

**EXPLORING THE UTILITY OF CARDIORESPIRATORY FITNESS
AS A POPULATION HEALTH SURVEILLANCE INDICATOR FOR
CHILDREN AND YOUTH: AN INTERNATIONAL ANALYSIS OF
RESULTS FROM THE 20 M SHUTTLE RUN TEST**

JUSTIN LANG

Dissertation submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the PhD degree in Population Health

PhD program in Population Health
University of Ottawa
Ottawa, Ontario, Canada

© Justin Lang, Ottawa, Canada, 2017

TABLE OF CONTENTS

Abstract	vi
Acknowledgments	viii
List of Figures	x
List of Tables	xiii
List of Abbreviations	xv
Chapter 1: Introduction	1
1.1 Dissertation Overview and Rationale	2
1.2 Conceptual Framework	3
1.2.1 Health-related fitness model	3
1.2.2 Ecological model	5
1.3 Research Objectives.....	6
1.4 Dissertation Organization and Outline.....	7
1.5 References.....	9
Chapter 2: Review of Literature	11
2.1 Measurement Instruments	11
2.1.1 Physical activity	11
2.1.1.1 Validity of accelerometers	12
2.1.2 Cardiorespiratory fitness	13
2.1.2.1 Validity of the 20m shuttle run test.....	16
2.2 Intra-Individual Variability	17
2.2.1 Physical activity versus cardiorespiratory fitness	17
2.3 Association with Health Markers.....	18
2.3.1 Adulthood	19
2.3.2 Childhood and adolescence.....	19
2.3.2.1 Physical activity and health.....	20
2.3.2.2 Cardiorespiratory fitness and health	20
2.3.3 Tracking from childhood to adulthood	21
2.4 Benefits and Drawbacks	22

2.5 Conclusion	24
2.6 References.....	25
 Chapter 3: Manuscript 1 - Systematic review of the relationship between 20 m shuttle run performance and health indicators among children and youth	
Introduction.....	32
Methods	33
Results.....	34
Discussion.....	35
Conclusion	43
 Chapter 4: Manuscript 2 - Review of criterion-referenced standards for cardiorespiratory fitness: What percentage of 1 142 026 international children and youth are apparently healthy?	
Introduction.....	48
Methods	49
Results.....	50
Discussion.....	50
Conclusion	52
 Chapter 5: Manuscript 3 - Cardiorespiratory fitness is associated with physical literacy among Canadian children aged 8 to 12 years	
Introduction.....	56
Methods	60
Results.....	61
Discussion.....	68
Conclusion	69
 Chapter 6: Manuscript 4 - International variability in 20 m shuttle run performance in children and youth: Who are the fittest from a 50-country comparison? A systematic review with pooling of aggregate results	
Introduction.....	93
Methods	94

Results.....	96
Discussion.....	97
Conclusion	100
 Chapter 7: Manuscript 5 - Making a case for cardiorespiratory fitness surveillance	
among children and youth.....	106
Introduction.....	109
Cardiorespiratory Fitness and Physical Activity.....	110
Cardiorespiratory fitness measurement.....	111
Health Indicators among Children and Youth	114
Cardiorespiratory fitness and health	115
Cardiorespiratory fitness and future health.....	116
Cardiorespiratory fitness standards.....	116
Population-based Surveillance Utility	117
Comparisons of surveillance measures across countries	118
Temporal trends in surveillance measures	120
Novel Population Health Surveillance Strategies	121
The scientific research setting.....	121
The physical education setting.....	122
The clinical setting.....	123
Importance of Guidelines for Public Health	124
Cardiorespiratory fitness guidelines.....	124
How to meet the cardiorespiratory fitness guidelines	125
Perspectives for Progress	126
References.....	128
 Chapter 8: General Discussion	140
8.1 Cardiorespiratory Fitness and Health among Children and Youth	140
8.2 Cardiorespiratory Fitness as a Population Health Surveillance Indicator.....	142
8.3 Summary of Strengths and Limitations	144
8.3.1 Strengths	144
8.3.2 Limitations	145
8.4 Summary of Important Future Directions.....	147

8.5 Suitability for a Doctoral Dissertation in Population Health	148
8.6 Conclusion	149
8.7 References.....	150
 Supplemental Manuscript 1 - International normative 20 m shuttle run values from	
1 142 026 children and youth representing 50 countries	151
Introduction.....	152
Methods	153
Results.....	155
Discussion	157
Conclusion	160
 Supplemental Manuscript 2 - Temporal trends in the cardiorespiratory fitness of	
965 264 children and youth representing 19 countries since 1981.....	169
Introduction.....	172
Methods	174
Results.....	177
Discussion	179
Conclusion	184
 Appendix A : Study Registration.....	
Appendix B : Chapter 3 Supplemental Materials	203
Appendix C : Ethics.....	310
Appendix D : In the News	312
Appendix E : Accepted ESSR Proposal Form	313
Appendix F : Summary of Included Studies (Supplemental Manuscript 1)	317
Appendix G : Summary of Conference Abstracts and Presentations	327

ABSTRACT

Emerging evidence has demonstrated the strong link between cardiorespiratory fitness and multiple aspects of health (i.e., physiological, physical, psychosocial, cognitive), independent of physical activity, among school-aged children and youth. Cardiorespiratory fitness is a trait that does not vary substantially from day-to-day, and provides an indication of recent physical activity levels, making it an important possible indicator of population health. Thus, the objective of this dissertation was to investigate the utility of cardiorespiratory fitness, measured using the 20 m shuttle run test, as a broad, holistic health indicator for population health surveillance among children and youth. To achieve this objective we completed seven manuscripts, all prepared for submission to peer-reviewed, scientific journals:

1. Systematic review of the relationship between 20 m shuttle run performance and health indicators among children and youth.
2. Review of criterion-referenced standards for cardiorespiratory fitness: What percentage of 1 142 026 international children and youth are apparently healthy?
3. Cardiorespiratory fitness is associated with physical literacy among Canadian children aged 8 to 12 years.
4. International variability in 20 m shuttle run performance in children and youth: Who are the fittest from a 50-country comparison? A systematic review with pooling of aggregate results.
5. Making a case for cardiorespiratory fitness surveillance among children and youth.
6. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries.

7. Temporal trends in the cardiorespiratory fitness of 965 264 children and youth representing 19 countries since 1981.

Combined, this dissertation provides support for the importance of cardiorespiratory fitness for health surveillance among school-aged children and youth. Results from the international analysis highlighted the variability across countries, with countries in North-Central Europe and Africa having the highest cardiorespiratory fitness, and countries in South America having the lowest cardiorespiratory fitness. The results indicated that declines in cardiorespiratory fitness may have stabilized in recent years in some high- and middle-income countries. This dissertation also identified two methods (criterion- and normative-referenced standards) to interpret cardiorespiratory fitness levels among children and youth, methods that could be used to inform future consensus, surveillance, and cardiorespiratory fitness guidelines.

ACKNOWLEDGEMENTS

“Make a conscious effort to surround yourself with positive, nourishing, and uplifting people — people who believe in you, encourage you to go after your dreams, and applaud your victories.”

- Jack Canfield

I have been extremely fortunate to be surrounded by people who have supported me throughout my academic journey – all of whom have played a significant role in making my dissertation work possible. I would like to express my sincere gratitude.

To Dr. Mark Tremblay: I am very lucky to have had you as a supervisor and mentor over the past four years. You have contributed significantly to my development as a researcher, and I have learned many skills from you that I will take with me in my future endeavors. In retrospect, my experience as a PhD student has been smooth and without barriers, something that most graduate students don’t experience. My drama free PhD was a result of your abilities as an effective supervisor and your strong interpersonal skills. Thank you for everything!

To Dr. Grant Tomkinson: Over the years you have been a supervisor and mentor to me, although this was never a requirement on your part. You have been selfless and generous in providing me with full access to your data, and you took the time to teach me your tremendous data management and analytical skills. I have been lucky to have you as a mentor, and appreciate you taking the time to help me progress to become a better researcher. This dissertation would not have been possible without you.

To my committee: My committee consisted of expert, world-renowned scientists who have guided me and provided me with the clear and concise direction over the course of my PhD

studies. Thank you for taking the time out of your day to read through my work. Thank you for encouraging me to become better at what I do.

To the HALO research group: You are some of the brightest, most energetic and caring people I know, and I have been lucky to have spent the last four years learning from you. Your work ethic is infectious. I am constantly inspired by the quality and quantity of work that you produce while leading such active lifestyles. I aspire to continue to lead the HALO lifestyle. Thank you for all the noon hour ball hockey/tennis games, and the 3pm Tim Horton coffee walks.

To my family and friends: I am very lucky to have a support system of family and freinds who are constantly there to help me, without question. You have been there during the highs and lows, you continued to believe in me, and you have always been there to celebrate my victories. I thank you for everything you have done along the way. I couldn't have done it without you.

To my future wife, Whitney: All these years you have known me only as a 'student.' As such, you have seen me at my worst – stressed out over various assignments, exams, papers, and presentations. You have constantly supported me and have always encouraged me to realize my dreams and aspirations. Your intelligence, perseverance, kind heart, and loving nature, constantly motivates me to become a better person. I'm excited to turn the page and move onto the next part of our lives together as 'parents and working professionals.'

LIST OF FIGURES

Chapter 1

Figure 1	Modified version of Bouchard and Shephard's health-related fitness model	4
Figure 2	Bauman and colleagues' ecological model	6

Chapter 3

Figure 1	PRISMA flowchart for the identification, screening, eligibility and inclusion of articles.	41
----------	---	----

Chapter 4

Figure 1	Study of criterion-referenced standards minus international normative 20-m shuttle run test (20mSRT) performance as a function of age for girls (A) and boys (B). The dash line, centred at 0, represents the international 20mSRT 50 th percentile (predicted VO _{2peak}) for 9- to 17-year-old children and youth. Lines below the dash line indicate that the criterion-referenced standards are below the international 50 th percentile of 20mSRT performance, and lines above represent standards above the 50 th percentile for each age group	52
Figure 2	Overall smoothed population-weighted mean percentages and 95% CIs for healthy CRF in girls and boys. Calculated across age-specific percentage estimates for each criterion-referenced standard (VO _{2peak}) available in table 3. CRF, cardiorespiratory fitness.	52

Chapter 5

Figure 1	Percentages of Canadian boys and girls with healthy cardiorespiratory fitness using the interim international standards. Data are presented as % and standard deviations.	91
----------	--	----

Chapter 6

Figure 1	Flow chart describing the general methodological procedure for this study. All protocol results (Léger, Eurofit or Queen's University of Belfast) were corrected to reflect the Léger protocol. All results were then corrected to reflect speed (km/h) at the last completed stage as a common metric. Next, sample-weighted mean 20mSRT performance score and SD were calculated for each sex by age by country group. Then z-scores were generated relative to the international norms for each	
----------	--	--

	sex by age by country group to calculate the mean country specific z-score, described as the ‘performance index.’ Each country specific performance index was corrected using a post-stratified population-weighted technique. Finally, a correlation analysis took place to analyse associations between the performance index and broad socioeconomic indicators. 20mSRT, 20 m shuttle run test.	95
Figure 2	Bar graph describing the distance (m) run by boys and girls on the 20mSRT based on country-specific and age-specific performance indices. Only countries (n=45) with 20mSRT performance indices for boys (represented by the black bars) and girls (represented by the grey bars) are shown. 20mSRT, 20 m shuttle run test.....	98
Chapter 7		
Figure 1	Modified version of Bouchard and Shephard’s health-related fitness, physical activity and health model.	1388
Supplemental Manuscript 1		
Figure 1	PRISMA flow chart outlining the identification of the included studies	153
Figure 2	Flow chart showing the methodological procedure used in the study. Results from the studies were first expressed in the common metric of running speed (km/h) at the final completed stage and then corrected to Léger’s original 1 min protocol according to which protocol they used (Léger, Eurofit or Queen’s University of Belfast). Following the estimation of missing means (from reported medians) and SDs (from calculated CVs), post-stratified population-weighted means and SDs were estimated for each sex by age group, with pseudo data then generated using Monte Carlo simulation. Smoothed centiles were then generated using the LMS method, with international normative 20mSRT values expressed in several different metrics (speed at the last completed stage, the number of completed full stages/minutes and relative VO_{2peak} (mL/kg/min)). 20mSRT, 20 m shuttle run test; CV, coefficient of variation; LMS, Lambda Mu and Sigma; and VO_{2peak} , peak oxygen uptake	154
Figure 3	Performance metric conversion flow chart adapted from Tomkinson et al. VO_{2peak} , peak oxygen uptake.....	155
Figure 4	World map representing the 50 countries (shown in black) for which 20mSRT data on 9–17 year old were available. 20mSRT, 20 m shuttle run test.....	156

Figure 5	Smoothed centile curves (P_{10} , P_{50} and P_{90}) for the 20mSRT (speed (km/h) at the last completed stage) performance of (A) boys and (B) girls. 20mSRT, 20 m shuttle run test.....	159
Figure 6	Prevalence of 10–17 year olds from 50 countries meeting the FITNESSGRAM Healthy Fitness Zone (V.10) thresholds for VO_{2peak} (mL/kg/min). VO_{2peak} , peak oxygen uptake	159
Figure 7	Standardised differences in mean 20mSRT performance (speed (km/h) at the least completed stage) between (A) age-matched boys and girls and (B) sex-matched children of different ages (anchored to age 17 years=0). Positive differences indicate that 20mSRT performances were higher for boys than for girls (top panel) or for older children than for younger children (bottom panel). The limits of the grey zone represent the thresholds for a large standardised difference (0.8 or –0.8). 20mSRT, 20 m shuttle run test.....	159
Supplemental Manuscript 2		
Figure 1	Flow chart outlining the identification of the included studies.....	192
Figure 2	International temporal trends in mean CRF (mL/kg/min) between 1981 and 2014. Note: Figure 2a shows the absolute differences in CRF, with the grey child-like figure indicating mean CRF in 1981 and the white child-like figure indicating mean CRF in 2014. Figure 2b shows the time-related patterns of change in CRF. Data are standardised to the year 2000=0 mL/kg/min, with positive values indicating better CRF and negative values poorer CRF; the solid lines represent the international change in mean CRF and the shaded areas the 95% CIs, with upward sloping lines indicating increases over time and downward sloping lines declines; mean (95%CI) absolute changes (per decade) are shown at the top of each panel	192
Figure 3	National temporal trends in mean CRF (mL/kg/min) between 1981 and 2014	192

LIST OF TABLES

Chapter 2

Table 1	Running speed (km/h) and number of laps for each stage of the 20mSRT according to the three major protocol variants. Published with permission from Tomkinson <i>et al.</i> ³³	15
---------	---	----

Chapter 3

Table 1	Association between 20mSRT performance and adiposity.....	36
Table 2	Association between 20mSRT performance and cardiometabolic biomarkers.....	36
Table 3	Association between 20mSRT performance and cognition/academic achievement.....	37
Table 4	Association between 20mSRT performance and mental health.....	38
Table 5	Association between 20mSRT performance and psychosocial health.....	39
Table 6	Association between 20mSRT performance and self-esteem and physical self-perception.....	40
Table 7	Association between 20mSRT performance and quality of life and wellbeing	41
Table 8	Association between 20mSRT performance and bone health.....	43
Table 9	Association between 20mSRT performance and musculoskeletal fitness	44
Table 10	Association between 20mSRT performance and motor skill development	44
Table 11	Association between 20mSRT performance and risk of injury/harm	45
Table 12	High-level summary of findings by health indicator	46

Chapter 4

Table 1	International 20mSRT norms for VO_{2peak} (mL/kg/min) obtained by converting Tomkinson et al's normative centile values using the Léger equation	50
Table 2	Descriptive characteristics of included studies	51
Table 3	Percentage of the 1 141 026 international 20mSRT performance pseudo-datasets that met each CRF criterion-referenced standard.....	52

Chapter 5

Table 1	Descriptive characteristics of participants.	79
Table 2	Tertiles of cardiorespiratory fitness across components of physical literacy for 8 year-old boys and girls.	81
Table 3	Tertiles of cardiorespiratory fitness across components of physical literacy for 9 year-old boys and girls.	83

Table 4	Tertiles of cardiorespiratory fitness across components of physical literacy for 10 year-old boys and girls.	85
Table 5	Tertiles of cardiorespiratory fitness across components of physical literacy for 11 year-old boys and girls.	87
Table 6	Tertiles of cardiorespiratory fitness across components of physical literacy for 12 year-old boys and girls.	89

Chapter 6

Table 1	International 20mSRT norms for running speed (km/h) at the last completed stage.	96
Table 2	Potential correlates of 20mSRT performance in children and youth.	96
Table 3	Descriptive data for countries presented in rank order by 2016 performance index	97
Table 4	Spearman's rank correlation coefficients (95%CI) between 20mSRT performance index and socioeconomic indicators.	98

Chapter 7

Table 1	20mSRT prediction equations to estimate $\dot{V}O_{2peak}$ in children and youth.	1366
Table 2	20mSRT minimum recommended standards of 42 and 35 ml·kg ⁻¹ ·min ⁻¹ for boys and girls, respectively.	1377

Supplemental Manuscript 1

Table 1	Twenty-metre shuttle run (speed (km/h) at the last completed stage) centile by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981	1566
Table 2	Twenty-metre shuttle run (number of completed stages/minutes) centile by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981	1577
Table 3	Twenty-metre shuttle run (number of laps) centile by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981	1577
Table 4	Twenty-metre shuttle run ($\dot{V}O_{2peak}$ mL/kg/min) centile by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981	1588

Supplemental Manuscript 2

Table 1	Summary of the included studies by country.	1900
Table 2	National and international temporal trends in mean CRF (per decade) from 965,264 children aged 9–17 years from 19 countries between 1981 and 2014.	1911

LIST OF ABBREVIATIONS

20mSRT	20 metre shuttle run test
%BF	Percent bodyfat
AUC	Area under the curve
BMI	Body mass index
CAMSA	Canadian Agility and Movement Skill Assessment
CES-DC	Center for Epidemiological Studies – Depression scale for Childhood
CI	Confidence interval
CRE	Cardiorespiratory endurance
CRF	Cardiorespiratory fitness
CRP	C-reactive protein
CSAPPA	Children’s Self-Perception and Adequacy in and Predilection for Physical Activity
FMSMCT	Fundamental Motor Skill Manual for Classroom Teachers
GRADE	Grading of Recommendations Assessment, Development and Evaluation
GSHS	Global School-based student Health Survey
HBSC	Health Behaviour in School-aged Children
HDI	Human Development Index
HDL	High-density lipoprotein
HELENA	Healthy Lifestyle in European by Nutrition in Adolescence
HFZ	FITNESSGRAM Healthy Fitness Zone
HGS	Hand-grip strength
HOMA-IR	Homeostatic model assessment of insulin resistance

HRQoL	Health-related quality of life
ICAD	International Children's Accelerometry Database
IFIS	International Fitness Scale
LDL	Low-density lipoprotein
MAP	Mean arterial pressure
NCD	Non-communicable disease
NIZ	FITNESSGRAM Needs Improvement Zone
PA	Physical activity
PICOS	Participant, Intervention, Comparisons, Outcome, and Study design
PL	Physical literacy
QUB	Queen's University of Belfast
RCT	Randomized controlled trial
SAR	Sit-and-reach
SBJ	Standing broad jump
SE	Standard error
SSF	Sum of skinfolds
TC	Total cholesterol
TG	Triglycerides
TGMD-2	Test of Gross Motor Development – second edition
UK	United Kingdom
USA	United States of America
$\dot{V}O_{2peak}$	Peak oxygen uptake
WC	Waist circumference

CHAPTER 1: INTRODUCTION

Worldwide, the average life expectancy at birth has increased by roughly 10 years to 71.8 years between 1980 and 2015,¹ but this increase has not been matched by an equivalent increase in healthy life expectancy (i.e., life expectancy void of non-communicable disease (NCD) and/or disability).² This trend has caused a situation where people now live longer with NCDs and require costly treatments resulting in worldwide financial and economic strain. If this current trend persists there is a projected cumulative cost of \$47 trillion in lost economic production and healthcare cost worldwide between 2011 and 2030.³ Thus, there is a recognized need to implement cost-effective, population-based strategies to help prevent NCDs through effective policies and interventions that promote healthy active living.

