Q Exerc Sport.* 2010;81(3 Suppl):S79-83.

33. Naughton GA, Carlson JS, Greene DA. A challenge to fitness testing in primary schools. *J Sci Med Sport.* 2006;9:40-45.

34. Association for Physical Education. Supporting the Delivery of Physical Education & School Sport. 2012; http://www.afpe.org.uk/. Accessed 27 March 2012.

35. Domangue E, Solmon M. Motivational responses to fitness testing by award status and gender. *Res Q Exerc Sport.* 2010;81(3 Suppl):310-318.

36. Martin SB, Ede A, Morrow JR, Jackson AW. Statewide physical fitness testing: perspectives from the gym. *Res Q Exerc Sport.* 2010;81(3 Suppl):S31-41.

37. Zhu W, Welk GJ, Meredith MD, Boiarskaia EA. A survey of physical education programs and policies in Texas schools. *Res Q Exerc Sport.* 2010;81(3 Suppl):S42-S52.

38. Dishman RK, Hales DP, Pfeiffer KA, et al. Physical self-concept and self-esteem mediate cross-sectional relations of physical activity and sport participation with depression symptoms among adolescent girls. *Health Psychol.* May 2006;25(3):396-407.

39. Ridler C, Townsend N, Dinsdale H, et al. National Child Measurement Programme: Detailed Analysis of the 2007/08 Dataset. 2009; http://www.noo.org.uk/securefiles/091016_1351//noo_NCMPreport1_110509.pdf. Accessed 27 March 2012.

40. Lloyd M, Colley RC, Tremblay MS. Advancing the Debate on 'Fitness Testing' for Children: Perhaps We're Riding the Wrong Animal. *Pediatr Exerc Sci.* May 2010;22(2):176-182.

41. Rowland TW. The Horse is Dead; Let's Dismount. *Pediatr Exerc Sci.* 1995;7:117-120.

42. Cooper KH, Everett D, Meredith MD, Kloster J, et al. Texas Statewide Assessment of Youth Fitness. *Res Q Exerc Sport.* Sep 2010;81(3 Suppl):ii-iv.

43. Texas Education Agency. *Texas fitness now. 2008-2009 Year end report.* Austin, TX2009.

44. Bacon L, Aphramor L. Weight Science: Evaluating the Evidence for a Paradigm Shift. *Nutr J.* Jan 24 2011;10(9):1-13.

45. Blair SN, Kohl HW, 3rd, Paffenbarger RS, Jr., Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *Journal of the American Medical Association.* Nov 3 1989;262(17):2395-2401.

46. Sui X, Laditka JN, Hardin JW, Blair SN. Estimated functional capacity predicts mortality in older adults. *J Am Geriatr Soc.* Dec 2007;55(12):1940-1947.

47. Sui X, LaMonte MJ, Blair SN. Cardiorespiratory fitness as a predictor of nonfatal cardiovascular events in asymptomatic women and men. *Am J Epidemiol.* Jun 15 2007;165(12):1413-1423.

48. Wei M, Kampert JB, Barlow CE, et al. Relationship between low cardiorespiratory fitness and mortality in normal-weight, overweight, and obese men. *JAMA.* Oct 27 1999;282(16):1547-1553.

49. Artero EG, España-Romero V, Ortega FB, et al. Health-related fitness in adolescents: underweight, and not only overweight, as an influencing factor. The AVENA study. *Scand J Med Sci Sports.* 2010;20(3):418-427.

50. Morano M, Colella D, Robazza C, et al. Physical self-perception and motor performance in normal-weight, overweight and obese children. *Scand J Med Sci Sports.* Jun 2011;21(3):465-473.

51. Anderson CB, Masse LC, Zhang H, et al. Contribution of Athletic Identity to Child and Adolescent Physical Activity. *Am J Prev Med.* Sep 2009;37(3):220-226.

52. Sollerhed AC, Apitzsch E, Rastam L, Ejlertsson G. Factors associated with young children's self-perceived physical competence and self-reported physical activity. *Health Educ Res.* Feb 2008;23(1):125-136.

53. Olds T, Tomkinson G, Leger L, Cazorla G. Worldwide variation in the performance of children and adolescents: an analysis of 109 studies of the 20-m shuttle run test in 37 countries. *J Sports Sci.* Oct 2006;24(10):1025-1038.

54. Garcia-Artero E, Ortega FB, Ruiz JR, et al. [Lipid and metabolic profiles in adolescents are affected more by physical fitness than physical activity (AVENA study)]. *Rev Esp Cardiol.* Jun 2007;60(6):581-588.

55. Trudeau F, Shephard RJ. Physical education, school physical activity, school sports and academic performance. *Int J Behav Nutr Phys Act.* 2008;5(10):1-12.

56. Casajus JA, Leiva MT, Villarroya A, et al. Physical performance and school physical education in overweight Spanish children. *Ann Nutr Metab.* 2007;51(3):288-296.

57. Mahar MT, Guerieri AM, Hanna MS, David Kemble C. Estimation of Aerobic Fitness from the 20-m Multistage Shuttle Run Test Performance. *Am J Prev Med.* 2011;41(4S2):S117-S123.

58. Wellcome Trust. In the zone. 2012; http://www.wellcome.ac.uk/Education-resources/Teaching-and-education/In-the-zone/index.htm. Accessed 27 March 2012.

59. McInnis K. Active Science. Where fitness and fun meet science and achievement. 2010; http://www.activescienceforkids.org/. Accessed 27 March 2012.

60. Welk GJ, Corbin CB, Dale D. Measurement issues in the assessment of physical activity in children. *Res Q Exerc Sport.* 2000;71(2 Suppl):S59-73.

61. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of Accelerometer Cut Points for Predicting Activity Intensity in Youth. *Med Sci Sport Exer.* Jul 2011;43(7):1360-1368.

62. Espana-Romero V, Artero EG, Jimenez-Pavon D, et al. Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA Study. *Int J Sports Med.* Jul 2010;31(7):490-497.

63. Morrow JR, Jr., Martin SB, Jackson AW. Reliability and validity of the FITNESSGRAM: quality of teacher-collected health-related fitness surveillance data. *Res Q Exerc Sport.* Sep 2010;81(3 Suppl):S24-30.

64. Blair SN, Cheng Y, Holder JS. Is physical activity or physical fitness more important in defining health benefits? *Med Sci Sports Exerc.* Jun 2001;33(6 Suppl):S379-399; discussion S419-320.

65. Williams PT. Physical fitness and activity as separate heart disease risk factors: a meta-analysis. *Med Sci Sports Exerc.* May 2001;33(5):754-761.