Childhood represents an important time to begin NCD prevention because vital risk factors that predict future NCDs could manifest during these years.⁴ Furthermore, recent research has identified unfavourable trends in many health-related behaviours and traits (e.g., physical activity (PA),^{5,6} cardiorespiratory fitness (CRF) levels,⁷ strength,⁸ sedentary behaviours, and screen time⁹) among children and youth supporting the need for health promoting policies and interventions in these age groups. However, to implement effective NCD prevention strategies there is a need for a sustainable population health surveillance system to help quantify and monitor NCD risk and the associated determinants.¹⁰ Thus, the overall aim of this dissertation was to investigate the utility of CRF as a broad, holistic health indicator for population health surveillance among children and youth.

1.1 Dissertation Overview and Rationale

This dissertation evaluated CRF as a population health indicator for NCD surveillance among children and youth. Throughout this dissertation population health surveillance is defined using the Centers for Disease Control and Prevention definition: “systematic, ongoing collection, management, analysis, and interpretation of data followed by the dissemination of these data to public health programs to stimulate [population] health action” (p. 3).¹¹ The measurement of CRF throughout this dissertation is in the context of the 20 m shuttle run test (20mSRT) – a progressive, multistage endurance test where children are asked to run back-and-forth between two parallel lines, 20 meters apart, following the pace of an audio recording.^{12,13} Chapter 2 provides a detailed description of this measure and related concepts.

The underlying rationale and motivation for this dissertation lies in a profound interest in healthy public policy as a tool to take action on improving health. Population health surveillance provides two unique opportunities with respect to healthy public policy: 1) it provides a basis to *guide* policy efforts by identifying populations at greater risk; and 2) it provides a basis to *monitor* and *evaluate* the progress and impact of implemented policies.¹⁴ Current population health surveillance initiatives among children and youth have focused on behaviours that are associated with health outcomes. For example, PA levels among children and youth are a common indicator of population health,^{15,15} and by extension, as a means to guide healthy public policy and monitor its impact in improving health (i.e., PA levels). Therefore, this dissertation will begin by comparing PA and CRF measurement, and it will continue by further exploring the utility

of CRF as a potential population health surveillance indicator among school-aged children and youth.

1.2 Conceptual Framework

This dissertation was guided by two conceptual frameworks: a modified version of the Bouchard and Shephard's health-related fitness model (primary)¹⁷ and Bauman and colleagues' ecological model of the determinants of PA (secondary).¹⁸ Together, these models helped describe how the surveillance of health indicators may inform and guide population-based healthy public policies and intervention. The primary conceptual model by Bouchard and Shephard describes the inter-relationship between potential health indicators (i.e., PA, health-related fitness, and health), heredity, and other factors which includes the physical and social environment. The secondary conceptual model by Bauman and colleagues' provides an in-depth description of the physical and social environmental, and how population health indicators might lead to the implementation of regional or national policies and interventions.

1.2.1 Health-related fitness model

Bouchard and Shephard conceptualized the relationships between health-related fitness, PA, health, heredity, and other factors in their health-related fitness model, which was modified for this dissertation (Figure 1.1).¹⁷ This modified model was the outcome of a 4-day consensus conference that included an international group of scientists with expertise in health-related fitness, physical activity, and health.¹⁹ Bouchard and Shephard's modified model describes health-related fitness as a combination of physical

and physiological components that relate to the risk of developing diseases and/or comorbidities. Health-related fitness was divided into five assessment components: (a) morphological (body mass index, abdominal visceral fat, flexibility); (b) muscular (endurance, power, strength); (c) cardiorespiratory (i.e., CRF; oxygen transport, function of heart and lungs, blood pressure); (d) motor (agility, balance, coordination, speed); and (e) metabolic (glucose tolerance, lipid and lipoprotein metabolism, and choice of metabolic substrate).¹⁷ The modified model postulates that PA has a positive relationship with health-related fitness, although this relationship is weak to moderate in children and adolescents.²⁰ Additionally, both PA and health-related fitness have an independent influence on health (well-being, morbidity and mortality).²¹ These relationships are mediated and/or moderated by heredity and other factors, which include lifestyle behaviours, personal attributes, physical and social environments.

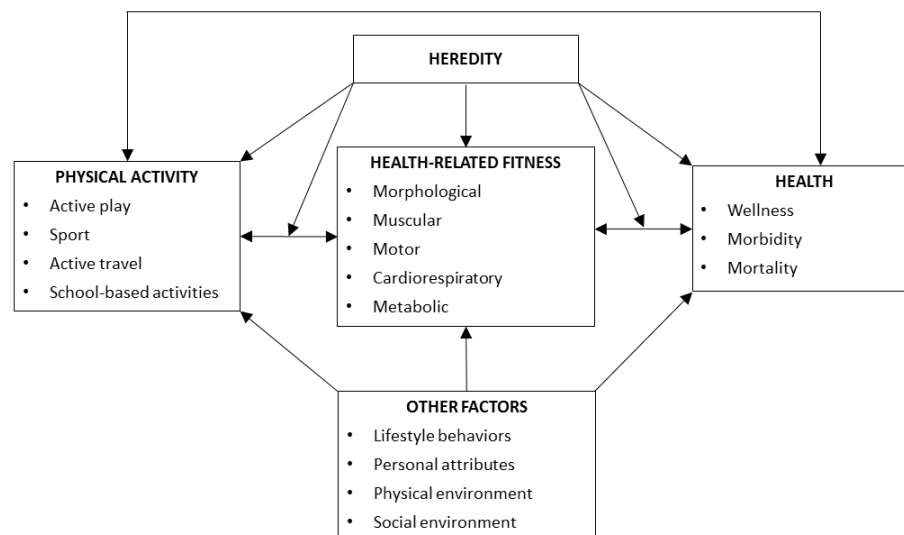


Figure 1.1 Modified version of Bouchard and Shephard's health-related fitness model.¹⁷ Reproduced with permission.

1.2.2 Ecological model

Bauman and colleagues' model suggests that environmental factors are present at the individual, interpersonal, environmental, regional/national and global levels, and that these nested environmental factors may influence behaviours differently across the life course (Figure 1.2).¹⁸ The model outlines environmental pathways that may influence PA levels, fitness and health at the individual level. For instance, effective healthy public policy at the regional and national levels could have an impact on our social and built environments, and in turn, this impact may influence interpersonal relationships and individual health through a top-down mechanism. Conversely, the surveillance of population health indicators (i.e., population CRF levels) provides data that can help identify the impact of implemented public policy and interventions aimed at improving health through a bottom-up mechanism. Additionally, surveillance indicators may help identify populations at risk of low health providing a way to promote equity and guide healthy public policies, interventions, and government investments.

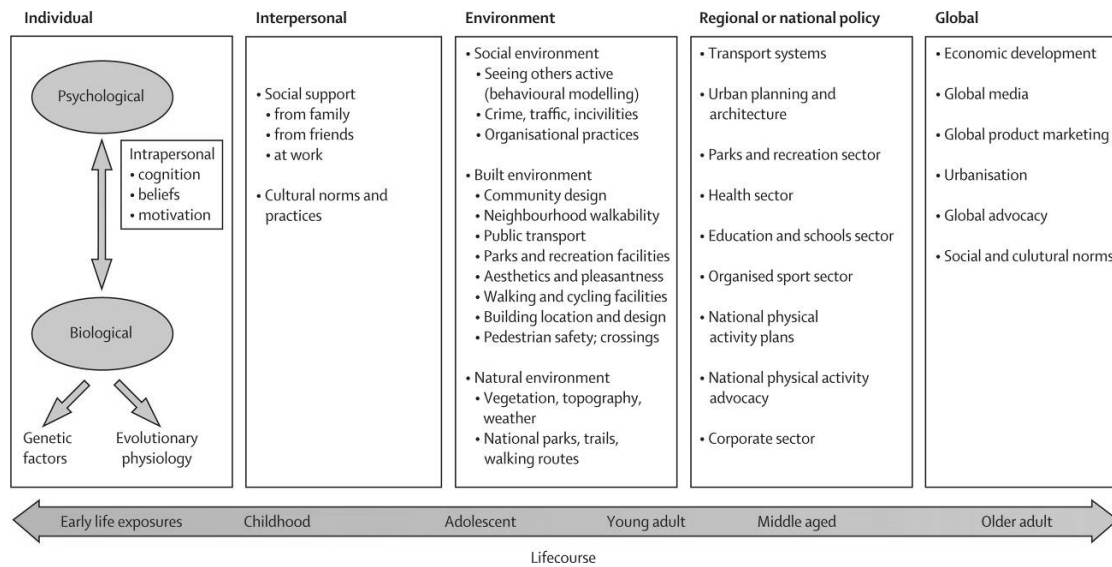


Figure 1.2 Bauman and colleagues' ecological model.¹⁸ Reproduced with permission.

1.3 Research Objectives

The purpose of this dissertation was to evaluate 20mSRT performances (i.e., CRF) as an important population health indicator, and the utility of the 20mSRT as a potential health surveillance instrument among children and youth. To achieve this purpose the following specific objectives were identified:

- 1) Perform a systematic review of the literature to identify the associations between 20mSRT and health markers among children and youth aged 5–17 years old;
- 2) Determine a criterion-referenced minimum 20mSRT performance to achieve health benefits among children and youth aged 9–17 years old;
- 3) Determine the association between cardiorespiratory fitness and components of physical literacy among Canadian children and youth aged 8–12 years old;
- 4) Determine the international variability in 20mSRT performance across countries in children and youth aged 9–17 years old;

- 5) Identify strategies to help promote cardiorespiratory fitness and the 20mSRT as a population health surveillance indicator.

1.4 Dissertation Organization and Outline

This dissertation is organized in an article format following the University of Ottawa dissertation guidelines. A brief overview of each chapter is provided below:

- Chapter 1 provides a brief introduction, overview, and justification of the dissertation topic.
- Chapter 2 presents a brief literature review that compares and contrasts the measurement of PA and CRF for population health surveillance.
- Chapter 3 is the first manuscript of this dissertation, entitled: *Systematic review of the relationship between 20 m shuttle run performance and health markers in children and youth*, submitted to the *Journal of Science and Medicine in Sport*. This manuscript is currently under peer-review.
- Chapter 4 is the second manuscript of this dissertation, entitled: *Review of criterion-referenced standards for cardiorespiratory fitness: What proportion of 142 026 international children and youth are apparently healthy?* This manuscript was accepted for publication with the *British Journal of Sports Medicine*.
- Chapter 5 is the third manuscript of this dissertation, entitled: *Cardiorespiratory fitness is associated with physical literacy among Canadian children aged 8 to 12 years*, submitted to *BMC Public Health*. This manuscript is currently under peer-review.

- Chapter 6 is the fourth manuscript of this dissertation, entitled: *International variability in 20 m shuttle run performance in children and youth: Who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results*. This manuscript was published online on September 20th, 2016 with the *British Journal of Sports Medicine*.
- Chapter 7 is the fifth and final manuscript of this dissertation, entitled: *Making a case for cardiorespiratory fitness surveillance among children and youth*, submitted to *Exercise and Sport Science Reviews*. This manuscript is currently under peer-review as an invited ‘perspectives for progress’ manuscript.
- Finally, Chapter 8 provides a general discussion, including the strengths, limitations, and future implications of the dissertation results.
- Supplemental Manuscript 1 provides a related manuscript entitled: *International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries*. This manuscript was published online on April 23rd, 2016 with the *British Journal of Sports Medicine*.
- Supplemental Manuscript 2 provides a second related manuscript entitled: *Temporal trends in the cardiorespiratory fitness of 965,264 children and youth representing 19 countries since 1981*, submitted to the *British Journal of Sports Medicine*. This manuscript is currently under peer-review.

1.5 References

1. GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*. 2016; 388:1459–544.
2. Salomon JA, Wang H, Freeman M, Vos T, Flaxman AD, Lopez AD, et al. Healthy life expectancy for 187 countries, 1990-2010: a systematic analysis for the Global Burden Disease Study 2010. *Lancet*. 2012; 380:2144–62.
3. Bloom DE, Cafiero ET, Jané-Llopis E, et al. The global economic burden of non-communicable disease. Geneva: World Economic Forum. 2011.
4. Berenson GS, Srinivasan SR, Bao W, Newman WP, Tracey RE, Wattigney WA. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults: The Bogalusa Heart Study. *New England Journal of Medicine*. 1998; 338(23):1650–5.
5. Dollman J, Norton K, Norton L. Evidence for secular trends in children's physical activity behaviour. *British Journal of Sports Medicine*. 2005; 39:892–7.
6. Ekelund U, Tomkinson G, Armstrong N. What proportion of youth are physically active? Measurement issues, levels and recent time trends. *British Journal of Sports Medicine*. 2011; 45:859–65.
7. Tomkinson GR, Olds TS. Secular changes in pediatric aerobic fitness test performance: The global picture. *Medicine and Sport Science*. 2007; 50:46–66.
8. Tremblay MS, Shields M, Laviolette M, Craig CL, Janssen I, Connor Gorber S. Fitness of Canadian children and youth: results from the 2007-2009 Canadian Health Measures Survey. *Health Reports*. 2010; 21(1):7–20.
9. Nelson MC, Neumark-Stzainer D, Hannan PJ, Sirard JR, Story M. Longitudinal and secular trends in physical activity and sedentary behavior during adolescence. *Pediatrics*. 2006; 118(6):e1627
10. Alwan A, MacLean DR, Riley LM, d'Espaignet ET, Mathers CD, Stevens GA, et al. Monitoring and surveillance of chronic non-communicable diseases: progress and capacity in high-burden countries. *Lancet*. 2010; 376:1861–68.
11. Centers for Disease Control and Prevention. Public health surveillance in the United States: Evolution and challenges. *Morbidity and Mortality Weekly Report*. 2012; 61(Suppl 1):3–9.
12. Léger L, Lambert J, Goulet A, Rowan C, Dinelle Y. Capacité aérobie des Québécois de 6 à 17 ans – test navette de 20 mètre avec paliers de 1 minute. *Canadian Journal of Applied Sport Science*. 1984; 9:64–9.

13. Léger LA, Lambert J. A maximal multistage 20-m shuttle run test to predict $\text{VO}_{2\text{max}}$. *European Journal of Applied Physiology and Occupational Physiology*. 1982; 49:1–12.
14. Hall HI, Correa A, Yoon PW, Braden CR. Lexicon, definitions, and conceptual framework for public health surveillance. *Morbidity and Mortality Weekly Report*. 2012; 61(Suppl):10–4.
15. Tremblay MS, Barnes JD, Gonzalez SA, Katzmarzyk PT, Onywera VO, Reilly JJ, et al. Global matrix 2.0: report card grades on the physical activity of children and youth comparing 38 countries. *Journal of Physical Activity and Health*. 2016; 13(Suppl 2):S343–S366.
16. Hallal PC, Andersen LB, Bull F, Guthold R, Haskell W, Ekelund U. Global physical activity levels: Surveillance progress, pitfalls, and prospects. *Lancet*. 2012; 380:247–57.
17. Bouchard C, Shephard RJ. Physical activity, fitness and health: the model and key concepts. In: Bouchard C, Shephard RJ, Stephens T, eds. *Physical Activity, Fitness and Health. International Proceedings and Consensus Statement*. Champaign III: Human Kinetics; 1994:77–88.
18. Bauman A, Reis RS, Sallis JF, Wells JC, Loos RJF, Martin BW, et al. Correlates of physical activity: why are some people physically active and others not? *Lancet*. 2012; 380:258–71.
19. Shephard RJ. Physical activity, fitness, and health: The current consensus. *Quest*. 1995; 47:288–303.
20. Kristensen PL, Moeller NC, Korsholm L, Kolle E, Wedderkopp N, Froberg K, et al. The association between aerobic fitness and physical activity in children and adolescents: The European Youth Heart Study. *European Journal of Applied Physiology*. 2010; 110:267–75.
21. Steele RM, Brage S, Corder K, Wareham NJ, Ekelund U. Physical activity, cardiorespiratory fitness, and metabolic syndrome in youth. *Journal of Applied Physiology*. 2008; 105:342–51.

CHAPTER 2: REVIEW OF LITERATURE

A strong population health surveillance indicator should demonstrate three main attributes. First, the indicator must have a robust and valid measurement instrument that can be applied with relative ease across a large population, while using minimal resources. Second, the indicator must not vary considerably from day-to-day, described as intra-individual variability. Third, the indicator must be substantially associated with health markers of interest. The following review will describe these attributes for physical activity (PA) and cardiorespiratory fitness (CRF) measurement instruments in an effort to compare their utility as a population surveillance indicator to identify variations in the health of school-aged children and youth.

2.1 Measurement Instruments

There are several field-based measurement instruments for CRF¹⁹ and PA² as described below. This section will conclude by selecting the most suitable PA and CRF field-based measure for population-based assessment, followed by a description of the selected instruments' validity.

2.1.1 Physical activity

There are several different measurement instruments that quantify PA such as self-report methods, pedometers, and accelerometers. Bauman *et al* provide a summary of several self-report instruments to estimate PA levels across populations.¹⁷ These self-

report methods are effective at estimating global physical activity levels,⁴ but, like any self-report measure, they are subject to recall bias (particularly apparent among children and youth),⁵ measurement error, and the instruments themselves often lack reliability and validity.⁶

A pedometer is a device that attaches to the hip and measures the number of steps taken per day. Although pedometers are relatively inexpensive, they do not capture vital information regarding the type, frequency, duration, and intensity of PA.⁷ However, new pedometers can estimate moderate-to-vigorous PA through step frequency.⁸

Accelerometers, also a device that attaches to the hip, ankle or wrist, solve many of the limitations associated with pedometers. Accelerometers are able to capture data on the frequency, duration, and intensity of PA;⁹ however, these devices are expensive and have limitations that relate to reliability,¹⁰ insensitivity to certain activities,¹¹ participant compliance,^{9,12} and comparability across studies.⁹ However, accelerometers are constantly improving, and thus, are becoming a common tool to assess PA in large groups of children and youth.¹³ For this reason, accelerometers are the choice PA measurement for this chapter.

2.1.1.1 *Validity of accelerometers*

Validity can be described as the extent to which an accelerometer measures the true PA patterns of an individual.¹⁴ The validity of accelerometers are derived from several components: accelerometer model, epoch length, cut-points, treatment of non-wear time, definition of a valid day, and requirements for minimum number of valid wear days.¹⁵ Currently, the regular application of five (Freedson,¹⁶ Puyau,¹⁷ Treuth,¹⁸

Mattocks,¹⁹ and Evenson²⁰) ActiGraph prediction equations to estimate cut-points for PA intensity levels and sedentary time represent the largest threat to the validity of accelerometers — described as the ‘cut point conundrum’.^{21,22} This conundrum plays a significant role in hindering research efforts to understand PA behaviours among children and youth. There is a need to arrive at consensus on the most valid approach to estimate PA using accelerometers in order to help compare PA levels across studies and populations.²¹

2.1.2 Cardiorespiratory fitness

CRF is defined as maximal aerobic power ($\dot{V}O_{2\text{peak}}$)²³ - “the greatest rate at which a person is able to consume oxygen during sustained, exhaustive exercise” (18, p. 113).¹⁹ The gold standard measurement for $\dot{V}O_{2\text{peak}}$ is a laboratory-based, maximal, graded treadmill or cycle ergometer exercise test.²³ Although the gold standard provides the most valid measurement of CRF, it is expensive and not feasible for assessing large groups of children. Instead, field-based measurements that demonstrate acceptable validity against the gold standard represent an attractive alternative for population-based assessment.²⁴

There are three general categories of CRF field-based measurements: (a) distance runs, (b) timed runs, and (c) endurance shuttle runs. Together, distance and timed runs provide an accurate measurement of CRF when the distance is greater than 1600 m or longer than 9 minutes.²⁵ Children also tend to improve their distance and timed runs up until the age of 15 years for boys and 13 years for girls, which may be related to the fractional utilization of oxygen (the fraction of $\dot{V}O_{2\text{peak}}$ that can be sustained for any

given length of time), mechanical efficiency, $\dot{V}O_{2\text{peak}}$ kinetics and anaerobic capacity (relatively more important for shorter distance runs), as well as affective (motivation) and cognitive (pacing) factors.²⁴ Additional limitations associated with distance and timed run performance include methodological factors such as the practice effect, weather conditions, ground conditions, running course standardization, footwear/clothing, and the test name (e.g., a run test vs. a run/walk test).²⁴

The endurance shuttle run, first described in 1982 by Léger and Lambert,¹² comprised 20 m shuttles (or laps) with 2-minute stages, where participants completed as many stages as possible in a standardized, progressive format. However, since 1982 the stage length was reduced from 2-minutes to 1-minute, and 3 different 20m shuttle run test (20mSRT) versions have emerged in the literature (see Table 1). The Léger 20mSRT is the most widely used, international protocol. The Eurofit protocol is common across Europe and Oceania, and results derived are comparable to the Léger protocol as the only difference between protocols is the initial running speed (Stage 1). The use of protocol 3 (Queen's University of Belfast; QUB) is rare and results are difficult to compare with other protocols because the running speed is 0.5 km/h slower across all stages.

Table 2.1 Running speed (km/h) and number of laps for each stage of the 20mSRT according to the three major protocol variants. Published with permission from Tomkinson *et al.*³³

Stage no. ^a	Protocol 1 (Léger et al.) ¹¹		Protocol 2 (Eurofit et al.) ²⁸⁻³¹		Protocol 3 (QUB) ³²	
	speed (km/h)	no. of laps	speed (km/h)	no. of laps ^b	speed (km/h)	no. of laps
1	8.5	7	8.0	7	8.0	7
2	9.0	8	9.0	8	8.5	7
3	9.5	8	9.5	8	9.0	8
4	10.0	8	10.0	8	9.5	8
5	10.5	9	10.5	9	10.0	8
6	11.0	9	11.0	9	10.5	9
7	11.5	10	11.5	10	11.0	9

^a The first stage is sometimes called ‘Stage 1’ (if ‘Stage 1’ is called at the start of the stage), and sometimes ‘Stage 0’ (if ‘Stage 1’ is called at the end of the first stage).

^b In some versions of Protocol 2 (e.g., the Eurofit, Australian Coaching Council) the number of laps per stage varies in places.

QUB = Queens University of Belfast.

In comparison with timed and distance runs, the 20mSRT provides several benefits. The progressive testing format of all 20mSRT protocols provide standardized pacing, eliminating the cognitive pacing component associated with timed and distance runs. These assessments also allow for indoor testing, solving the limitations of running course standardization and climate. Children also experience an age-related improvement in 20mSRT performance,³³ which is similar, but more gradual than the age-related improvements reported for distance runs.^{24,33} For these reasons, among others, there exists a general consensus that the 20mSRT is the preferred field-based measure of CRF among school-aged children and youth.³³⁻³⁴ Thus, the 20mSRT is the choice measurement of CRF for this chapter.

2.1.2.1 *Validity of the 20m shuttle run test*

To determine if the 20mSRT is a ‘good’ estimate of CRF among children and youth Mayorga-Vega *et al* performed a meta-analysis of 57 studies that reported on the criterion-related validity of the 20mSRT.³⁶ The meta-analysis results indicated that the Léger protocol had a greater average criterion-related validity coefficient when compared to the Eurofit and QUB protocols, suggesting that the use of protocols other than the Léger protocol is not justified.³⁶ Furthermore, the criterion-related validity of the 20mSRT was considered moderate-to-high ($r_p = 0.66-0.84$) among children and youth, but the coefficients improved greatly ($r_p = 0.78-0.95$) when combined with other variables (e.g., age, sex, body mass index) to estimate CRF.³⁶ As a result, several prediction equations have been developed to help predict $\dot{V}O_{2peak}$ from the maximal running speed attained during the 20mSRT.

Ruiz *et al* ⁴¹ have identified four, among many other 20mSRT prediction equations in the literature: Léger,³⁸ Barnett,³⁸ Matsuzaka,⁴⁰ and Ruiz.⁴¹ To determine the criterion-related validity of each equation the predicted $\dot{V}O_{2peak}$ values are compared to objectively measured $\dot{V}O_{2peak}$. The results suggested that all equations, except for Barnett,³⁹ significantly underestimated $\dot{V}O_{2peak}$, and as a result, the study promoted the use of the Barnett equation. However, one study suggested the use of the Ruiz³⁷ and/or Barnett³⁹ equations;³⁴ and another study by Moreira *et al*⁴² concluded that the best equations to estimate $\dot{V}O_{2peak}$ were that of Barnett³⁹ and/or Matsuzaka.⁴⁰ Inconsistencies throughout the literature make it difficult to identify the strongest prediction equation. However, it is conceivable that researchers may select a given prediction equation based on the input variables available. For instance, the Léger equation requires speed of the

last completed stage and age as inputs, whereas the Ruiz equation requires sex, age, weight, height, and stage attained – variables that may not be available to all researchers. Thus, regardless of the prediction equation used, it is important to recognize that systematic error is inevitably introduced, which may result in the over- or under-estimation of $\dot{V}O_{2\text{peak}}$ among children and youth.

2.2 Intra-Individual Variability

Large intra-individual variability (i.e., day-to-day variability) can cause problems when estimating the proportion of the population meeting recommendations and/or guidelines.⁴³ As a result, there are fundamental differences in measuring PA and CRF as described below.

2.2.1 Physical activity versus cardiorespiratory fitness

PA is a behaviour that potentially varies substantially from day-to-day, causing high intra-individual variability. Therefore, to obtain a reliable PA measurement it is necessary to collect data over 4 to 9 days (including one weekend day) to achieve a reliability coefficient of 0.8.^{12,11,44} Intra-individual variability of PA causes considerable measurement difficulty that results in an increase in financial cost,¹⁰ an increase in participant burden, a decrease in compliance,¹² and difficulties with cross-country comparability due to climate and weather variations.⁴³

In contrast, CRF is a trait that is innate, and modified by PA, resulting in relatively low intra-individual variability. Thus, CRF varies considerably less than PA on a day-to-day basis. Léger *et al* determined that the test-retest reliability of the

20mSRT,³⁸ after one week, was $r = 0.89$ in children and youth (8 to 19 years). A replication study with 12 to 15 year-old youth concluded that reliability was slightly better ($r = 0.93$) with older children.⁴⁵ Additionally, Tomkinson and Olds conducted a 20mSRT review of 15 reliability studies and reported a sample weighted average coefficient of determination (r^2) of 0.73 for reliability among children and youth.⁶ These studies indicate that the 20mSRT has strong to near perfect reliability in children and youth between 8 to 19 years of age.

High intra-individual variability of PA presents difficulties in describing relationships with health. At times, it is difficult to obtain a reliable estimation of PA, resulting in an uncertain and possibly erroneous association with health markers. In contrast, the low intra-individual variability of CRF often provides a clear association with health markers.⁴⁷ Thus, intra-individual variability is important to consider for population health surveillance indicators.

2.3 Association with Health Markers

PA and CRF are independently associated with disease risk factors among children and with chronic disease outcomes among adults,^{20–50} highlighting the need to treat CRF and PA as separate variables in analyzing their associations with health. In addition, PA, CRF, and health vary across the life course.⁵¹ Therefore, taking a life course approach to understanding the associations between PA and CRF with health among children and adults might provide insight into current and future population health. The following sections will describe the associations of PA and CRF with health, and how these associations differ across the life course.

2.3.1 Adulthood

In adults, chronic disease health outcomes are prevalent throughout the population.⁵⁰ Thus, it is possible to determine the associations between PA and CRF with a variety of disease health outcomes, including mortality. High levels of CRF associate moderately to strongly with a decrease in cardiovascular disease,⁵² type 2 diabetes,⁵³ and all-cause mortality in adults.^{54,55} There is also strong evidence that PA plays a role in reducing cardiovascular disease, type 2 diabetes, breast and colon cancer, depression, and the loss of cognitive functions.⁵⁶ Although both PA and CRF are beneficial for health in adults, Blair *et al* indicated that the “gradient for various health outcomes is steeper across categories of CRF than across PA groups” (p.S396).⁵⁰ Thus, CRF in adults is possibly a better population health indicator than PA levels.

2.3.2 Childhood and adolescence

In children and youth it is difficult to determine the association between fitness and adverse health outcomes (e.g., cancer, cardiovascular disease, metabolic syndrome), because these chronic diseases don’t often manifest until the fifth decade of life or later.²⁰ As a result, adverse health outcomes are generally not observed in children and youth, but evidence suggests that risk factors for chronic diseases (e.g., high blood pressure, dyslipidemia, obesity) may manifest in childhood.³ This makes it possible to determine the relationship of CRF and PA with disease risk factors (health markers) among children and youth.

2.3.2.1 *Physical activity and health*

In recent years two comprehensive systematic reviews were commissioned to inform the development of PA guidelines in Canada,^{58,59} and to ultimately help shape and decipher our knowledge of the health benefits of PA among school-aged children and youth. The original review by Janssen and LeBlanc⁵⁸ identified that PA was associated with numerous health benefits (e.g., favourable cholesterol, blood lipids, blood pressure, mental health) among children and youth, and that many of these associations displayed a dose-response gradient. Thus, more PA is associated with greater health benefits. The study concluded with 3 main recommendations, one of which being that “children and youth 5-17 years of age should accumulate an average of at least 60 minutes per day and up to several hours of at least moderate intensity physical activity...” (p. 11).⁵⁸ These recommendations and the overall health benefits of PA were later supported by a systematic review of objectively measured PA.⁵⁹ The findings also highlighted the benefits of light intensity PA, albeit at a reduced level of impact on health than moderate and vigorous intensity PA. Overall, these studies support the need for children and youth to attain a minimum benchmark of 60 minutes of PA per day in order to experience health benefits – a recommendation that is fairly consistent across the world.⁶⁰

2.3.2.2 *Cardiorespiratory fitness and health*

There is one notable narrative synthesis of the literature investigating the association between multiple measures of CRF and health benefits among school-aged children and youth. Similar to PA, this review highlighted the numerous health benefits associated with high levels of CRF among children, which include favourable effects on

depression, anxiety, and self-esteem.⁶¹ However, when compared with PA, CRF levels among children and youth may be a stronger predictor of metabolic and cardiovascular disease risk.^{62–64} This highlights the potential importance of measuring CRF to help understand population health.

Overall, the evidence suggests that CRF is favourably associated with health markers among children and youth. However, the literature is lacking robust systematic reviews that would help shape and decipher the overall understanding of the relationship between CRF and health. Furthermore, the limited efforts to consolidate the literature often include several different CRF measurement tools (i.e., various field-based and objective measures), which may present inconsistencies as the construct validity of different measures often vary substantially. Thus, it may be beneficial to further explore the associations between a single measure of CRF (i.e., the 20mSRT) and health markers among school-aged children and youth.

2.3.3 Tracking from childhood to adulthood

With longitudinal studies it is possible to identify if PA and/or CRF levels in childhood track (correlate) with PA and/or CRF in adulthood. Tracking studies suggest that correlations do not only rely on the strength of the measure, but also the amount of time between observations, with longer periods of time resulting in smaller tracking coefficients.⁶⁵ In general, PA tracking from adolescence to young adulthood is low, with coefficients between 0.20 and 0.31.⁶⁶ In contrast, CRF tracks moderately well from adolescence to young adulthood with coefficients around 0.50.^{66,67} Accounting for

roughly the same timespan (adolescent to young adulthood) CRF seems to track better than PA.

Another tracking method is to determine if PA and/or CRF in adolescence is associated with future disease in adulthood. Twisk *et al* assessed CRF and PA in 13 and 16 year old boys and girls, and measured disease prevalence in the same individuals 20 years later as part of the Amsterdam Growth and Health Longitudinal Study.⁶⁶ The researchers determined that CRF was related to future cardiovascular disease risk profile, whereas PA levels were not. Furthermore, Ruiz *et al* conducted a systematic review indicating that there was strong evidence suggesting that high CRF levels in childhood and adolescence was related to a healthier cardiovascular profile in adulthood.⁶⁸ These results were also supported by two recent cohort studies of more than 700,000 Swedish adolescents with a 29–34 year follow-up which suggested that high CRF levels in late adolescence was associated with a reduced risk of myocardial infarction and premature mortality in adulthood.^{69,70} Therefore, in comparison with PA surveillance, the surveillance of CRF among children and youth might not only provide a better summative measure of current health status, but it may also provide a good indication of future population health and chronic disease burden.

2.4 Benefits and Drawbacks

Over the past 2 decades there has been strong support for PA research. It is possible that official PA guidelines for children and youth provided a structural foundation to help support the PA research movement. The guidelines have been an effective knowledge translation tool which helped describe the importance of PA for

public health and well-being, while also setting an activity target of at least 60 minutes per day.⁷¹ This minimum PA target also helped describe an underlying problem where not enough children are meeting the guidelines, and thus, not enough children are physically active. It is conceivable that, as a foundation, the PA guidelines have helped spur other PA initiatives, such as the Toronto Charter for Physical Activity⁷² and the Bangkok Declaration on Physical Activity.⁷³ There have also been several important international surveillance efforts to describe PA levels among children and youth, and by association, the health of populations.^{4,74-76} Although PA is an important health-related behaviour, the measurement of PA as a population health surveillance indicator has many potential drawbacks related to measurement as previously discussed (e.g., cost, validity,^{21,22} participant burden and compliance,^{9,12} administration and data interpretation expertise, cross-country comparisons⁹) and intra-individual variability.¹⁰

In contrast, the measurement of CRF as a population health surveillance indicator demonstrates several important benefits. For instance, the 20mSRT is a cost-effective, valid and simple measure that can assess large groups of children and youth, simultaneously.²⁴ At a population level, CRF has the ability to describe both the current health of children and youth as well as potential future non-communicable disease burden. The simplicity of CRF measurement can also help screen and identify groups of children in need of PA-based interventions to improve health. However, CRF-based research is missing fundamental population-based studies that are quite apparent in PA-based research (i.e., guidelines). In addition, there has been a relative lack of CRF research since the 1990's which is potentially due to fitness testing stigma that has developed over the years in western culture, and perhaps in other regions worldwide.

Thus, there is a need to overcome the cultural stigma of fitness testing by establishing the importance of measuring CRF among children and youth as a population health initiative. Perhaps fundamental research, similar to PA-based research, can help further develop the measurement of CRF as a population health surveillance indicator.

2.5 Conclusion

Population-based surveillance of PA levels among children and youth has been a notable public health accomplishment over the course of the 21st century. However, evidence suggests that the measurement of CRF is also a strong population health indicator, independent of PA levels, among children and youth. There is certainly a need for more comprehensive efforts to better describe population health worldwide in an effort to help monitor and evaluate existing health promoting policies and interventions, and/or to help identify and guide the implementation of new health promoting policies and interventions. Thus, the measurement of CRF may be an important feature to help complement existing PA surveillance efforts, or in some circumstances even replace it. For instance, CRF measurement could be a cost-effective population health surveillance indicator for countries and regions that lack the financial resources to implement expensive PA surveillance. However, there are several CRF literature gaps, as identified through this review, which may have created a barrier to implement national and international CRF measurement as a population health surveillance indicator. Thus, the remainder of this dissertation will make efforts to help fill the CRF literature gaps needed to support regional, national, and international CRF population health surveillance efforts.

2.6 References

1. Institute of Medicine (IOM). Fitness measures and health outcomes in youth. Washington, DC: The National Academics Press; 2012.
2. Lindsay AR, Hongu N, Spears K, Idris R, Dyrek A, Manore MM. Field assessments for obesity prevention in children and adults: physical activity, fitness, and body composition. *Journal of Nutrition Education and Behavior*. 2014; 46(1):43–53.
3. Bauman A, Phongsavan P, Schoeppe S, Owen N. Physical activity measurement – a primer for health promotion. *International Union for Health Promotion and Education*. 2006; 13(2):92–103.
4. Hallal PC, Andersen LB, Bull F, Guthold R, Haskell W, Ekelund U. Global physical activity levels: Surveillance progress, pitfalls, and prospects. *The Lancet*. 2012; 380:247–257.
5. Welk GJ, Corbin CB, Dale D. Measurement issues in the assessment of physical activity in children. *Research Quarterly for Exercise and Sport*. 2000; 71(Suppl 2):S59–S73.
6. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaire. *British Journal of Sports Medicine*. 2003; 37:197–206.
7. Corder K, Ekelund U, Steele RM, Wareham NJ, Brage S. Assessment of physical activity in youth. *Journal of Applied Physiology*. 2008; 105:977–87. doi:10.1152/00094.2008
8. Saunders TJ, Gray CE, Borghese MM, McFarlane A, Mbonu A, Ferraro ZM, et al. Validity of SC-StepRx pedometer-derived moderate and vigorous physical activity during treadmill walking and running in a heterogeneous sample of children and youth. *BMC Public Health*. 2014; 14(1):519. doi:10.1186/1471-2458-14-519
9. Rowlands AV. Accelerometer assessment of physical activity in children: An update. *Pediatric Exercise Science*. 2007; 19:252–66.
10. Chen KY, Brasset Jr. DR. The technology of accelerometry-based activity monitors: Current and future. *Medicine and Science in Sports and Exercise*. 2005; 37(Suppl 11):S490–S500.
11. Trost SG. Objective measurement of physical activity in youth: Current issues, future directions. *Exercise and Sport Sciences Reviews*. 2001; 29(1):32–6. doi:0091-6631/2901/32-36
12. Colley RC, Garriguet D, Janssen I, Craig CL, Tremblay MS. Physical activity of Canadian children and youth: Accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. *Health Reports*. 2011; 22(1):15–23.

13. Mattocks C, Ness A, Leary S, Tilling K, Blair SN, Shield J, et al. Use of accelerometers in a large field-based study of children: Protocols, design issues, and effects on precision. *Journal of Physical Activity and Health*. 2008; 5(Suppl 1):S98–S111.
14. Drost EA. Validity and reliability in social science research. *Education Research and Perspectives*. 2011; 38(1):105–23.
15. Cain KL, Sallis JF, Conway TL, Dyck DV, Calhoun L. Using accelerometers in youth physical activity studies: A review of methods. *Journal of Physical Activity and Health*. 2013; 10:437–50.
16. Freedson PS, Pober D, Janz KF. Calibration of accelerometer output for children. *Medicine and Science in Sports and Exercise*. 2005; 37(Suppl 11):S523–S530.
17. Puyan MR, Adolph AL, Vohra FA, Butte NF. Validation and calibration of physical activity monitors in children. *Obesity Research*. 2002; 10:150–7.
18. Treuth MS, Schmitz K, Catellier DJ, McMurray RG, Murray DM, Joao Almeida M, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. *Medicine and Science in Sports and Exercise*. 2004; 36(7):1259–66.
19. Mattocks C, Leary S, Ness A, Deere K, Saunders J, Tilling K, et al. Calibration of an accelerometer during free-living activities in children. *International Journal of Pediatric Obesity*. 2007; 2:218–26.
20. Evenson KR, Cattellier D, Gill K, Ondrak K, McMurray RG. Calibration of two objective measures of physical activity for children. *Journal of Sports Science*. 2008; 26:1557–65.
21. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. *Medicine and Science in Sports and Exercise*. 2011; 43(7):1360–8.
22. Trost SG. Measurement of physical activity in children and adolescents. *American Journal of Lifestyle Medicine*. 2007; 1:299–314.
23. Jackson AS. The evolution and validity of health-related fitness. *Quest*. 2006; 58:160–75.
24. Tomkinson GR, Olds TS. Field tests of fitness. In: Armstrong N, van Mechelen W, eds. *Paediatric Exercise Science and Medicine*. New York, NY: Oxford University Press; 2008:109–24.
25. Krahenbuhl GS, Pangrazi RP, Petersen GW, Burkett LN, Schneider MJ. Field testing of cardiorespiratory fitness in primary school children. *Medicine in Science and Sports*. 1978; 10:208–13.