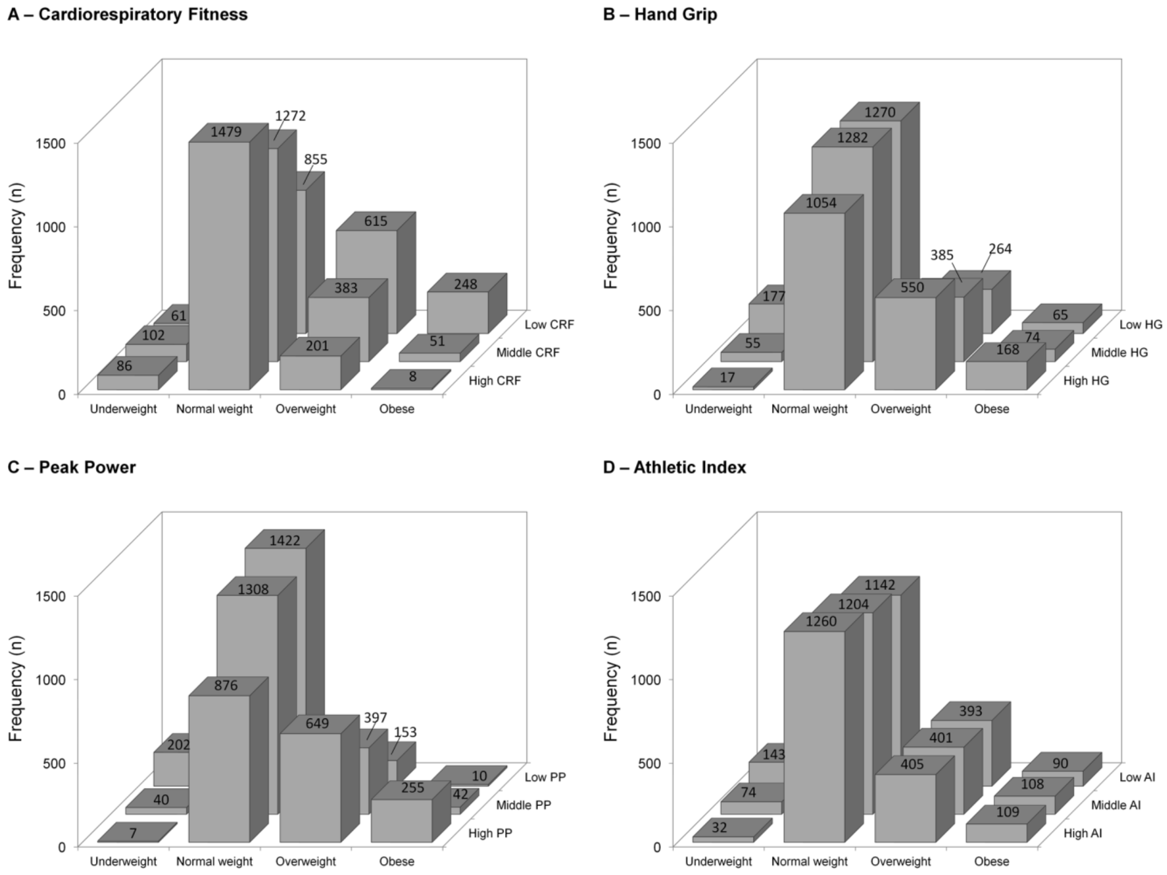
66. McCarthy HD, Ellis SM, Cole TJ. Central overweight and obesity in British youth aged 11-16 years: cross sectional surveys of waist circumference. *BMJ.* Mar 22 2003;326(7390):624.

67. Ness AR, Leary SD, Mattocks C, et al. Objectively measured physical activity and fat mass in a large cohort of children. *PLoS Med.* Mar 2007;4(3):e97.

68. Naylor PJ, McKay HA. Prevention in the first place: schools a setting for action on physical inactivity. *Br J Sports Med.* Jan 2009;43(1):10-13.

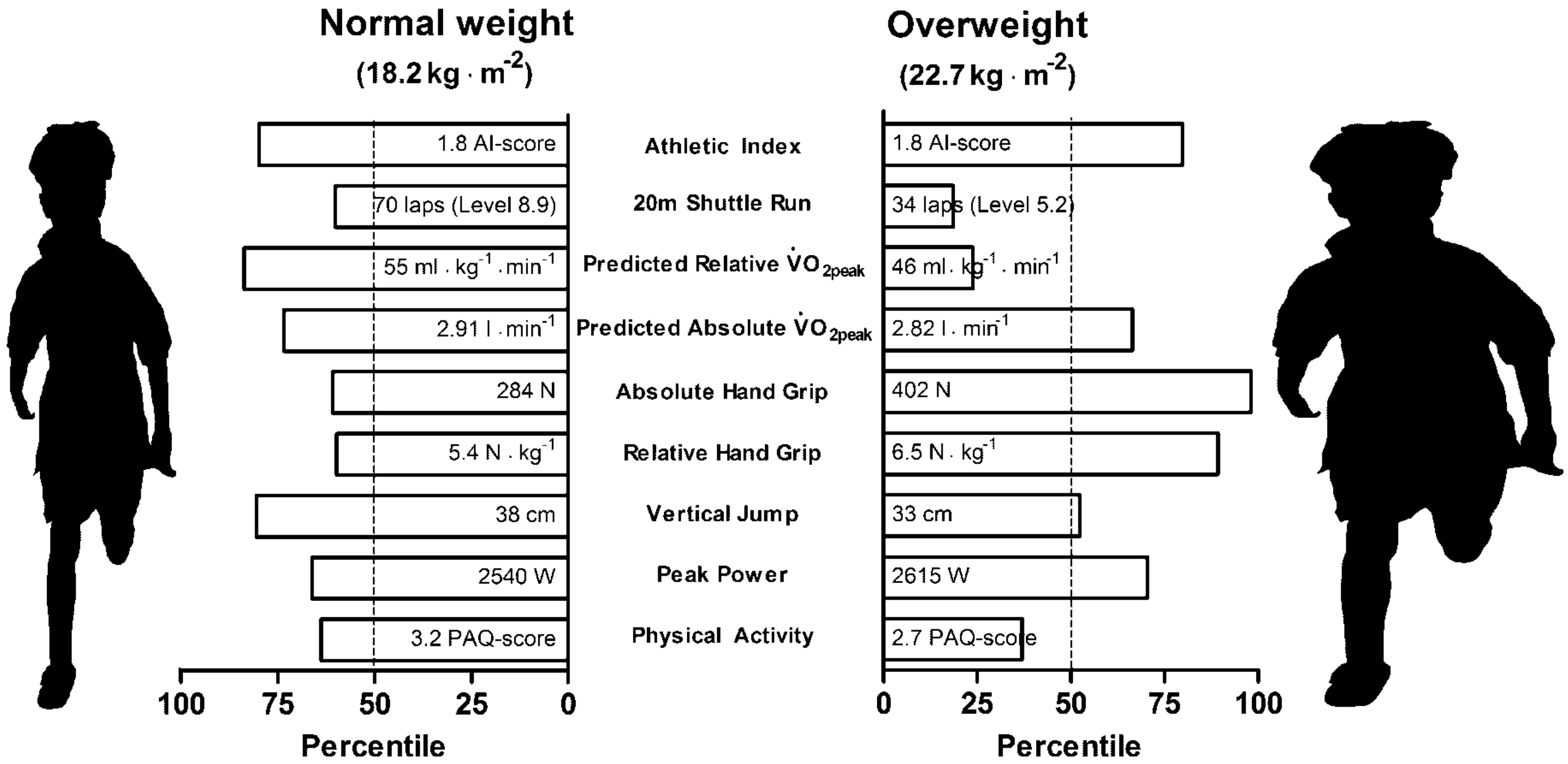
69. Kowalski KC, Crocker RE, Donen RM. (2004). *The Physical Activity Questionnaire for Older Children (PAC-C) and Adolescents (PAQ-A) Manual*. Saskatoon, Canada: Universtiy of Saskatchewan.

Figure 1. Prevalence of performance tertiles by IOTF weight category in 10-15 year-olds from the East of England Healthy Hearts Study (n = 5361, 53% boys)



CRF – cardiorespiratory fitness (Panel A), tertiles created based on final running speed z-scores at last completed level;[53] HG – hand grip (Panel B), tertiles created based on dominant arm hand grip z-scores;[24] PP – peak power (Panel C), tertiles created based on peak power z-scores;[26] AI – athletic index (Panel D), tertiles created based on composite z-scores of CRF, HG, vertical jump (not shown) and PP.