26. Léger LA, Lambert J. A maximal multistage 20-m shuttle run test to predict $\text{VO}_{2\text{max}}$. *European Journal of Applied Physiology and Occupational Physiology*. 1982; 49:1–12.
27. Léger L, Lambert J, Goulet A, Rowan C, Dinelle Y. Capacité aérobie des Québécois de 6 à 17 ans – test navette de 20 mètre avec paliers de 1 minute. *Canadian Journal of Applied Sport Science*. 1984; 9:64–9.
28. Council of Europe. Eurofit: handbook for the Eurofit tests of physical fitness. Rome: Council of Europe; 1988.
29. Australian Sports Commission. 20m shuttle run test: a progressive shuttle run test for measuring aerobic fitness. Belconnen (ACT): Australian Coaching Council; 1999.
30. Brewer J, Ramsbottom R, Williams C. Multistage fitness test: a progressive shuttle-run test for the prediction of maximal oxygen uptake. Leeds: National Coaching Foundation; 1988.
31. Cooper Institute for Aerobic Research. The Prudential FITNESSGRAM test administration manual. Dallas (TX): Cooper Institute for Aerobic Research; 1992.
32. Riddoch CJ. The Northern Ireland health and fitness survey – 1989: the fitness, physical activity, attitudes and lifestyle of Northern Ireland post-primary schoolchildren. Belfast: The Queen's University of Belfast; 1990.
33. Tomkinson GR, Léger LA, Cazorla G. Secular trends in the performance of children and adolescents (1980 – 2000): An analysis of 55 studies of the 20 m shuttle run test in 11 countries. *Sports Medicine*. 2003; 33(4):285–300. doi:0112-1642/03/0004-0285
34. Castro-Piñero J, Artero EG, España-Romero V, Ortega FB, Sjöström M, Suni J, et al. Criterion-related validity of field-based fitness tests in youth: A systematic review. *British Journal of Sports Medicine*. 2010; 44:934–43.
35. Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9 - 17 –year-old Australians since 1985. *British Journal of Sports Medicine*. 2013; 47:98–109. doi: 10.1136/bjsports-2011-090218
36. Mayorga-Vega D, Aguilar-Soto P, Viciano J. Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *Journal of Sports Science and Medicine*. 2015; 14: 536–47.
37. Ruiz JR, Silva G, Oliveira N, Ribeiro JC, Oliveira JF, Mota J. Criterion-related validity of the 20-m shuttle run test in youth aged 13–19 years. *Journal of Sports Sciences*. 2009; 27(9):899–906.
38. Léger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. *Journal of Sport Sciences*. 1988; 6:93–101.

39. Barnett A, Chan LYS, Bruce IC. A preliminary study of the 20-m multistage shuttle run as a predictor of peak VO₂ in Hong Kong Chinese students. *Pediatric Exercise Science*. 1993; 5:42–50.
40. Matsuzaka A, Takahashi Y, Yamazoe M, Kumakura N, Ikeda A, Wilk B, et al. Validity of the multistage 20-m shuttle run test for Japanese children, adolescents, and adults. *Pediatric Exercise Science*. 2004; 16:113–25.
41. Ruiz JR, Ramirez-Lechuge J, Ortega FB, Castro-Piñero J, Benitez JM, Arauzo-Azofra A, et al. Artificial neural network-based equation for estimating VO₂max from the 20 m shuttle run test in adolescents. *Artificial Intelligence in Medicine*. 2008; 44:233–45.
42. Moreira C, Santos R, Ruiz JR, Vale S, Soares-Miranda L, Marques AI, et al. Comparison of different VO₂max equations in the ability to discriminate the metabolic risk in Portuguese adolescents. *Journal of Science and Medicine in Sport*. 2011; 14:79–84.
43. Ridley K, Olds T, Hands B, Larkin D, Parker H. Intra-individual variations in children's physical activity patterns: Implications for measurement. *Journal of Science and Medicine in Sport*. 2008; 12(5):568–72.
44. Fairclough SJ, Butcher ZH, Stratton G. Whole-day and segmented day physical activity variability of northwest England school children. *Preventive Medicine*. 2007; 44(5):421–5.
45. Liu NYS, Plowman SA, Looney MA. The reliability and validity of the 20-meter shuttle test in American students 12 to 15 years old. *Research Quarterly for Exercise and Sport*. 1992; 83(4):360–5.
46. Tomkinson GR, Olds TS. Secular changes in pediatric aerobic fitness test performance: The global picture. *Medicine and Sport Science*. 2007; 50:46–66.
47. Froberg K, Andersen LB. Mini review: Physical activity and fitness and its relations to cardiovascular disease risk factors in children. *International Journal of Obesity*. 2005; 29:S34–S39. doi:10.1038/sj.ijo.0803096
48. Steele RM, Brage S, Corder K, Wareham NJ, Ekelund U. Physical activity, cardiorespiratory fitness, and metabolic syndrome in youth. *Journal of Applied Physiology*. 2008; 105:342–51.
49. Eklund U, Anderssen SA, Froberg K, Sardinha LB, Andersen LB, Brage S. Independent associations of physical activity and cardiorespiratory fitness with metabolic risk factors in children: The European youth heart study. *Diabetologia*. 2007; 50:1832–40. doi:10.1007/s00125-007-0762-5
50. Blair SN, Cheng Y, Holder JS. Is physical activity or physical fitness more important in defining health benefits? *Medicine and Science in Sports and Exercise*. 2001; 33(Suppl 6):S370–S399. doi:0195-9131/01/3306-0379

51. Boreham C, Twisk J, Neville C, Savage M, Murray L, Gallagher A. Associations between physical fitness and activity patterns during adolescence and cardiovascular risk factors in young adulthood: The Northern Ireland Young Heart Project. *International Journal of Sports Medicine*. 2002; 23:S22–S26.
52. Carnethon MR, Gidding SS, Nehgme R, Sidney S, Jacobs DR, Liu K. Cardiorespiratory fitness in young adulthood and the development of cardiovascular disease risk factors. *JAMA*. 2003; 290(23):3092–3100. doi:10.1001/jama.290.23.3092
53. Lynch J, Helmrigh SP, Lakka TA, Kaplan GA, Cohen RD, Salonen R, et al. Moderately intense physical activity and high levels of cardiorespiratory fitness reduce the risk of non-insulin-dependent diabetes mellitus in middle-aged men. *JAMA Archives of Internal Medicine*. 1996; 156(12):1307–14.
54. Blair SN, Kohl HW, Paffenbarger RS, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality: A prospective study of healthy men and women. *JAMA*. 1989; 262(17):2395–2401. doi:10.1001/jama.1989.03430170057028
55. Laakkanen JA, Lakka TA, Rauramaa R, Kuhanen R, Venalainen JM, Salonen R, et al. Cardiovascular fitness as a predictor of mortality in men. *JAMA Archives of Internal Medicine*. 2001; 161(6):825–31.
56. Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Report. Washington, DC: U.S. Department of Health and Human Services; 2008.
57. Berenson GS, Srinivasan SR, Bao W, Newman WP, Tracey RE, Wattigney WA. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults: The Bogalusa Heart Study. *New England Journal of Medicine*. 1998; 338(23):1650–5.
58. Janssen I, LeBlanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity*. 2010; 7:40.
59. Poitras VJ, Gray CE, Borghese MM, Carson V, Chaput JP, Janssen I, et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology Nutrition and Metabolism*. 2016; 41:S197–S239.
60. Tremblay MS, Carson V, Chaput JP, Gorber SC, Dinh T, Duggan M, et al. Canadian 24-hour movement guidelines for children and youth: An integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology Nutrition and Metabolism*. 2016; 41(6):S311–S327.

61. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity*. 2008; 32:1–11.
62. Anderssen SA, Cooper AR, Riddoch C, Sardinha LB, Harro M, Brage S, et al. Low cardiorespiratory fitness is a strong predictor for clustering of cardiovascular disease risk factors in children independent of country, age and sex. *European Journal of Cardiovascular Prevention and Rehabilitation*. 2007; 14:526–31.
63. Hurtig-Wennlof A, Ruiz JR, Harro M, Sjöström M. Cardiovascular fitness relates more strongly than physical activity to cardiovascular disease risk factors in healthy children and adolescents: the European Youth Heart Study. *European Journal of Cardiovascular Preventive Rehabilitation*. 2007; 14:575–81.
64. Rizzo NS, Ruiz JR, Hurtig-Wennlof A, Ortega FB, Sjöström M. Relationship of physical activity, fitness, and fatness with clustered metabolic risk in children and adolescents: The European Youth Health Study. *Journal of Pediatrics*. 2007; 150:388–94. doi: 10.1016/j.jpeds.2006.12.039
65. Malina RM. Physical activity and fitness: Pathways from childhood to adulthood. *American Journal of Human Biology*. 2001; 13:162–72.
66. Twisk JW, Kemper HC, van Mechelen W. Tracking of activity and fitness and the relationship with cardiovascular disease risk factors. *Medicine and Science in Sports and Exercise*. 2000; 32(8):1455–61.
67. Twisk JWR, Kemper HC, van Mechelen W. The relationship between physical fitness and physical activity during adolescence and cardiovascular disease risk factors at adult age: The Amsterdam Growth and Health Longitudinal Study. *International Journal of Sports Medicine*. 2002; 23:S8–S14.
68. Ruiz JR, Castro-Piñero J, Artero EG, Ortega FB, Sjöström M, Suni J, et al. Predictive validity of health-related fitness in youth: A systematic review. *British Journal of Sports Medicine* Published Online First: 21 January 2009.
69. Högström G, Nordström A, Nordström P. High aerobic fitness in late adolescence is associated with a reduced risk of myocardial infarction later in life: a nationwide cohort study in men. *European Heart Journal*. 2014; 35:3133–40.
70. Högström G, Nordström A, Nordström P. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. *International Journal of Epidemiology* Published Online First: 20 December 2015.
71. Tremblay MS, Shephard RJ, Brawley LR, Cameron C, Graig CL, Duggan M, et al. Physical activity guidelines and guides for Canadians: facts and future. *Applied Physiology Nutrition and Metabolism*. 2007; 32:S218–S224.

72. Global Advocacy Council for Physical Activity, International Society for Physical Activity and Health. *The Toronto Charter for Physical Activity: A Global Call to Action*. 20 May 2010.
73. International Congress on Physical Activity and Public Health. *The Bangkok Declaration on Physical Activity for Global Health and Sustainable Development*. 16-19 November 2016.
74. Rütten A, Vuillemin A, Ooijendijk WT, Schena F, Sjöström M, Stahl T, et al. Physical activity monitoring in Europe. The European Physical Activity Surveillance System (EUPASS) approach and indicator testing. *Public Health Nutrition*. 2002; 6(4):377–84.
75. Tremblay MS, Barnes JD, Gonzalez SA, Katzmarzyk PT, Onywera VO, Reilly JJ, et al. Global matrix 2.0: report card grades on the physical activity of children and youth comparing 38 countries. *Journal of Physical Activity and Health*. 2016; 13(Suppl 2):S343–S366.
76. Cooper AR, Goodman A, Page AS, Sherar LB, Esliger DW, van Sluijs EMF, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). *International Journal of behavioral Nutrition*. 2015; 12:113.

CHAPTER 3: MANUSCRIPT 1

This manuscript has been written and formatted following the publication standards of the peer-reviewed *Journal of Science and Medicine in Sport*. It appears in its published format.

Title:

Systematic review of the relationship between 20 m shuttle run performance and health indicators among children and youth

Citation:

Lang JJ, Belanger K, Poitras V, Janssen I, Tomkinson GR, Tremblay MS. Systematic review of the relationship between 20 m shuttle run performance and health indicators among children and youth. *Journal of Science and Medicine in Sport*. 2017.

doi:10.1016/j.jsams.2017.08.002

Online link:

[http://www.jsams.org/article/S1440-2440\(17\)30990-8/fulltext](http://www.jsams.org/article/S1440-2440(17)30990-8/fulltext)



Contents lists available at ScienceDirect

Journal of Science and Medicine in Sport

journal homepage: www.elsevier.com/locate/jsams



Review

Systematic review of the relationship between 20 m shuttle run performance and health indicators among children and youth

Justin J. Lang^{a,b,*}, Kevin Belanger^a, Veronica Poitras^a, Ian Janssen^c,
Grant R. Tomkinson^{d,e}, Mark S. Tremblay^a

^a Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of Eastern Ontario Research Institute, Canada

^b PhD Program in Population Health, University of Ottawa, Canada

^c School of Kinesiology and Health Studies, Queen's University, Canada

^d Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health Sciences & Sansom Institute for Health Research, University of South Australia, Australia

^e Department of Kinesiology and Public Health Education, University of North Dakota, USA

ARTICLE INFO

Article history:

Received 22 January 2017

Received in revised form 14 June 2017

Accepted 1 August 2017

Available online xxx

Keywords:

Cardiorespiratory endurance

Aerobic fitness

Adiposity

Cardiomatabolic biomarkers

PACER

Beep test

ABSTRACT

Objective: This systematic review aimed to summarize research that assessed the associations between 20 m shuttle run test (20mSRT) performance and indicators of physiological, psychosocial and cognitive health among school-aged children and youth.

Design: Systematic review.

Methods: Five online databases were used to identify peer-reviewed studies published from 1980 to 2016. Studies were included if they matched these criteria: population (children and youth with a mean age of 5–17 years and/or in Grades 1–12), intervention/exposure (performance on the 20mSRT), and outcomes (health indicators: adiposity, cardiometabolic biomarkers, cognition, mental health, psychosocial health, self-esteem and physical self-perception, quality of life and wellbeing, bone health, musculoskeletal fitness, motor skill development, and injuries and/or harm). Narrative syntheses were applied to describe the results. A lack of homogeneity precluded a meta-analysis approach.

Results: Overall, 142 studies that determined an association between 20mSRT performance and a health indicator were identified, representing 319,311 children and youth from 32 countries. 20mSRT performance was favourably associated with indicators of adiposity, and some indicators of cardiometabolic, cognitive, and psychosocial health in boys and girls. Fewer studies examined the relationship between 20mSRT performance and measures of quality of life/wellbeing, mental health and motor skill development, and associations were generally inconsistent. The quality of the evidence ranged from very low to moderate across health indicators.

Conclusion and Implications: These findings support the use of the 20mSRT as a holistic indicator of population health in children and youth.

© 2017 Sports Medicine Australia. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Cardiorespiratory fitness (CRF) is measured as the highest rate of oxygen consumed ($\dot{V}O_2$) during a bout of maximal exercise to voluntary exhaustion, described as $\dot{V}O_{2peak}$ in children and youth.¹ CRF provides an indication of the function of the cardiovascular and respiratory systems and it provides a strong summative measure of physical health among children and youth.² Furthermore, youth with low CRF in late adolescence have an increased risk of

developing cardiovascular disease in adulthood³ and succumbing to premature mortality.⁴ For these reasons, CRF may be a strong population health indicator and surrogate measure that could help researchers and public health officials understand the health of children and youth and how health varies across these populations. Objective measures of CRF can be obtained in lab-based settings by collecting expired gases during a graded exercise test. Limitations of this approach include the high cost, time intensiveness, laboratory infrastructure requirements, unfamiliar equipment, and unnatural environment yielding heavy participant burden, making these tests not suitable for mass testing. To help address these limitations, the 20 m shuttle run test (20mSRT)⁵ was developed as a field-based measure to estimate CRF.

* Corresponding author.

E-mail address: jlang033@uottawa.ca (J.J. Lang).

<http://dx.doi.org/10.1016/j.jsams.2017.08.002>

1440-2440/© 2017 Sports Medicine Australia. Published by Elsevier Ltd. All rights reserved.

Please cite this article in press as: Lang JJ, et al. Systematic review of the relationship between 20 m shuttle run performance and health indicators among children and youth. *J Sci Med Sport* (2017), <http://dx.doi.org/10.1016/j.jsams.2017.08.002>

The 20mSRT is a progressive CRF test where children and youth are asked to run laps back and forth between two parallel lines 20 m apart. An audio recording paces the participants beginning at a speed of 8.5 km/h and increasing by 0.5 km/h every consecutive minute. Participants continue until they are no longer able to keep pace with the audio recording for two consecutive laps, at which point their score is recorded (i.e., total laps, stages, time, etc.).⁵ The 20mSRT is the most widely used field-based assessment of CRF, with recent research indicating that at least 50 countries from six continents have used the assessment.^{6,7} This popularity is likely due to the low cost of equipment, simplicity in administering and scoring the test, flexibility in testing (e.g., surface, space, location, etc.), low participant burden and ability to simultaneously test large groups of children and youth.⁸ These characteristics allow the 20mSRT to overcome many cultural and language barriers that limit other fitness assessments. Moreover, the 20mSRT demonstrates strong test-retest reliability⁸ and moderate-to-strong validity⁹ which suggests that the 20mSRT is a good measure of CRF. Although recent systematic reviews have explored the associations between physical fitness, physical activity and health among children and youth,^{1,10,11} no study has comprehensively assessed the associations between 20mSRT performance and health indicators among children and youth.

Thus, this systematic review aimed to summarize and evaluate available research that assessed the association between 20mSRT performance and health indicators (e.g., physiological, psychosocial, cognitive) among children and youth aged 5–17 years (or in Grades 1–12).

2. Methods

The systematic review protocol is registered with the International Prospective Register of Systematic Reviews (PROSPERO; registration no. CRD42015032250; available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015032250). This study was conducted and reported in accordance with the PRISMA 2009 statement for reporting systematic reviews.¹²

The participants, intervention, comparison, outcome and study designs (PICOS) framework¹³ was used to help delineate the study parameters of our research question, and to develop the search strategy.

School-aged participants who were between the mean ages of 5–17 years and/or between grades 1–12, if in the North America (or equivalent grades in other countries), were included in this study.

Performance results for the 20mSRT protocol beginning at 8.5 km/h and increasing by 0.5 km/h every minute was our main exposure. Results reported as raw scores (e.g., speed, laps, stages, etc.) and/or predicted relative $\dot{V}O_{2peak}$ were included.

In certain cases 20mSRT performance can be dichotomized into healthy and unhealthy CRF using criterion-referenced standards. In these cases comparisons for health indicators would take place between the different CRF groups. Otherwise, the relationship between 20mSRT performance and continuous health indicators were assessed according to comparisons or relationships provided by the authors.

Health indicators for this study were selected based on a recent physical activity systematic review,¹¹ with additional health indicators being added based on the literature (i.e., previous research). Eleven health indicators were included that broadly represented three categories: physiological, psychosocial and cognitive health. Health indicators included: (i) adiposity (e.g., body mass index, waist circumference), (ii) cardiometabolic biomarkers (e.g., lipid profile, insulin resistance, blood pressure), (iii) cognition/academic achievement (e.g., concentration, memory, school grades), (iv) mental health (e.g., depression, anxiety, risk of mental illness), (v)

psychosocial health (e.g., behaviour problems, self-efficacy, social development), (vi) self-esteem and physical self-perception (e.g., perceived competence), (vii) quality of life and wellbeing, (viii) bone health (e.g., bone mineral density, bone area), (ix) musculoskeletal fitness (e.g., muscular strength, muscular endurance, flexibility), (x) motor skill development (e.g., gross/fine motor skills, coordination), and (xi) injuries and/or harm (e.g., adverse events).

No restrictions were placed on study design. A minimum sample size for all study designs was set at 40 participants. Studies were screened if they were written in English only. Peer-reviewed original research articles that were published or in-press were eligible for inclusion. All grey literature, abstracts, reviews, theses and conference proceedings were excluded from this review. One additional variant of the 20mSRT protocol was also accepted (i.e., the Eurofit 20mSRT protocol begins at 8.0 km/h, increases to 9.0 km/h for the second minute, and then increases by 0.5 km/h every minute thereafter¹⁴). No restrictions were placed on the included health indicators (e.g., validity and reliability criteria).

A systematic search of the literature was conducted in MEDLINE (OVID interface; 1980 to March 2nd, 2016), EMBASE (OVID interface; 1980 to March 3rd, 2016), PsychINFO (OVID interface; 1980 to March 4th, 2016), Cochrane Central Register of Controlled Trials (CENTRAL, OVID interface; 1980 to March 4th, 2016) and SPORT-Discus (EBSCO interface; 1980 to March 2nd, 2016). The search strategy was developed by a librarian experienced in systematic review searching and peer-reviewed by another librarian using the PRESS standard.¹⁵ Search strategies are presented in Supplement 1.