Figure 2. Performance portrait of a normal weight (left) and overweight (right) 13-year-old boy with identical athletic index scores from the East of England Healthy Hearts Study



Individual performance scores (text), and corresponding age-sex specific percentile scores (bars) based on East of England Healthy Hearts cross-sectional cohort. AI – athletic index; composite z-scores of cardiorespiratory fitness, hand grip, vertical jump and peak power. $\dot{V}O_{2peak}$ – peak oxygen uptake; relative $\dot{V}O_{2peak}$ predicted;[57] total $\dot{V}O_{2peak}$ derived from relative $\dot{V}O_{2peak}$ and body mass; Hand Grip – dominant arm maximum grip strength; Vertical Power – calculated based on Taylor, et al., (2010);[26] PAQ-score – Physical Activity Questionnaire for Adolescents,[69] scored between 1 (inactive) and 5 (very active).

Appendix 1: East of England Healthy Heart Study Methods & Results

Methods

Sample

After gaining institutional ethics approval, parental consent and participant assent, schoolchildren were tested during regularly scheduled physical education (PE) classes during the summer months of 2006 to 2010 as part of the East of England Healthy Hearts Study. To date, over 8600 schoolchildren from 26 state-run primary and secondary schools have participated in all or most aspects of this ongoing health and fitness survey, with an overall response rate of 98%. The present analyses were restricted to individuals who completed all three performance tests (20m shuttle run, hand grip strength, vertical jump) and who had valid information on sex, age (10.0-15.9 years) and body mass (n = 5366, 53% boys). Reasons for missing data points ranged from illness or injury as well as child refusal or partial parental withdrawal which prohibited individuals from taking part in all aspects of the protocol.

Protocol

Participants had their body mass measured to the nearest 0.1 kg and their stature to the nearest 1 mm whilst wearing regular PE clothing. We have previously described in detail the school-based measurement protocols for the three performance tests.[1-3] In brief, cardio-respiratory fitness (CRF) was estimated using the 20m shuttle run test[4] in the form of the FITNESSGRAM PACER™ .[5] Handgrip (HG) strength was measured with a Takei T.K.K.5001 GRIP A dynamometer (Takei Scientific Instruments Co. Ltd, Tokyo, Japan) in the standing position during arm movement from 180° of flexion to near 0°; two trials in the dominant limb were carried out and the highest score was recorded.[1] Vertical jump height (VJ) and peak power (PP) were calculated from the best of two countermovement jumps on a timing mat (NewTest, Timing Mat, NewTest Ltd., UK).[2] Participants also completed a 7-day recall physical activity (PA) questionnaire,[6] which is an average score of responses to 8 questionnaire items, with 1 indicating low PA and 5 indicating very high PA.

Data treatment

Body mass index (BMI) was calculated (kg·m^{-2}) and categorised into underweight, normal weight, overweight or obese according to International Obesity Task Force (IOTF) criteria .[7,8] Relative $\dot{V}O_{2peak}$ (ml·min^{-1}·kg^{-1}) was predicted based on Mahar *et al.*'s quadratic equation,[9] which utilises the final shuttle count as well as sex and BMI. Total $\dot{V}O_{2peak}$ (l·min^{-1}) was derived from predicted relative $\dot{V}O_{2peak}$ (ml·kg^{-1}·min^{-1}) and body mass (kg). Shuttle count was also rounded down to the last completed level, expressed as final running speed (km·h^{-1}) and converted to CRF z-scores based on global norms. [10] HG, VJ and PP z-scores were created based on existing reference data (Cohen, et al., 2010; Taylor, et al., 2010). An athletic index (AI) score was calculated by creating a composite score of CRF, HG, VJ and PP z-scores. All performance z-scores were individually ranked and collapsed into tertiles, with the bottom tertile corresponding to the bottom third of performance, and the top tertile corresponding to the top third of performance. Age-and-sex-specific percentiles were also created for relative $\dot{V}O_{2peak}$ (ml·kg^{-1}·min^{-1}), total $\dot{V}O_{2peak}$ (l·min^{-1}), relative HG (N·kg^{-1}) and PAQ-score.