Database specific search strategies were executed by a single researcher through the OVID and EBSCOhost interfaces. All bibliographic records were extracted as text files and imported into Reference Manager Software (version 11; Thompson Reuters, San Francisco, CA, USA) where duplicate files were manually removed. Titles and abstracts for the remaining relevant articles were imported into DistillerSR (Evidence Partners; Ottawa, ON, Canada). All potentially relevant titles and abstracts were screened against the inclusion criteria by two reviewers. Exclusion by both reviewers was required to exclude records at level 1. The remaining records passed to level 2 where full-text articles were screened against the inclusion criteria independently by two reviewers. Consensus by both reviewers was required at level 2 to include records. All discrepancies were resolved between reviewers or by a third reviewer, if necessary. Furthermore, the reference lists of included articles and relevant reviews in the area were checked for additional studies not identified through the systematic search.

A study specific database was created in Excel (Microsoft Corp.; Redmond, Wash., USA) and piloted by reviewers. Data were extracted from included studies into the database by one reviewer, and subsequently checked for accuracy by a second reviewer.

Study specific descriptive characteristics were extracted (i.e., author, study year, study country, study design, sample size, participant age and sex), 20mSRT type (i.e., result units, equation used if predicted $\dot{V}O_{2peak}$ was calculated, CRF standard, etc.), health indicator (including measurement method and units), results (i.e., correlation coefficient, odds ratio, p-value, chi-square values, 95% CIs) and confounding variables (included in multivariate analyses). When studies presented multiple models, only the results from the fully adjusted model were extracted. Statistically significant findings were set at $p < 0.05$. Furthermore, correlations < 0.1 were considered as trivial and correlations of 0.1, 0.3 and 0.5 were used as thresholds for weak, moderate and strong, respectively.

The risk of bias across all included studies was evaluated using the methods described in the Cochrane Handbook.¹⁶ A modified version of the Cochrane Handbook was used to assess risk of bias in observational studies. Each study was assessed for selec-

tion bias (e.g., failure to use an appropriate sampling technique), performance bias (e.g., inappropriate measurement of exposure) and selective reporting bias (e.g., selective/incomplete reporting of results); detection (e.g., inappropriate measurement of outcomes) and attrition (e.g., amount and treatment of incomplete data) bias for each health indicator; and other sources of bias (e.g., inadequate control for important confounders).

The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework¹⁷ was used to assess the quality of evidence for each health indicator category across each study design. The GRADE framework categorizes the quality of evidence into 4 groups (high, moderate, low or very low). According to GRADE, quality of evidence is high for randomized controlled trials (RCTs) and low for all other study designs. Quality of evidence is downgraded if there is a serious risk of bias, imprecision (e.g., insufficiently precise findings to support a recommendation),¹⁸ inconsistency (e.g., differences in the direction of the effects not explained by methodological differences),¹⁹ indirectness (e.g., whether studies included the population, assessment, or outcomes of interest),²⁰ or other factors. If there is no justification for downgrading, quality of evidence can be upgraded if there is a large effect size, a dose-response gradient, or if all plausible confounders would decrease an apparent treatment effect.²¹ Overall, quality of evidence was assessed independently by two reviewers and subsequently verified by the larger co-author team.

Meta-analyses were planned for sufficiently homogeneous health indicators with regard to statistical and methodological characteristics. Narrative syntheses of results were planned for health indicators that were considered to be too heterogeneous.

3. Results

A total of 11,745 records were identified through database searches. After de-duplication, 2,447 records remained. After screening titles and abstracts, 470 full-text articles were obtained for further review. A total of 328 studies were excluded for the reasons described in Fig. 1. Full consensus was attained between both reviewers for all included studies.

The remaining 142 studies met all inclusion criteria and represented 319,311 participants (301,105 from unique samples) from 32 different countries. Study-specific sample sizes ranged from 51 [S138] to 141,169 [S35] participants, and spanned a mean age from 5.1 to 17.9 years. Included studies employed either a longitudinal cohort ($n = 14$) or cross-sectional ($n = 128$) design. Individual study characteristics are summarized in Supplement 2, and Tables 1–11.

In total, 89 studies (63% of included studies) investigated the relationship between 20mSRT performance and measures of adiposity (Table 1, Supplement Table 1). Six studies used a longitudinal cohort study design, 82 studies used a cross-sectional study design, and the remaining study used a mixed longitudinal cohort and cross-sectional design. Five of seven longitudinal cohort studies reported that better 20mSRT performance at baseline was associated with lower adiposity levels at 1–9 years of follow-up [S1–S3, S6, S7], while two longitudinal cohort studies reported mixed favourable and null associations for girls [S4] and boys [S5]. The quality of the evidence for longitudinal cohort studies remained rated as low as there were no reasons to downgrade or upgrade the quality of the evidence. A total of 66 cross-sectional studies (out of 83) reported favourable relationships between 20mSRT performance and adiposity, 15 studies reported mixed favourable and null findings, and two studies reported mixed favourable and unfavourable findings (Table 1) [S16, S41]. Of the 89 included studies, 41 studies assessed adiposity using only BMI, 25 of the studies used various measures, including waist circumference, percent body fat, and sum of skinfolds, and the remaining 23 studies

used a combination of BMI and other measures to assess adiposity. The quality of the evidence was upgraded from low to moderate because of a large effect and evidence of a dose-response relationship.

A total of 42 studies (30% of included studies) investigated the relationship between 20mSRT performance and various cardiometabolic biomarkers (e.g., HOMA-IR, cholesterol, C-reactive protein) in children and youth (Table 2, Supplement Table 2). One study used a longitudinal cohort design, 39 studies used a cross-sectional design, and the remaining two studies used mixed longitudinal cohort and cross-sectional designs. One longitudinal cohort study reported mixed null and favourable associations between 20mSRT performance and cardiometabolic biomarkers (HOMA-IR) at 2 years of follow-up [S91], and the remaining two longitudinal cohort studies reported null findings at follow-up [S90, S92]. The quality of the evidence for longitudinal cohort studies remained rated as low. Of the 41 cross-sectional studies, 13 reported favourable findings, 21 reported mixed favourable and null findings, 6 reported null findings, and one study [S97] reported a mixed null and unfavourable findings between 20mSRT performance and cardiometabolic biomarkers for boys only. The quality of the evidence for cross-sectional studies was also rated as low.

Fifteen studies (11% of included studies) investigated the associations between 20mSRT performance and cognition/academic achievement (Table 3, Supplement Table 3). Only one study used a longitudinal cohort design, and the remaining 14 studies used a cross-sectional design. The longitudinal cohort study reported that 20mSRT performance was favourably associated with attention, but null associations were reported for spatial working memory at 9 months of follow-up [S114]. The quality of the evidence for the longitudinal cohort study was downgraded from low to very low due to serious imprecision (insufficient evidence). Three cross-sectional studies assessed the relationship between 20mSRT performance and cognition, and reported favourable or mixed favourable and null findings [S9, S33, S74]. The remaining 11 cross-sectional studies investigated the association between 20mSRT performance and academic achievement. Six studies reported favourable associations [S9, S42, S116, S119–S121], four studies reported mixed favourable and null findings [S115, S117, S118, S122], one study reported null findings [S37], and one study reported mixed unfavourable and null findings [S123]. The quality of the evidence for cross-sectional studies was downgraded from low to very low due to a serious risk of bias. Specifically, nine studies had unknown reliability and validity for the outcome measures, and four studies did not control for sex or age differences in their analyses.

Two cross-sectional studies (1% of included studies) investigated the association between 20mSRT performance and mental health (Table 4, Supplement Table 4). One study reported no association between 20mSRT performance and depression in children [S124]. Another study reported a favourable association between 20mSRT performance and levels of depression in girls, but reported no association in boys [S125]. The quality of the evidence remained rated as low.

Five cross-sectional studies (4% of included studies) investigated the associations between 20mSRT performance and psychosocial health (Table 5, Supplement Table 5). Four of five studies reported favourable associations between 20mSRT performance and self-efficacy [S126–S128], enjoyment [S13], and outcome expectancy [S128]. One study reported favourable associations between 20mSRT performance and mood/emotions in girls, but the associations were null in boys [S129]. The majority of studies did not adequately control for differences in sex or age within their analyses [S13, S126–S128], and one study did not describe the sampling methods employed [S129], resulting in a risk of bias. As a result, the quality of the evidence was downgraded from low to very low.

Table 1
Association between 20mSRT performance and adiposity.

No. of studies	Design	Quality assessment				No. of participants	Absolute effect	Quality
		Risk of bias						
		Inconsistency	Indirectness	Imprecision	Other			
The range of mean ages was 5.1–17.4 years; 37 studies reported an age range in lieu of the mean, and in these cases the included ages ranged from 5 to 18 years. Data were collected cross-sectionally and up to 9 years of follow-up. Adiposity was assessed as body mass index (absolute, z-score or percentile), waist circumference, waist-to-hip ratio, waist-to-height ratio, hip circumference, % body fat (bioelectrical impedance, dual-energy X-ray absorptiometry, skinfolds) or Marshal Visual Rating Scale								
7	Longitudinal ^a	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	18,396	BMI (5 studies) 5 of 7 studies reported favourable associations between 20mSRT performance and BMI at 1 to 6.1 years of follow-up 2 studies also reported null findings between 20mSRT performance and BMI for girls but not boys at 1 year [S5] and boys but not girls at 18 months [S4] of follow-up Other measures (2 studies) 2 studies reported favourable associations between 20mSRT performance and SSF [S2] and WC [S3] at 5–9 years of follow-up BMI only (37 studies)	Low ^b
83	Cross-sectional ^c	No serious risk of bias	No serious inconsistency	No serious indirectness	Evidence of a large dose-response effect and a gradient. Example, better 20mSRT performance associated with lower adiposity ^d	220,971		Moderate ^e
32 studies reported favourable relationships between 20mSRT performance and BMI 3 studies reported mixed favourable and null associations between 20mSRT performance and BMI , 2 studies reported null associations between 20mSRT performance and BMI Other measures (23 studies) 20 studies reported favourable relationships between 20mSRT and other measures of adiposity 2 studies reported mixed favourable and null associations between 20mSRT performance and other measures of adiposity 1 study reported null associations between 20mSRT performance and %BF BMI and other measures (23 studies) 14 studies reported favourable associations between 20mSRT performance and various measures of adiposity 7 studies reported mixed favourable and null associations between 20mSRT and various measures of adiposity 2 studies reported mixed favourable and unfavourable associations between 20mSRT performance and adiposity Due to heterogeneity in the measurement of 20mSRT performance and adiposity a meta-analysis was not possible								

^aNote 20mSRT, 20 m shuttle run test; BMI, body mass index; SSF, sum of skinfolds; WC, waist circumference; %BF, percent body fat.

^b Includes 7 longitudinal studies [S1–S7].

^c The quality of the evidence remained as “low” as there were no serious concerns about the quality of included longitudinal studies or reasons to increase the rating.

^d Includes 83 cross-sectional studies [S5, S8–S89].

^e The general favourable gradient and large effect of the findings (e.g., approximately 60% of reported unstandardized correlation coefficients were greater than $r = 0.30$) suggests that better 20mSRT performance is associated with lower adiposity.

^f The quality of the evidence was upgraded from “low” to “moderate” because of the large effect observed and the evidence of a dose-response gradient between 20mSRT performance and adiposity.

Table 2
Association between 20mSRT performance and cardiometabolic biomarkers.

No. of studies	Design	Quality assessment					No. of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other			
The range of mean ages was 9–17.9 years; 13 studies reported an age range in lieu of the mean, and in these cases the included ages ranged from 7 to 18 years. Data were collected cross-sectionally and up to 4 years of follow-up. Lipid, Glucose/Insulin and inflammatory indicators were measured via fasting blood samples. Blood pressure was measured either manually or with an automated sphygmomanometer. Cardiometabolic risk score was calculated according to published methods using varying combinations of pertinent health indicators ^a									
3	Longitudinal ^b	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	1,115	Glucose, insulin, HOMA-IR 1 study reported mixed favourable and null associations between 20mSRT performance and HOMA-IR at 2 years follow-up [S91] 2 studies reported no association between 20mSRT performance and HOMA-IR at 9 months to 4 years follow-up [S90,S92] Inflammatory markers 1 study reported null findings between 20mSRT performance and CRP at 9 months follow-up [S90]	Low ^c
41	Cross-sectional ^d	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	52,102	Blood pressure (16 studies) 3 studies reported favourable associations between 20mSRT performance and BP . 4 studies reported mixed favourable and null associations between 20mSRT performance and BP . 9 studies reported null findings between 20mSRT performance and BP Blood lipids (TC, HDL, LDL; 13 studies) 4 studies reported favourable associations between 20mSRT performance and blood lipids 7 studies reported mixed favourable and null associations between 20mSRT performance and blood lipids 2 studies reported null findings between 20mSRT performance and blood lipids Glucose, insulin, HOMA-IR (15 studies) 7 studies reported favourable associations between 20mSRT performance and fasting glucose , fasting insulin , and/or HOMA-IR levels 6 studies reported mixed favourable and null associations between 20mSRT performance and fasting glucose , fasting insulin , and/or HOMA-IR levels 2 studies reported null findings between 20mSRT performance and fasting glucose , fasting insulin , and/or HOMA-IR levels Inflammatory Markers (CRP, homocysteine, Interleukin-6; 13 studies) 5 studies reported favourable associations between 20mSRT performance and inflammatory markers 5 studies reported mixed favourable and null associations between 20mSRT performance and inflammatory markers 3 studies reported null associations between 20mSRT performance and inflammatory markers Inflammatory markers Cardiometabolic risk (14 studies) 10 studies reported favourable associations between 20mSRT performance and cardiometabolic risk 3 studies reported mixed favourable and null associations between 20mSRT performance and cardiometabolic risk 1 study reported null findings between 20mSRT performance and cardiometabolic risk Due to heterogeneity in the measurement of 20mSRT performance and cardiometabolic biomarkers a meta-analysis was not possible	Low ^e

^aNote 20mSRT, 20 m shuttle run test; BP, blood pressure; CRP, c-reactive protein; HDL, high density lipoprotein; HOMA-IR, homeostatic model assessment – insulin resistance; LDL, low density lipoprotein; MAP, mean arterial pressure; TC, total cholesterol; TG, triglycerides.

^b Different combinations of the following health indicators were used to calculate a cardiometabolic risk score: waist circumference, BMI z-score, HDL, LDL, glucose, insulin, HOMA-IR, MAP, triglycerides, blood pressure, total cholesterol:HDL ratio, c-reactive protein.

^c Includes 3 studies [S90–S92].

^d The quality of the evidence remained as “low” as there were no serious concerns about the quality of included longitudinal studies or reasons to increase the rating.

^e The quality of the evidence remained as “low” as there were no serious concerns about the quality of included cross-sectional studies or reasons to increase the rating.

Table 3
Association between 20mSRT performance and cognition/academic achievement.

No. of studies	Design	Quality assessment				No. of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other		
The range of mean ages was 5.2 and 17.3 years; 5 studies reported an age range in lieu of the mean, and in these cases the included ages ranged from 8 to 17 years. Data were collected cross-sectionally and up to 9 months of follow-up. Cognition was measured using computer testing modalities and other tests or questionnaires: FittSmart, Intelligence and Development Scales, KHV-VK test, Flanker task, spatial n-back task, Stroop Colour and Word test, and Kaufman Brief Intelligence Test. Academic achievement was measured using student grade averages across various subjects, performance on standardized tests, WESTEST, Weschsler Individual Achievement Test, or the HSBC questionnaire								
1	Longitudinal ^a	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious imprecision ^b	None	Niederer et al. found that 20mSRT performance at baseline was favourably associated with attention ($\beta = 0.25$, 95%CI 0.03–0.48, $p = 0.03$) but not associated with spatial working memory ($\beta = 0.08$, 95%CI –0.09 to 0.24, $p = 0.37$) at 9 months follow-up [S114]. Cognition (3 studies) Williams et al. found favourable associations between 20mSRT performance and health knowledge ($\beta = 0.56$, $p < 0.001$) [S9].	Very low ^c
14	Cross-sectional ^d	Serious risk of bias ^e	No serious inconsistency	No serious indirectness	No serious imprecision	None	Scudder et al. found that children's 20mSRT performance was favourably associated with working memory , reaction time and accuracy but not incongruent accuracy during tasks of inhibitory control [S33]. Buck et al. found that 20mSRT performance was favourably related to word, colour, and colour/word concentration tasks , but not interference concentration tasks. Performance on the 20mSRT was not related to IQ in children [S74]. Academic achievement (11 studies) 5 studies found favourable associations between 20mSRT performance and measures of overall academic achievement , English language arts , math , science , social studies , reading or spelling courses [S42, S116, S119–S121]. 4 studies reported mixed favourable associations between 20mSRT performance and numeracy , writing and spelling , but no association with reading skills [S115, S117, S118, S122]. 1 study reported null associations between 20mSRT performance and math , English and social study grades [S37]. 1 study found mixed null associations between 20mSRT performance and academic achievement in children, but unfavourable associations in adolescents [S123]. Due to heterogeneity in the measurement of 20mSRT performance and cognition a meta-analysis was not possible	Very low ^f

^aNote: 20mSRT, 20 m shuttle run test; β indicates standardized beta; HSBC, Health Behaviour in School-Aged Children.

^b Includes 1 longitudinal study [S114].

^c Serious imprecision. Only 1 study investigated the longitudinal associations between 20mSRT performance and cognition.

^d The quality of the evidence was downgraded from "low" to "very low" because of serious imprecision.

^e Includes 14 cross-sectional studies [S9, S33, S37, S42, S74, S115–S123].

^f Serious risk of bias. A total of 10 out of 14 studies did not provide detail on the validity and reliability of outcome measures [S9, S37, S42, S74, S116].

[†] The quality of the evidence was downgraded from "low" to "very low" because of a serious risk of bias that diminished the level of confidence in the observed effects.

Table 4
Association between 20mSRT performance and mental health.

No. of studies	Design	Quality assessment				No. of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other		
The range of mean ages was 12.4–12.5 years. Data were collected cross-sectionally. Depression was measured using the CES-DC								
2	Cross-sectional ^a	No serious risk of bias	No serious inconsistency	No serious indirectness	No serious imprecision	None	2,008	Rieck et al. found no association between 20mSRT performance and depression in girls (OR = 2.09, 95%CI 0.96–4.57) and boys (OR = 1.63, 95%CI 0.08–3.17) [S124] Greenleaf et al. reported a favourable association between 20mSRT performance and depression in girls ($F = 8.64$, $r = -0.54$, $p = 0.01$), but not boys ($F = 5.14$, $r = -0.25$) [S125] Due to heterogeneity in the measurement of 20mSRT performance and mental health a meta-analysis was not possible

*Note 20mSRT, 20 m shuttle run test; CES-DC, Center for Epidemiology Studies–Depression Scale for Children.

^a Includes 2 studies [S124, S125].

^b The quality of the evidence remained as “low” as there were no serious concerns about the quality of included cross-sectional studies or reasons to increase the rating.

Table 5
Association between 20mSRT performance and psychosocial health.

No. of studies	Design	Quality assessment				No. of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision			
The range of mean ages was 9.5–13.3 years. Data were collected cross-sectionally. Psychosocial health was measured using self-report questionnaires: KIDSCREEN-52 and various adapted questionnaires based on Bandura's guide for constructing self-efficacy scales								
5	Cross-sectional ^a	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	No serious imprecision	1,949	4 of 5 studies reported favourable relationships between 20mSRT performance and either self-efficacy, enjoyment, or outcome expectancy [S13, S126–S128] 1 study reported favourable associations between 20mSRT performance and moods/emotions for girls ($\beta=0.14$, $p=0.009$), but no association for boys ($\beta=0.05$, $p=0.43$) [S129] Due to heterogeneity in the measurement of 20mSRT performance and psychosocial health a meta-analysis was not possible	Very low ^c

*Note 20mSRT, 20 m shuttle run test.

^a Includes 5 studies [S13, S126–S129].

^b Serious risk of bias. 4 out of 5 studies did not control for sex and/or age in the analyses [S13, S126–S128].

^c The quality of the evidence was downgraded from “low” to “very low” because of a serious risk of bias that diminished the level of confidence in the observed effects.

Table 6
Association between 20mSRT performance and self-esteem and physical self-perception.

No. of studies	Design	Quality assessment					No. of participants	Absolute effect	Quality
		Risk of bias							
		Random risk of bias ^a	Inconsistency	Indirectness	Imprecision	Other			
The range of mean ages was 8.2–13.7 years; 3 studies reported an age range in lieu of the mean, and in these cases the included ages ranged from 8 to 17 years. Data were collected cross-sectionally and up to 4 years of follow-up. Self-esteem and/or physical self-perception was measured using an adapted version of the Body Self-Esteem Scale for children, Perceived Competence in Physical Education Scale, Behavioural Regulation in Exercise-2 questionnaire, KIDSCREEN-52 questionnaire, Self-Description Questionnaire II, CSAPPA questionnaire, and the Intrinsic Motivation Inventory questionnaire									
1	Longitudinal ^b	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious imprecision ^c	None	821	20mSRT performance at baseline was not associated with body dissatisfaction at 4 years follow-up in girls ($\beta = -4.96$, $p = 0.21$) and boys ($\beta = -4.67$, $p = 0.21$) [S130]	Very low ^d
9	Cross-sectional ^b	Serious risk of bias ^d	No serious inconsistency	No serious indirectness	No serious imprecision	None	6,973	Favourable associations Colquitt et al. found 20mSRT performance was favourably associated with perceived usefulness ($r = 0.35$, $p < 0.01$; $\beta = 0.27$, $p = 0.005$) [S13]	Very low ^e
Gao et al. found 20mSRT performance was favourably associated with perceived sport competence ($\beta = 0.27$, $p < 0.01$) [S131] Risser et al. found 20mSRT performance was favourably associated with self-determined motivation ($\beta = 2.08$, $p < 0.01$) [S133] Greenleaf et al. found boys ($F = 10.2$, $r = 0.35$, $p < 0.01$) and girls ($F = 6.87$, $r = 0.48$, $p = 0.01$) in the Healthy Fitness Zone (HFZ) for 20mSRT performance reported better self-perceived strength than those not in the HFZ. Additionally, boys ($F = 80.2$, $r = 0.97$, $p < 0.001$) and girls ($F = 26.1$, $r = 0.93$, $p = 0.0001$) in the HFZ for 20mSRT performance reported better self-perceived endurance and better body satisfaction (Boys: $F = 8.3$, $r = 0.31$, $p = 0.01$; Girls: $F = 8.4$, $r = 0.53$, $p = 0.01$) than those not in the HFZ [S125] Cairney et al. found 20mSRT performance was favourably associated with perceived adequacy ($\beta = 0.08$, $SE = 0.19$, $p < 0.001$) and predilection ($\beta = 0.03$, $SE = 0.1$, $p < 0.05$) [S73] Shen et al. found 20mSRT performance favourably associated with perceived competence in physical education ($r = 0.21$, $p < 0.01$) [S134] Jurimae et al. found 20mSRT performance was favourably associated with perceived endurance (r range = -0.31 – 0.51 , $p < 0.05$) [S132] Olive et al. found 20mSRT performance was favourably associated with body dissatisfaction in boys ($\beta = -9.87$, $SE = 2.6$, $p < 0.001$) and girls ($\beta = -11.0$, $SE = 3.5$, $p = 0.002$) [S130] Mixed findings Morales et al. found 20mSRT performance was favourably associated with self-perception in girls ($\beta = 0.13$, $p = 0.01$) but not boys ($\beta = 0.04$, $p = 0.5$) [S129] Greenleaf et al. found girls in the HFZ for 20mSRT performance reported better self-esteem ($F = 8.46$, $p = 0.01$) compared to those not in the HFZ. No differences were observed in boys ($F = 5.24$, $p = \text{not significant}$) [S125] Null findings Morales et al. found 20mSRT performance was not associated with autonomy in girls ($\beta = 0.05$, $p = 0.42$) and boys ($\beta = 0.00$, $p = 0.42$) [S129] Due to heterogeneity in the measurement of 20mSRT performance and self-esteem and physical self-perception a meta-analysis was not possible									

Note* 20mSRT, 20 m shuttle run test; HFZ, Healthy Fitness Zone; NIZ, Needs Improvement Zone.

^a Includes 1 longitudinal study [S130].

^b Serious risk of bias. 1 study scale used to measure Body Dissatisfaction was adapted by the authors and had an equal number of positive and negative items, but did not mention changing the Likert scale of responses (ranged from "very much like I feel" to "not at all like I feel") to match the item's direction [S130].

^c Serious imprecision. Only one longitudinal study was published examining the association between 20mSRT performance and physical self-perception.

^d The quality of evidence was downgraded from "low" to "very low" because of serious imprecision and a serious risk of bias that diminished the level of confidence in the observed effects.

^e Includes 9 cross-sectional studies [S13, S73, S125, S129–S134].

^f Serious risk of bias. 5 of 9 studies reported using convenience sampling [S13, S130, S131, S133, S134], and it was unclear whether 3 of the remaining 4 studies used convenience sampling methods [S73, S125, S132]. Additionally, 3 studies did not control for age or sex differences in 20mSRT performance [S13, S131, S134].

^g The quality of evidence was downgraded from "low" to "very low" because of a serious risk of bias that diminished the level of confidence in the observed effects.

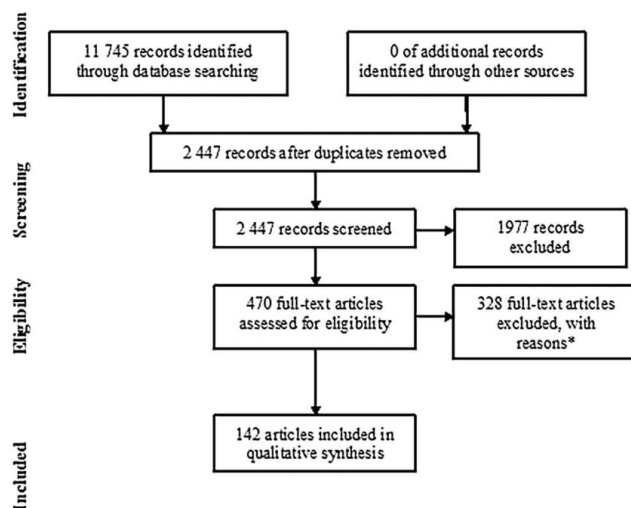


Fig. 1. PRISMA flowchart for the identification, screening, eligibility and inclusion of articles. *Reasons for exclusion: study did not include the proper 20mSRT protocol ($n=55$); study did not assess the relationship between 20mSRT performance and a relevant health indicator ($n=137$); sample size ($n=2$); study participants were not in the target age range ($n=2$); grey literature and/or not original research (i.e., review, conference proceeding, abstract; $n=82$); populations with known disease or injury ($n=8$); could not obtain full-text article ($n=12$); non-English language article ($n=26$); and duplicate paper discovered during level 2 screening ($n=4$). Some articles were excluded for more than one reason.

Nine studies (6% of included studies) investigated the associations between 20mSRT performance and indicators of self-esteem and physical self-perception (Table 6, Supplement Table 6). One study used a mixed longitudinal cohort and cross-sectional design, and the remaining eight studies used a cross-sectional design. The longitudinal cohort study reported a null association between 20mSRT performance and body dissatisfaction after 4-years of follow-up [S130]. However, because this study used an unvalidated scale to measure body dissatisfaction, there was a serious risk of bias. There was also serious imprecision (insufficient evidence). As a result, the quality of the evidence from the longitudinal cohort study was downgraded from low to very low. Six cross-sectional studies reported favourable associations between 20mSRT performance and various indicators of self-esteem and physical self-perception [S13, S128, S130, S132–S134]. Three studies reported mixed favourable and null associations between 20mSRT performance and various indicators of self-esteem or physical self-perception [S73, S125, S129]. The quality of the evidence was downgraded from low to very low as a result of a serious risk of bias in sampling methods and not controlling for age and sex differences in some studies [S13, S128, S134].

Six cross-sectional studies (4% of included studies) investigated the relationship between 20mSRT performance and indicators of quality of life and wellbeing (Table 7, Supplement Table 7). Three studies reported favourable associations between 20mSRT performance and various indicators of quality of life and wellbeing [S123, S133, S135], two studies reported mixed favourable and null associations [S123, S129], and one study reported mixed null and unfavourable findings [S10]. One study used a questionnaire that was not validated for the age groups in which it was used, thereby creating a serious risk of bias [S123]. There was also serious imprecision (i.e., wide confidence intervals) in the estimates reported by two studies [S123, S133]. As a result, the quality of the evidence was downgraded from low to very low.

One cross-sectional study (1% of included studies) investigated the relationship between 20mSRT performance and bone health

(Table 8, Supplement Table 8) [S137]. This study found that 20mSRT performance was a favourably associated with bone mineral content in boys and girls. However, the study did not include lean body mass as a covariate in the analysis although it accounted for 67% of the variance in bone mineral content [S137]; this resulted in a serious risk of bias. In addition, the lack of studies in this area created a serious risk of imprecision. As a result, the quality of the evidence was downgraded from low to very low.

Nine cross-sectional studies (6% of included studies) investigated the relationship between 20mSRT performance and musculoskeletal fitness (Table 9, Supplement Table 9). The musculoskeletal fitness outcomes were divided into four categories: strength and power, muscular endurance, flexibility, and agility. Of the seven studies that investigated the association between 20mSRT performance and indicators of strength and power, four studies reported favourable findings [S13, S84, S108, S138], two studies reported mixed favourable and null findings [S12, S46], and one study reported null findings [S62]. All six studies assessing the relationship between 20mSRT performance and measures of muscular endurance reported favourable associations [S12, S13, S37, S42, S46, S62]. All four studies reported null associations between 20mSRT performance and flexibility [S13, S37, S42, S62]. Lastly, of the two studies that investigated the associations between 20mSRT performance and measures of agility, one reported favourable associations [S12] while the other reported null associations [S62]. The quality of the evidence was downgraded from low to very low as a result of a serious risk of bias. More specifically, the majority of studies used convenience sampling methods or did not describe their sampling procedures, and most did not describe the validity and reliability of their measurement instruments.

Four cross-sectional studies (3% of included studies) investigated the association between 20mSRT performance and motor skill development (Table 10, Supplement Table 10). Two of the four studies reported favourable associations between 20mSRT performance and motor skill development [S18, S139], and the remaining two studies reported null associations [S62, S140]. The quality of

Table 7
Association between 20mSRT performance and quality of life and wellbeing.

No. of studies	Design	Quality assessment					No. of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other			
6	Cross-sectional ^a	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious imprecision ^c	None	3,508	Favourable associations Riser et al. found 20mSRT performance was favourably associated with HRQoL ($\beta = 4.16$, 95%CI 0.3–8.02, $p < 0.05$) [S133] Castro-Pinero et al. found 20mSRT performance was favourably associated with health complaints in children ($\beta = -0.15$, $p = 0.005$) and adolescents ($\beta = -0.18$, $p = 0.002$). Additionally, children (OR = 2.56, 95%CI: 1.3–5.0) and adolescents (OR = 1.99, 95%CI: 1.2–3.4) with poor performance on the 20mSRT had increased odds of reporting health complaints [S135] Mota et al. found unfit girls had higher odds of reporting negative self-rated health (OR = 2.93, 95%CI: 1.8–4.7, $p < 0.001$) [S136] Padilla-Moledo et al. found 20mSRT performance was positively associated with life satisfaction in children ($\beta = 0.14$, $p = 0.006$) and adolescents ($\beta = 0.17$, $p = 0.002$) [S123] Mixed findings Morales et al. found 20mSRT performance was favourably associated with physical ($\beta = 0.16$, $p = 0.001$) and psychological ($\beta = 0.11$, $p = 0.03$) wellbeing in girls, but no associations were observed in boys [S129] Padilla-Moledo et al. found 20mSRT performance was favourably associated with perceived health in adolescents ($\beta = 0.15$, $p = 0.008$), but no association was found in children ($\beta = 0.08$, $p = 0.11$) [S123] Padilla-Moledo et al. found adolescents with a poor performance on the 20mSRT had increased odds of reporting their health as 'fair' instead of 'excellent' (OR = 2.76, 95%CI 1.1–7.2) and of reporting low life satisfaction (OR = 1.66, 95%CI 1.0–2.7). No increased odds were observed in children [S123] Null findings Borras et al. found 20mSRT performance was not associated to satisfaction with health, emotional comfort or restricted activity [S10] Unfavourable findings Borras et al. found 20mSRT performance was unfavourably associated to physical comfort ($r = -0.13$, $p = 0.03$) [S10] Due to heterogeneity in the measurement of 20mSRT performance and quality of life and wellbeing, a meta-analysis was not possible	Very low ^d

^aNote 20mSRT, 20m shuttle run test; HRQoL, Health Related Quality of Life.

^b Includes 6 studies [S10, S123, S129, S133, S135, S136].

^c Serious risk of bias. 2 of 6 studies employed convenience sampling methods and reported only favourable findings related to 20mSRT performance [S133, S136]. 1 study used a questionnaire that was not validated for the age groups in their study [S123].

^d Serious imprecision. 2 studies reported wide confidence intervals [S123, S133].

^e The quality of evidence from cross-sectional studies was downgraded from 'low' to 'very low' because of serious risk of bias and serious imprecision that reduced the level of confidence in the observed effects.

Table 8
Association between 20mSRT performance and bone health.

No. of studies	Design	Quality assessment				No. of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other		
1	One study included age from 13 to 18.5 years. Data were collected cross-sectionally. Cross-sectional ^a	No serious risk of bias ^b Serious risk of bias ^c	No serious inconsistency	Bone mineral content was measured objectively with dual-energy X-ray absorptiometry	Serious imprecision ^d	None	Vincent-Rodriguez et al. found that performance on the 20mSRT was a favourably associated with bone mineral content (Boys: B = 0.05, SE = 0.24, p < 0.0001; Girls: B = 0.18, p < 0.0001) [S137]	Very low ^d

^aNote 20mSRT, 20 m shuttle run test; B indicates unstandardized beta; SE, standard error.

^bIncludes 1 cross-sectional study [S137].

^cSerious risk of bias. Lean body mass accounted for 67% of the total variance in whole-body bone mineral content, independent of age and sex, yet lean body mass was not included in the analysis as a covariate [S137].

^dSerious imprecision. Only one study investigated the relationship between 20mSRT performance and bone health.

The quality of the evidence was downgraded from "low" to "very low" because of a serious risk of bias and serious imprecision that diminished the level of confidence in the observed effects.

the evidence was downgraded from low to very low as a result of a serious risk of bias and serious imprecision. More specifically, three studies used convenience sampling methods, and one study did not control for sex differences in their analysis [S18, S62, S139], subsequently resulting in a serious risk of bias. Serious imprecision was recorded because one study reported wide confidence intervals [S139].

Two cross-sectional studies (1% of included studies) investigated the relationship between 20mSRT performance and the risk of injury/harm (Table 11, Supplement Table 11). One study reported a favourable association where children with better 20mSRT performance had lower odds of sustaining serious injury [S141]. Another study reported opposing findings where better 20mSRT performance was associated with higher odds of sustaining a bone fracture [S142]. The quality of the evidence was downgraded from low to very low due to this serious inconsistency.

A high-level summary of findings is reported in Table 12. Due to heterogeneity across all health indicators a meta-analysis of results was not possible.

4. Discussion

This systematic review combined results from 142 studies that investigated the associations between 20mSRT performance and various health indicators among children and youth aged 5–17 years. The included studies represented 319,311 children and youth (301,105 from unique samples) from 32 countries across 6 continents, which comprised of approximately 39% of the world population.²² The quality of the evidence ranged from very low to moderate. Overall, the evidence suggests that better 20mSRT performance is associated with a variety of positive health indicators that range from lower adiposity levels to favourable cardiometabolic biomarker profiles and physical self-perception.

Some believe 20mSRT performance is strongly influenced by adiposity, suggesting that poor 20mSRT performers are on average the heavier children and youth.²³ However, our results indicate that the unadjusted correlation coefficients ranged widely from $r = -0.70$ (strong) [S47] to $r = -0.10$ (weak) [S28], with some studies having reported positive coefficients as high as $r = 0.13$ (weak unfavourable) [S83]. These varying results suggest that adiposity alone does not explain 20mSRT performance. Similarly, Olds and Dollman reported that adiposity explained 40–60% of the longitudinal declines seen in 1.6 km running speed in 10- to 12-year-old children between 1985 and 1997.²⁴ In fact, it is likely that a combination of various physiological and psychological aspects help explain 20mSRT performance, including mechanical efficiency,²⁵ sustainable fractional of $\dot{V}O_{2\max}$,²⁶ $\dot{V}O_2$ kinetics,²⁷ lactate threshold,²⁷ anaerobic capacity,²⁸ and motivation.⁸

In some cases 20mSRT performance was strongly associated with one subcategory, but not others, within similar health indicator categories. For instance, 20mSRT performance displayed favourable associations with some subcategories of the cardiometabolic biomarker category: among the most favourable were HOMA-IR (effect size: $r = 0.02$ [S101] to $r = -0.60$ [S26]; negligible to strong), C-reactive protein (effect size: $r = 0.01$ [S36] to $r = -0.15$ [S63]; negligible to weak) and aggregate indices of cardiometabolic disease risk (effect size: $r = 0.03$ [S101] to $r = -0.31$ [S52]; negligible to moderate). In addition, other health indicator subcategories also displayed consistent favourable associations: academic achievement (effect size: $r = -0.06$ [S117] to $r = 0.27$ [S121]; negligible to weak), muscular endurance (effect size: $r = 0.08$ [S37] to $r = 0.70$ [S13]; negligible to very strong), psychosocial health (effect size: $r = 0.23$ [S131] to $r = 0.71$ [S126]; weak to very strong) and physical self-perception (effect size: $r = 0.21$ [S134] to $r = 0.51$ [S132]; weak to strong); whereas, some health indicator subcategories dis-

Table 9
Association between 20mSRT performance and musculoskeletal fitness in children and youth.

No. of studies	Design	Quality assessment					No of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other			
The range of mean ages was 5.4 and 15.3 years; 3 studies reported an age range in lieu of the mean, and in these cases the included ages ranged from 7 to 18 years. Data were collected cross-sectionally. Measures of musculoskeletal fitness were obtained using standing broad jump (SBJ), sit-and-reach (SAR), sit-and-reach (SAR) test, hand grip strength (HGS), Bent arm hang, 10 × 5 m shuttle, trunk lift, max repetition curl-ups and push-ups, and Bruininks–Oseretsky Test of Motor Proficiency									
9	Cross-sectional ^a	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	No serious imprecision	None	3,784	Strength and power (7 studies) 2 studies reported favourable associations between 20mSRT performance and SBJ [S12, S46]	Very low ^c
1 study reported null associations between 20mSRT performance and SBJ ($r = 0.32$, 95%CI 0.05–0.55) [S62]									
3 studies reported null associations between 20mSRT performance and HGS [S12, S46, S62]									
4 studies reported favourable associations between 20mSRT performance and other measures of strength and power [S13, S84, S108, S138]									
Muscular endurance (6 studies)									
6 studies reported favourable associations between 20mSRT performance and curl-up and [S12, S13, S37, S42, S46, S62] push-ups [S13, S37, S42]									
Flexibility (4 studies)									
4 studies reported null associations between 20mSRT performance and the SAR [S13, S37, S42, S62]									
Agility (2 studies)									
1 study reported favourable associations between 20mSRT performance and agility ($F = 8.32$, $p < 0.001$) [S12]									
1 study reported a null association between 20mSRT performance and agility ($r = -0.10$, 95%CI -0.37 to 0.18) [S62]									
Due to heterogeneity in the measurement of 20mSRT performance and musculoskeletal fitness a meta-analysis was not possible									

Note* 20mSRT, 20 m shuttle run test; HGS, hand-grip strength; SAR, sit and reach; SBJ, standing broad jump.

^a Includes 9 studies [S12, S13, S37, S42, S46, S62, S84, S108, S138].

^b Serious risk of bias. 7 of 9 studies did not describe their sampling methods [S12, S13, S37, S42, S84, S107, S138]. 7 of 9 studies did not sufficiently describe the validity and reliability of their measurement procedures [S12, S13, S37, S42, S84, S108, S138].

^c The quality of the evidence was downgraded from “low” to “very low” because of a serious risk of bias that diminished the level of confidence in the observed effects.

Table 10
Association between 20mSRT performance and motor skill development.