Statistical analyses

Tertiles for each performance indicator were cross-tabulated by IOTF weight categories and Pearson's χ^2 analyses were run to identify if prevalence of performance tertiles were different by IOTF weight categories ($p < 0.05$; *post hoc* standardised residual of ≤ -1.96 or ≥ 1.96). SPSS 18.0 for windows (SPSS Inc., an IBM Company, Chicago, IL) was used for all analyses.

Results

In agreement with other recent British data,[11] 5% of the present sample were underweight, 67% were normal weight, 22% were overweight, and 6% were obese by IOTF criteria (Cole, et al., 2000). Prevalence of CRF tertiles (Figure 1, Panel A; $\chi^2 = 716.8$, $p < 0.001$) differed significantly between IOTF categories. Low CRF was more common in overweight and obese participants compared with normal weight children, whereas the reverse was true for high CRF. HG tertiles (Figure 1, Panel B; χ^2 = 363.2, $p < 0.001$) and PP tertiles (Figure 1, Panel C; χ^2 = 1054.9, $p < 0.001$) were also significantly different between IOTF categories. Poor performance (low HG and PP) was more common in underweight and normal weight participants when compared with overweight and obese youth, whereas good performance (high HG and PP) was more common in overweight and obese participants. Although prevalence of AI tertiles differed significantly between IOTF categories (Figure 1, Panel D; χ^2 = 83.9, $p < 0.001$), this was almost exclusively due to the few good and abundant poor AIs within the underweight population. There were no noteworthy differences in AI tertiles between normal weight, overweight and obese participants.

References

1. Cohen DD, Voss C, Taylor M J, et al. Handgrip strength in English schoolchildren. *Acta Paediatr.* 2010;99(7), 1065-1072.

2. Sandercock G, Voss C, Cohen D, et al. Centile curves and normative values for the twenty metre shuttle-run test in English schoolchildren. *J Sports Sci*. 2012;30(7), 679-87.

3. Taylor M J, Cohen D, Voss C, Sandercock G R. Vertical jumping and leg power normative data for English school children aged 10-15 years. *J Sports Sci*. 2010; 28(8), 867-872.

4. Leger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci.* 1988; 6(2), 93-101.

5. Meredith MD, Welk GJ. Chapter 5. Aerobic Capacity. In M. D. Meredith & G. J. Welk (Eds.), FITNESSGRAM / ACTIVITYGRAM ADMINISTRATION MANUAL (4th ed., pp. 27-34). 2007. Champaign, IL: Human Kinetics.

6. Kowalski KC, Crocker RE, Donen RM. The Physical Activity Questionnaire for Older Children (PAC-C) and Adolescents (PAQ-A) Manual. Saskatoon, Canada: University of Saskatchewan. 2004

7. Cole TJ, Bellizzi MC, Flegal, KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ.* 2000;320(7244), 1240-1243.

8. Cole TJ, Flegal KM, Nicholls D, Jackson AA. Body mass index cut offs to define thinness in children and adolescents: international survey. *BMJ*. 2007;335(7612), 194-201.

9. Mahar MT, Guerieri AM, Hanna MS, Kemble DC. Estimation of Aerobic Fitness from the 20-m Multistage Shuttle Run Test Performance. *Am J Prev Med*. 2011; 41(4S2), S117-S123.

10. Olds T, Tomkinson G, Leger L, Cazorla G. Worldwide variation in the performance of children and adolescents: an analysis of 109 studies of the 20-m shuttle run test in 37 countries. *J Sports Sci.* 2006; 24(10), 1025-1038.

11. Boddy LM, Hackett AF, Stratton G. Changes in BMI and prevalence of obesity and overweight in children in Liverpool, 1998-2006. *Perspect Public Health.* 2009; 129(3), 127-131.