No. of studies	Design	Quality assessment			No. of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other	
The range of mean ages was 8.4–8.9 years; 2 studies reported an age range in lieu of the mean, and in these cases the included ages ranged from 5 to 16 years. Data were collected cross-sectionally and up to 7 years of follow-up. Motor skill development was measured by the TGMD-2 and an observational behavioural instrument							
4	Cross-sectional ^a	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious imprecision ^c	None	Very low ^d 2 of 4 studies reported favourable associations between 20mSRT performance and motor skill development [S18, S139] 2 of 4 studies found no association between 20mSRT performance and motor skill development [S62, S140] Due to heterogeneity in the measurement of 20mSRT performance and motor skill development a meta-analysis was not possible

Note^a 20mSRT, 20 m shuttle run test; FMSMCT, Fundamental Motor Skills Manual for Classroom Teachers; TGMD-2, Test of Gross Motor Development-Second Edition.

^a Includes 4 cross-sectional studies [S18, S62, S139, S140].

^b Serious risk of bias. 3 studies reported using convenience sampling [S18, S62, S139], and 1 study did not control for sex differences in 20mSRT performance [S139].

^c Serious imprecision. 1 study reported an OR = 70.3 (95%CI 2.17–269.3). These wide confidence intervals would make a difference in the interpretation of the effect [S139].

^d The quality of the evidence was downgraded from “low” to “very low” because of a serious risk of bias and a serious imprecision that diminished the level of confidence in the observed effects.

Table 11
Association between 20mSRT performance and risk of injury/harm.

No. of studies	Design	Quality assessment			No. of participants	Absolute effect	Quality
		Risk of bias	Inconsistency	Indirectness	Imprecision	Other	
The range of mean ages was 7.9 to 13.9 years. Data were collected cross-sectionally. Measures of risk of injury were collected using self-reported parental questionnaires							
2	Cross-sectional ^a	No serious risk of bias	Serious inconsistency ^b	No serious indirectness	No serious imprecision	None	Very low ^c 1 study reported favourable associations between high 20mSRT performance and lower odds of sustaining serious injury (OR 0.29, 95%CI 0.16–0.63) [S141] 1 study reported unfavourable associations between high 20mSRT performance and higher odds of sustaining bone fractures (OR 1.50, 95%CI 1.08–1.60) [S142] Due to heterogeneity in the measurement of 20mSRT performance and risk of injury a meta-analysis was not possible

Note^a 20mSRT, 20 m shuttle run test.

^a Includes 2 studies [S141, S142].

^b Serious inconsistency. Results across studies conflict and these conflicting results are unexplainable, resulting in serious inconsistency.

^c The quality of the evidence was downgraded from “low” to “very low” because of serious inconsistency across studies.

Table 12
High-level summary of findings by health indicator.

Health indicator	No. of studies	Quality of the evidence	Summary of findings
Adiposity	89	Low to moderate	71 studies reported favourable findings between 20mSRT performance and adiposity 16 studies reported mixed favourable and/or null findings 2 studies reported mixed favourable and unfavourable findings
Cardiometabolic biomarkers	42	Low	13 studies reported favourable findings between 20mSRT performance and cardiometabolic biomarkers 28 studies reported mixed favourable and/or null findings 1 study reported mixed unfavourable and null findings
Cognition/academic achievement	15	Very low	6 studies reported favourable findings between 20mSRT performance and cognition/academic achievement 8 studies reported mixed favourable and/or null findings 1 study reported mixed unfavourable and null findings
Mental health	2	Low	1 study reported mixed favourable and null findings between 20mSRT performance and mental health 1 study reported null findings
Psychosocial health	5	Very low	4 studies reported favourable findings between 20mSRT performance and psychosocial health 1 study reported mixed favourable and null findings
Self-esteem/physical self-perception	9	Very low	5 studies reported favourable findings between 20mSRT performance and self-esteem/physical self-perception 4 studies reported mixed favourable and null findings
Quality of life/wellbeing	6	Very low	3 studies reported favourable findings between 20mSRT performance and quality of life/wellbeing 2 studies reported mixed favourable and null findings 1 study reported mixed unfavourable and null findings
Bone health	1	Very low	1 study reported favourable findings between 20mSRT performance and health
Musculoskeletal fitness	9	Very low	3 studies reported favourable findings between 20mSRT performance and musculoskeletal health 6 studies reported mixed favourable and null findings
Motor skill development	4	Very low	2 studies reported favourable findings between 20mSRT performance and motor skill development 2 studies reported mixed favourable and null findings
Risk of injury/harm	2	Very low	1 study reported favourable findings between 20mSRT performance and harms/risk of injury 1 study reported unfavourable findings

played consistent null associations: mean arterial pressure (effect size: $r=0.065$ [S3] to $r=-0.156$ [S11]; negligible to weak) and flexibility (effect size: $r=-0.01$ [S37] to $r=0.20$ [S62]; negligible to weak). The variability across different subcategories within each health indicator was the main reason for mixed results and inconsistencies. However, it is apparent that very few studies reported unfavourable associations between 20mSRT performance and health indicators suggesting that there is an overall gradient between better performance and better holistic health among children and youth.

This study represents the largest effort to date to synthesize the associations between 20mSRT performance and health indicators in children and youth. The results from this study brought together findings from 142 studies using a rigorous systematic review process. Previous research has suggested that the 20mSRT would be a good measure to compare the CRF and health of children and youth across countries and continents.^{6,7} Indeed, the findings from our review support the 20mSRT as an appropriate population health indicator for children and youth.

A common issue with systematic reviews is that the synthesis of evidence is only as good as the papers that they contain.³⁰ Thus, the primary limitation of this review is that the quality of the evidence for all health indicator categories was rated as very low to moderate as a result of pooling findings that were always observational in nature. Nonetheless, the relative uniformity of findings with very few unfavourable relationships, despite the flaws inherent in several studies, should not prevent drawing confident conclusions about the strong and favourable relationships identified herein. Although the nature of the 20mSRT does not allow for randomized controlled trial (RCT) designs (the 20mSRT is a single measurement of a trait), the results from this study encourage the need for RCTs that investigate the sensitivity of 20mSRT performance improve-

ments in children and youth and their relationship with changes in health indicators. In addition, a meta-analysis was not appropriate due to the heterogeneity between studies, which was mainly identified in the reporting of 20mSRT results as the exposure measure. For example, results were recorded as laps, stages, time, speed, distance, or predicted $\dot{V}O_{2peak}$, which in itself has many variations in the different equations used (e.g., Léger et al.³¹ Ruiz et al.³² Fitnessgram³³). There is certainly a need for consistent reporting of 20mSRT results across studies, as recommended by Tomkinson et al.⁷ Furthermore, we were not able to separate our results by sex or age groups because of the heterogeneity across studies, although study specific details regarding sex and age are available in Supplement 2. Lastly, we did not include studies written in languages other than English which may have introduced a bias.

This review also highlights many gaps in the literature. There is a need to further investigate the associations between 20mSRT performance and indicators of bone health, mental health, psychosocial health, and quality of life/wellbeing. These are areas of research that are important as mental health disorders continue to grow and result in healthcare challenges worldwide.³⁴ Furthermore, there is a need for more global and coordinated efforts to measure 20mSRT performance to better understand the health of children and youth across developed and developing countries. Evidence suggests that 20mSRT performance varies substantially worldwide⁶; thus, global measurement efforts could provide a proactive method to identify populations of youth at risk of poor health outcomes and future chronic diseases.³⁵

5. Conclusion

Our study provides a systematic review of the associations between 20mSRT performance and health indicators among

school-aged children and youth. The results demonstrate the broad utility of the 20mSRT as an indicator of not only physiological and physical health, but also psychosocial and cognitive health in children and youth. As a result, the 20mSRT would be an excellent tool to screen children and youth at risk of poor health and in need of increasing physical fitness levels. These results also support the notion that the 20mSRT could be used as an indicator to compare the holistic health of children and youth through temporal trends and between geographic locations (e.g., cities, provinces, states and countries).

Contributorship

Justin Lang, Kevin Belanger, Mark Tremblay and Grant Tomkinson developed the research questions and objectives. Justin Lang and Kevin Belanger contributed equally in screening articles, extracting data and producing result tables. Justin Lang drafted the manuscript. All authors contributed to interpreting results, editing, reviewing and approved the final manuscript.

Funding statement

This study received no financial support.

Ethical approval statement

Not applicable.

Data sharing statement

Not applicable.

Acknowledgements

We thank Margaret Sampson, MLIS, PhD, AHIP (Children's Hospital of Eastern Ontario) for developing the electronic search strategies, and Lise Dubas, MLIS student, University of Alberta, for peer review of the search strategies.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jsams.2017.08.002>.

References

1. Armstrong N, Tomkinson GR, Ekelund U. Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during youth. *Brit J Sports Med* 2011; 45:849–858.
2. Ortega FB, Ruiz JR, Castillo MJ et al. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes* 2008; 32(1):1–11.
3. Höglström G, Nordström A, Nordström P. High aerobic fitness in late adolescence is associated with a reduced risk of myocardial infarction later in life: a nationwide cohort study in men. *Eur Heart J* 2014; 35:3133–3140.
4. Höglström G, Nordström A, Nordström P. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. *Int J Epidemiol* 2015. <http://dx.doi.org/10.1093/ije/dyv321>. Published Online First: 20 December.
5. Léger L, Lambert J, Goulet A et al. Capacité aérobie des Québécois de 6 à 17 ans-Test navette de 20 mètres avec paliers de 1 minute Aerobic capacity of 6 to 17 year old Québécois – 20 metre shuttle run test with 1 minute stages. *Can J Appl Sport Sci* 1984; 9(2):64–69.
6. Lang JJ, Tremblay MS, Léger L et al. International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country com-

- parison? A systematic literature review with pooling of aggregate results. *Sports Med* 2016. Published Online First: 20 September.
7. 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. *Brit J Sports Med* 2016. <http://dx.doi.org/10.1136/bjsports-2016-095987>. Published Online First: May 20th.
8. Tomkinson GR, Olds TS. Field tests of fitness, in *Paediatric exercise science and medicine*, Armstrong N, van Mechelen W, editors, New York, NY, Oxford University Press, 2000, p. 109–128.
9. Mayorga-Vega M, Aguilar-Soto P, Viciano J. Criterion-related validity of the 20-M shuttle run test for estimating cardiorespiratory fitness: a meta-analysis. *J Sports Sci Med* 2015; 14:536–547.
10. Janssen I, LeBlanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int J Behav Nutr Phys Act* 2010; 7:40.
11. Poitras VJ, Gray CE, Borghese MM et al. Systematic review of the relationship between objectively measured physical activity and health indicators in school-aged children and youth. *Appl Physiol Nutr Metab* 2016; 41:S197–S239.
12. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; 6(6):e1000097.
13. Schardt C, Adams MB, Owens T, Keits S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inf Decis Mak* 2007; 7:16.
14. Council of Europe. Eurofit: handbook for the Eurofit tests of physical fitness, Rome, Council of Europe, 1988.
15. McGowan J, Sampson M, Lefebvre C. An evidence based checklist for the peer review of electronic search strategies (PRESS EBC). *Evid Based Libr Inf Pract* 2010; 5(1):149–154.
16. Higgins J, Green S. *Cochrane handbook for systematic reviews of interventions version 5.1.0*, The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
17. Guyatt GH, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J et al. GRADE guidelines: 1. Introduction–GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; 64(4):383–394.
18. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D et al. GRADE guidelines: 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol* 2011; 64(12):1283–1293.
19. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol* 2011; 64(12):1294–1302.
20. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M et al. GRADE guidelines: 8. Rating the quality of the evidence—indirectness. *J Clin Epidemiol* 2011; 64(12):1303–1310.
21. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P et al. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 2011; 64(12):1311–1316.
22. United Nations, Department of Economic and Social Affairs, Population Division. *World population prospects: the 2015 revision, key findings and advance tables*. 2015; Working Paper No. ESA/P/WP.241.
23. Rowland T. Declining cardiorespiratory fitness in youth: fact or supposition? *Pediatr Exerc Sci* 2002; 14:1–8.
24. Olds T, Dollman J. Are changes in distance-run performance of Australian children between 1985 and 1997 explained by changes in fatness? *Pediatr Exerc Sci* 2004; 16:201–209.
25. Lussier L, Buskirk ER. Effects of an endurance training regimen of assessment of work capacity in prepubertal children. *Ann N Y Acad Sci* 1977; 301:734–747.
26. Krahenbuhl GS, Pangrazi RP, Chomokos EA. Aerobic responses of young boys to submaximal running. *Res Q* 1979; 50:413–421.
27. Péronnet F, Thibault G. Mathematical analysis of running performance and world records. *J Appl Physiol* 1989; 67:453–465.
28. Sjödén B. The relationship among running economy, aerobic power, and onset of blood lactate accumulation in young boys (11–15 years), in *Exercise and sport biology*, Komi PV, editor, Champaign, IL, Human Kinetics, 1982, p. 57–60.
29. Weir A, Rabia S, Ardern C. Trusting systematic reviews and meta-analyses: all that glitters is not gold! *Brit J Sports Med* 2016. Published online first: 15 October.
30. 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.
31. Ruiz JR, Ramirez-Lechuga J, Ortega F et al. Artificial neural network-based equation for estimating VO2max for the 20 m shuttle run test in adolescents. *Artif Intell Med* 2008; 44:233–245.
32. Meredith MD, Welk GJ. *FITNESSGRAM/ACTIVITYGRAM test administration manual (updated 4th ed)*, Meredith MD, Welk GJ, editors, Champaign, IL, Human Kinetics, 2010.
33. Whiteford HA, Degenhardt L, Rehm J et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 2013; 382(9904):1575–1586.
34. Ruiz JR, Castro-Pinero J, Artero EG et al. Predictive validity of health-related fitness in youth: a systematic review. *Brit J Sports Med* 2009; 43(12):909–923.

Please cite this article in press as: Lang JJ, et al. Systematic review of the relationship between 20 m shuttle run performance and health indicators among children and youth. *J Sci Med Sport* (2017), <http://dx.doi.org/10.1016/j.jsams.2017.08.002>

CHAPTER 4: MANUSCRIPT 2

This manuscript has been written and formatted following the publication standards of the peer-reviewed journal, *British Journal of Sport Medicine*. It appears in its published format.

Title:

Review of criterion-referenced standards for cardiorespiratory fitness: what percentage of 1 142 026 international children and youth are apparently healthy?

Citation:

Lang JJ, Tremblay MS, Ortega FB, Ruiz JR, Tomkinson GR. 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*. Published Online First: 2 March 2017. doi:10.1136/bjsports-2016-096955

Online link:

<http://bjsm.bmj.com/content/early/2017/03/02/bjsports-2016-096955>

Review of criterion-referenced standards for cardiorespiratory fitness: what percentage of 1 142 026 international children and youth are apparently healthy?

Justin J Lang,¹ Mark S Tremblay,¹ Francisco B Ortega,^{2,3} Jonatan R Ruiz,^{2,3} Grant R Tomkinson^{4,5}

¹Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

²PROMoting FITness and Health through physical activity research group (PROFITH), Department of Physical Education and Sports, Faculty of Sport Sciences, University of Granada, Granada, Spain

³Karolinska Institutet, Huddinge, Sweden

⁴Department of Kinesiology and Public Health Education, University of North Dakota, Grand Forks, North Dakota, USA

⁵Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health Sciences & Sansom Institute for Health Research, University of South Australia, Adelaide, South Australia, Australia

Correspondence to

Dr Justin J Lang, Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Rd, Ottawa, ON, K1H 8L1, Canada; jlang@cheo.on.ca

Accepted 2 January 2017

ABSTRACT

Purpose To identify criterion-referenced standards for cardiorespiratory fitness (CRF); to estimate the percentage of children and youth that met each standard; and to discuss strategies to help improve the utility of criterion-referenced standards for population health research.

Methods A search of four databases was undertaken to identify papers that reported criterion-referenced CRF standards for children and youth generated using the receiver operating characteristic curve technique. A pseudo-dataset representing the 20-m shuttle run test performance of 1 142 026 children and youth aged 9–17 years from 50 countries was generated using Monte Carlo simulation. Pseudo-data were used to estimate the international percentage of children and youth that met published criterion-referenced standards for CRF.

Results Ten studies reported criterion-referenced standards for healthy CRF in children and youth. The mean percentage ($\pm 95\%$ CI) of children and youth that met the standards varied substantially across age groups from $36\% \pm 13\%$ to $95\% \pm 4\%$ among girls, and from $51\% \pm 7\%$ to $96\% \pm 16\%$ among boys. There was an age gradient across all criterion-referenced standards where younger children were more likely to meet the standards compared with older children, regardless of sex. Within age groups, mean percentages were more precise (smaller CI) for younger girls and older boys.

Conclusion There are several CRF criterion-referenced standards for children and youth producing widely varying results. This study encourages using the interim international criterion-referenced standards of 35 and 42 mL/kg/min for girls and boys, respectively, to identify children and youth at risk of poor health—raising a clinical red flag.

INTRODUCTION

Cardiorespiratory fitness (CRF) is often measured as peak oxygen uptake ($\dot{V}O_{2peak}$) in children and youth, which is described as the highest rate of consumed oxygen during a bout of exercise.¹ CRF in children and youth is a good determinant of sporting success¹ and is meaningfully associated with health.² In children and youth, CRF is moderately to strongly associated with cardiometabolic,^{3–4} mental⁵ and physical⁶ health markers. Being a trait, CRF levels in adolescence track moderately well into adulthood.⁷ Furthermore, recent cohort studies of over 700 000 Swedish adolescents with a median follow-up of 29–34 years have identified that individuals with

low CRF in late adolescence have an increased risk of myocardial infarction⁸ and early mortality⁹ in adulthood. Together, these findings suggest that the measurement and surveillance of CRF in childhood and adolescence can help describe the current and future health of populations.

The 20-m shuttle run test (20mSRT), originally published in 1984 by Léger and colleagues,¹⁰ is the most widely reported field-based assessment of CRF with published normative data available on over 1.1 million children from 50 countries.^{11–12} This test displays moderate-to-high criterion related validity¹³ and high reliability,¹⁴ suggesting that the 20mSRT is a good field-based measure of $\dot{V}O_{2peak}$. It is also considered an excellent population health surveillance tool because of its simplicity, low cost, flexibility (eg, indoor/outdoor), minimal equipment and ability to test large groups of children and youth simultaneously.¹⁴ As a result, the 20mSRT is included as part of several national and international physical fitness test batteries, such as the Eurofit,¹⁵ FITNESSGRAM,¹⁶ Canadian Assessment of Physical Literacy,¹⁷ Australian Council for Health, Physical Education and Recreation,¹⁸ Assessing Levels of Physical Activity and Fitness,¹⁹ and Assessing FITness in PRESchoolers.^{20–22}

There are two distinct approaches to interpret 20mSRT performance results. Individual results can be compared with a normative-referenced standard—described as a *relative* comparison with the performance of peers, and expressed as a percentile rank. Normative-referenced standards for the 20mSRT exist for international²³ and various national populations.^{24–29} Alternatively, results can be converted to predicted $\dot{V}O_{2peak}$ using one of several equations²³ to compare with a criterion-referenced standard—described as an *absolute* measure of the minimum performance required to 'achieve' the criterion-referenced (eg, healthy metabolic or cardiovascular profile) standard. These data are often expressed as the percentage (prevalence) of the sample meeting the criterion-referenced standard. Several of these standards have recently emerged applying a receiver operating characteristic (ROC) curve analysis—a diagnostic evaluation that identifies the optimal standards to distinguish a health outcome in a population.³⁰ The recent emergence of multiple criterion-referenced standards may potentially create confusion regarding the healthy CRF status of children and youth.



CrossMark

To cite: Lang JJ, Tremblay MS, Ortega FB, et al. *Br J Sports Med* Published Online First: [please include Day Month Year]. doi:10.1136/bjsports-2016-096955



Copyright Article author (or their employer) 2017. Produced by BMJ Publishing Group Ltd under licence.

Original article

Furthermore, international percentages of children and youth meeting these criterion-referenced standards do not exist. These referenced percentages would provide utility for international surveillance, screening and CRF comparisons across populations.

Therefore, the purpose of this study was threefold: (a) to identify all published ROC curve-generated criterion-referenced standards for CRF in children and youth; (b) to estimate the percentage of children and youth that met each criterion-referenced standard using an international sample of 20mSRT performance scores from 1 142 026 children and youth aged 9–17 years; and (c) to discuss strategies to improve the utility of CRF criterion-referenced standards for population health research.

METHODS

Search strategy

A literature search was conducted to identify all studies that reported health-related criterion-referenced standards for CRF in children and youth. A single author (JL) searched MEDLINE, PsycINFO, EMBASE and SPORTDiscus (1980–2016) using a combination of three separate search strategies: (a) child* OR adoles* OR paediatric* OR schoolchild; (b) cardiorespiratory* OR cardiovascular* OR aerobic* OR VO₂* OR criterion*; and (c) cardiovascular OR metabolic OR insulin OR blood pressure OR cardiometabolic OR psychosocial OR well-being OR self-efficacy. All search strategies were combined using the Boolean operator 'AND.' No restrictions were placed on language. Furthermore, the reference lists of all relevant studies were reviewed for additional studies not included in the original search.

Inclusion criteria

One author (JL) assessed titles and abstracts for those likely to report criterion-referenced standards for children and youth. Studies were included if they reported: (a) CRF standards (VO_{2peak}/max, mL/kg/min) associated with cardiovascular, metabolic and/or psychosocial health markers; (b) standards generated using ROC curve statistics; and (c) standards generated for children and youth. No restrictions were placed on study design (eg, cohort, cross-sectional).

Data extraction

One author (JL) independently reviewed all the included studies and extracted the following information: authors, year of publication, country, health criterion, CRF assessment, VO_{2peak} prediction equation, sex, age (years), sample size, criterion-referenced standards (mL/kg/min), area under the ROC curve and 95% CI. A second author (GT) reviewed the data for transcription errors.

International sample and percentage analysis

To achieve the second objective of this study, international normative values for the 20mSRT performance were obtained from Tomkinson and colleagues (table 1).²³ In brief, the international norms were developed using a systematic review protocol to identify 177 studies that reported 20mSRT performance scores for apparently healthy (eg, free from known disease) children and youth aged between 9 and 17 years. A total of 1 142 026 20mSRT performances were extracted and standardised to running speed (km/hour) at the last completed stage using the data treatment procedure reported by Tomkinson *et al.*²³ Sample-weighted means and SD at the sex by age by country level were corrected for oversampling and/or undersampling using a poststratification population-weighted procedure.³¹ The

Table 1 International 20mSRT norms for predicted VO_{2peak} (mL/kg/min) obtained by converting Tomkinson *et al.*'s²³ normative centile values using the Léger equation³⁴

Girls				Boys			
Age (years)	n	50th percentile	SD	n	50th percentile	SD	
9	98 166	46.7	3.2	101 532	48.1	4.0	
10	81 415	45.1	3.7	81 538	46.9	4.3	
11	78 029	43.5	4.1	79 660	45.8	5.0	
12	69 497	42.0	4.4	70 266	45.2	5.7	
13	62 491	40.4	4.7	72 913	45.0	6.2	
14	55 035	38.8	5.0	65 317	44.6	6.8	
15	43 730	37.2	5.3	53 287	44.0	7.3	
16	34 978	35.6	5.4	35 490	43.2	7.7	
17	29 559	34.1	5.6	29 123	42.6	8.2	
Total	552 900			589 126			

20mSRT, 20-m shuttle run test.

corrected means and SD at the sex by age level were imported into LMSchartmaker³² to generate the international 20mSRT normative-referenced values using the Lambda Mu Sigma (LMS) method.

International normative running speed values were used to randomly generate pseudo-data points for each sex by age group using Monte Carlo simulation, a technique described in detail elsewhere.^{12–33} This technique assumes that data are normally distributed, which is true for 20mSRT running speed at the last completed stage. Pseudo-data were repeatedly generated until the calculated mean differed from the reported mean by less than 0.5%, and the calculated SD differed from the reported SD by less than 2.5%. These international running speed pseudo-data were then converted to VO_{2peak} using the Léger prediction equation:

$$\dot{V}O_{2peak} \text{ (mL/kg/min)} = 31.025 + 3.238 \cdot X_1 - 3.248 \cdot X_2 + 0.1536 \cdot X_1 \cdot X_2$$

where X_1 is the running speed at the last completed stage (km/hour) and X_2 is the age at the last birthday.³⁴

The Léger equation was applied to help compare outcomes across different criterion-referenced standards. This equation was developed using the 20mSRT, a wide range of age groups and two inputs that were readily available and similar to the pseudo-data used in this study.³⁴ In addition, the Léger equation is the most widely used equation to predict VO_{2peak} in children and youth.

The international 20mSRT pseudo-dataset was then applied to each reported CRF criterion-referenced standard:

$$\text{Healthy CRF (\%)} = \frac{\text{number of participants that attained the criterion - referenced standard}}{\text{total number of participants}}$$

In addition, these sex by age by study percentages were used to calculate the overall smoothed population-weighted mean percentages and 95% CIs using the poststratification population-weighted procedure described by Levy and Lemeshow.³¹ These percentages represent the gross estimate of the international percentage of children and youth with apparently healthy CRF.

RESULTS

Criterion-referenced standards

Ten studies reported CRF criterion-referenced standards for children and youth between 8 and 19 years of age. These studies were published over a 10-year period from 2006 to 2016 (table 2).

Table 2 Descriptive characteristics of included studies

Study	Year	Country	Criterion	CRF test	Sex	Age (years)	Sample size (n)	Healthy CRF (mL/kg/min)	AUC (95% CI)
Mesa <i>et al</i> ³⁹	2006	Spain	TC, HDLc, LDLc, Lp(a)	20-m shuttle run test with estimated $\dot{V}O_{2peak}$	Boys	13–18.5	248	45–47	0.72 (0.65 to 0.78)
Ruiz <i>et al</i> ⁴¹	2007	Estonia and Sweden	Insulin, glucose, HDLc, TG, sum of five skinfolds, BP	Maximal cycle ergometer test with estimated $\dot{V}O_{2peak}$	Girls	9–10	444	37	0.68 (0.62 to 0.73)
					Boys	9–10	429	42	0.67 (0.61 to 0.73)
Lobelo <i>et al</i> ⁴³	2009	USA	SBP, triceps and subscapular skinfolds, HOMA, TG, TC/HDLc	Submaximal treadmill walking test with estimated $\dot{V}O_{2peak}$	Girls	12–15	315	36	0.54 (0.42 to 0.66)
					Boys	16–19	255	36	0.54 (0.38 to 0.70)
						12–15	308	44	0.77 (0.68 to 0.86)
						16–19	369	40	0.72 (0.64 to 0.79)
Adegboye <i>et al</i> ⁴⁵	2011	Denmark, Estonia, Portugal and Norway	SBP, TG, TC/HDLc, HOMA, sum of four skinfolds	Maximal cycle ergometer test with direct measurement of $\dot{V}O_{2peak}$	Girls	8–11	1 181	37.4	0.67
					Boys	14–17	1 051	33	0.69
						8–11	1 219	43.6	0.68
						14–17	1 049	46	0.55
Moreira <i>et al</i> ⁴⁰	2011	Portugal	TG, HDLc, LDLc, waist to hip ratio, TC, glucose, BP	20-m shuttle run test with estimated $\dot{V}O_{2peak}$	Girls	10–18	255	39.5	0.65 (0.59 to 0.71)
					Boys	10–18	195	41.8	0.65 (0.58 to 0.72)
Welk <i>et al</i> ⁴⁶	2011	USA	Waist circumference, SBP/DBP, HDLc, TG, glucose	Submaximal treadmill walking test with estimated $\dot{V}O_{2peak}$	Girls	12–18	591	38–40	0.77
					Boys	12–18	649	40–44	0.83
Boddy <i>et al</i> ⁴⁴	2012	England	BMI, waist circumference, %BF	20-m shuttle run test with estimated $\dot{V}O_{2peak}$	Girls	9–10	8 237	41.9	0.58 to 0.77
					Boys	9–10	8 382	46.6	0.74 to 0.80
Silva <i>et al</i> ⁴⁵	2012	Portugal	Waist circumference, HDLc, TG, glucose, MAP	20-m shuttle run test with estimated $\dot{V}O_{2peak}$	Girls	10–18	2 789	32.6–43.9	NA
					Boys	10–18	2 770	43.1–46.3	
Ruiz <i>et al</i> ⁴⁷	2014	Greece, Germany, Belgium, France, Hungary, Italy, Sweden, Austria and Spain	Smoking, BMI, physical activity, diet, TC, BP, glucose	20-m shuttle run test with estimated $\dot{V}O_{2peak}$	Girls	12.5–17.5	259	34.6	0.69 (0.61 to 0.78)
					Boys	12.5–17.5	251	43.8	0.81 (0.73 to 0.88)
Silva <i>et al</i> ⁴⁰	2016	Brazil	BP	Modified Canadian Aerobic Fitness Test with estimated $\dot{V}O_{2max}$	Girls	14–19	464	32	0.70 (0.65 to 0.74)
					Boys	14–19	411	40	0.63 (0.58 to 0.67)

AUC, area under the curve; BMI, body mass index; BP, blood pressure; CRF, cardiorespiratory fitness; DBP, diastolic blood pressure; HDLc, high-density lipoprotein cholesterol; HOMA, homeostasis model assessment; Lp(a), lipoprotein(a); LDLc, low-density lipoprotein cholesterol; MAP, mean arterial pressure; NA, not applicable; SBP, systolic blood pressure; TC, total cholesterol; TG, triglycerides; %BF, per cent body fat.

The criterion-referenced standards for girls and boys ranged from 32.6 to 43.9 mL/kg/min and from 40.0 to 47.0 mL/kg/min, respectively. Standards were developed using European, North American or South American samples of children and youth. The health criteria were largely based on summative markers of metabolic and/or cardiovascular health (eg, triglycerides, cholesterol, glucose, body composition, etc). In general, each study estimated $\dot{V}O_{2peak}$ using a combination of various CRF assessment protocols, exercise modes (eg, walk/run field tests, graded treadmill test, cycle ergometry) and $\dot{V}O_{2peak}$ prediction equations. Only one study included a direct measure of $\dot{V}O_{2peak}$ for a portion of the study participants.⁴⁵ Five studies assessed CRF using the 20mSRT, representing half of the included studies.

Figure 1 illustrates the variability of reported criterion-referenced standards for each sex by age by study group across the international 20mSRT normative values reported by Tomkinson and colleagues.²³ For both girls and boys, this figure follows a similar pattern where the criterion-referenced standards consistently fall below the international 20mSRT 50th percentile for younger children, indicating that the standards are attained at lower 20mSRT performance percentile ranks. However, for both sexes, there are age gradients where older children must attain a higher 20mSRT performance percentile rank to attain the criterion-referenced standards when compared with younger children. For girls, the variability across age-matched

criterion-referenced standards spanned ~25 to 50 percentile points using Tomkinson and colleagues' normative-referenced standards.²³ The variability across age-matched standards was slightly smaller for boys, spanning ~20 to 45 percentile points.

Percentage estimates for 'healthy' CRF

Table 3 displays 105 sex by age by study percentages (%) of children and youth that achieved each reported criterion-referenced standard. These represent the study-specific international percentage estimates of apparently 'healthy' children and youth calculated using the international 20mSRT norms. In general, younger children were more likely to meet the standards than older children, and mean percentages were more precise (smaller CI) for younger girls and older boys. More specifically, the percentage pattern across all studies was such that girls aged 9–11 years were apparently healthier than age-matched boys, girls and boys aged between 12 and 15 years were similarly healthy, and older boys aged between 16 and 17 years were apparently healthier than girls of the same age. This pattern was true for all studies, except for the Welk (FITNESSGRAM) criterion-referenced standards,⁴⁶ where boys were considered to be healthier than girls at every age group.

In addition, the smoothed population-weighted mean percentage ($\pm 95\%$ CI) of girls and boys meeting the

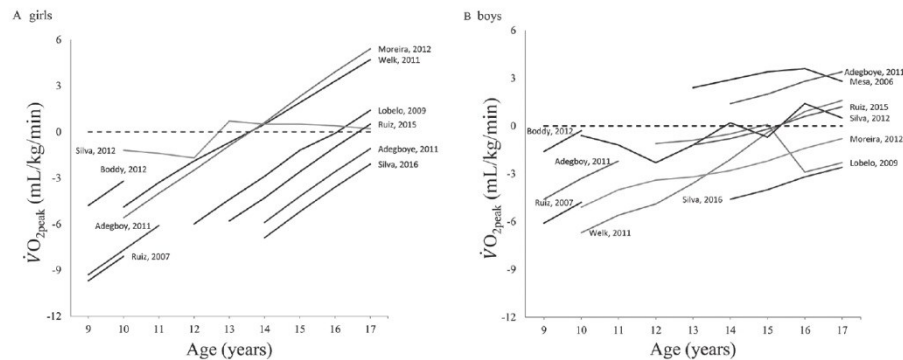


Figure 1 Study of criterion-referenced standards minus international normative 20-m shuttle run test (20mSRT) performance as a function of age for girls (A) and boys (B). The dash line, centred at 0, represents the international 20mSRT 50th percentile (predicted $\dot{V}O_{2peak}$) for 9- to 17-year-old children and youth.²³ Lines below the dash line indicate that the criterion-referenced standards are below the international 50th percentile of 20mSRT performance, and lines above represent standards above the 50th percentile for each age group.

criterion-referenced CRF standards ranged from $36\% \pm 13\%$ to $95\% \pm 4\%$ and from $51\% \pm 7\%$ to $96\% \pm 16\%$, respectively (figure 2). These smoothed population-weighted mean percentages represent our best estimate of the international mean percentage of children and youth aged 9–17 years old with apparently healthy CRF.

DISCUSSION

The main objectives of this paper were to identify all criterion-referenced standards for CRF generated using ROC curve

technique; to estimate the international percentage of children and youth that achieved each CRF standard; and to discuss strategies to enhance the utility of CRF criterion-referenced standards for population health research.

Criterion-referenced standards

Our paper identified 10 studies that reported CRF criterion-referenced standards using samples of children and youth from Europe, North America or South America. The reported criterion-referenced standards for CRF were developed using a variety of methods and protocols, and as a result, the standards vary considerably across sex by age groups. Recent research suggests that temporal trends exist in CRF³³ and body composition in children,³⁷ and metabolic health markers in adults.³⁸ These trends suggest that the relationship between CRF and health is not static, which may also help describe the different sex by age specific standards across studies given our 10-year time period between the first³⁹ and last⁴⁰ reported standards. For these reasons, comparisons of percentages of healthy CRF while using different standards should be interpreted with caution.

Table 3 Percentage of the 1 142 026 international 20mSRT performance pseudo-datasets that met each CRF criterion-referenced standard

Age (years)	9	10	11	12	13	14	15	16	17
Girls n=552 900									
Ruiz, 2007 ⁴¹	99.9	98.6							
Lobelo, 2009 ⁵³				90.8	82.9	71.7	59.7	49.3	40.0
Welk, 2011 ³⁶		91.3	77.0	67.8	55.6	47.1	37.8	24.7	19.8
Adegboye, 2011 ³⁵	99.9	93.5	93.5			87.7	78.8	68.9	57.3
Moreira, 2012		93.3	82.5	72.0	56.6	46.6	34.5	21.8	16.6
Silva, 2012 ²⁹		64.5	61.0	65.8	45.1	47.1	47.9	45.6	48.5
Boddy, 2012 ⁵⁴	93.7	81.3							
Ruiz, 2015					89.2	80.7	70.1	57.8	46.6
Silva, 2016 ⁴⁰						92.3	85.2	74.8	64.5
Boys n=589 126									
Mesa, 2006 ³⁹					36.7	34.3	32.2	32.5	36.9
Ruiz, 2007 ⁴¹	92.9	87.0							
Lobelo, 2009 ⁵³				59.4	58.8	52.4	51.9	65.2	59.2
Welk, 2011 ³⁶		94.8	87.6	81.5	73.7	60.4	53.9	45.8	41.6
Adegboye, 2011 ³⁵	88.3	79.0	67.4			41.4	40.8	35.5	32.9
Moreira, 2012		88.8	79.7	73.0	71.2	64.0	63.1	57.2	53.3
Silva, 2012 ²⁹		56.8	60.6	66.9	59.8	48.6	55.5	44.0	47.6
Boddy, 2012 ⁵⁴	66.1	53.4							
Ruiz, 2015					59.8	54.3	53.0	47.3	43.7
Silva, 2016 ⁴⁰						73.8	71.8	67.0	60.2

20mSRT, 20-m shuttle run test; CRF, cardiorespiratory fitness.

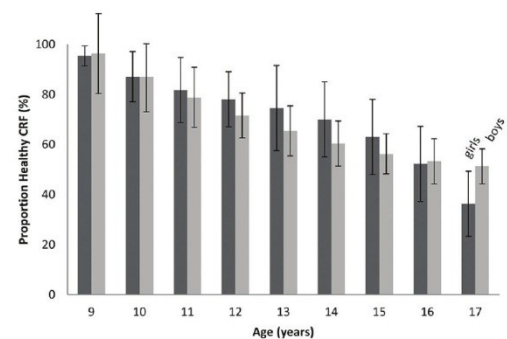


Figure 2 Overall smoothed population-weighted mean percentages and 95% CIs for healthy CRF in girls and boys. Calculated across age-specific percentage estimates for each criterion-referenced standard ($\dot{V}O_{2peak}$) available in table 3. CRF, cardiorespiratory fitness.

Percentage estimates for 'healthy' CRF

Our study also provides percentage estimates for each reported criterion-referenced standard using international data that represent 1 142 026 children and youth aged 9–17 years. These data demonstrate that a high percentage of young children (9–10 years old) are apparently healthy. In some cases (Ruiz⁴¹ and Adegboye³⁵), girls aged 9 and 10 years would only need to achieve a 20mSRT performance below the 5th percentile to achieve the criterion-referenced standard when using the Léger prediction equation, equivalent to completing the first stage (1 min) of the 20mSRT. Comparatively, boys aged 9–10 years would have to perform above the 5th percentile to achieve most criterion-referenced standards, suggesting that a relatively better performance is required for younger boys than younger girls. Furthermore, these data illustrate an age gradient where older children must achieve a higher 20mSRT percentile than younger children to achieve the criterion-referenced standards. This translates into a higher overall percentage of healthy CRF in younger children than older children. For example, data in table 3 indicate that young children (9–10 years old) are apparently healthy with 72% of all cases having a success rate over 80%, in comparison with only 7% of older children (14–17 years old).

The FITNESSGRAM criterion-referenced standards (Welk *et al*³⁶) for CRF are probably the most widely used standards to date. National percentages of children and youth who met these standards exist for USA,⁴² Chile⁴³ and Hungary.⁴⁴ In USA and Hungary, a sex gradient existed where boys met the standards more often than girls at every age group. In addition, an age gradient existed for these countries where younger children met the standards more often than older children, supporting the pattern observed in Welk's percentages in table 3. The US-based percentages⁴² indicated that 26%–49% of girls and 38%–62% of boys had healthy CRF, which was lower than our international percentages (using the Welk standards). This result is expected given that US children and youth have low CRF relative to international norms.^{11 12} Our international percentage estimates are more similar to the Chilean estimates⁴³ which reported healthy CRF percentages in grade 8 children (13- and 14-year-olds) for 45% and 74% of girls and boys, respectively. Furthermore, percentage estimates for Hungarian children and youth aged 11–18 years⁴⁴ are between 32% to 86% for girls and between 45% to 86% for boys. These percentage estimates fall above our international percentage estimates, being consistent with previous findings showing that Hungarian (central European) children have better CRF relative to international norms.^{11 12}

In addition, the Welk standards for boys are among the easiest to achieve and resulted in a higher percentage of boys with healthy CRF than girls. In comparison, most other reported criterion-referenced standards identified higher percentages of girls with healthy CRF than boys (table 3). This highlights the fact that using standards other than the Welk standards would have most likely resulted in a higher percentage of girls than boys with healthy CRF in USA, Chile and Hungary. Thus, the criterion-referenced standard applied to the sample makes a difference in the overall interpretation of results.

International percentages of healthy children and youth

Figure 2 displays the international percentages of children and youth with apparently healthy CRF. However, these percentages should be interpreted with caution because of the variability across all 10 criterion-referenced standards, resulting in large 95% CIs. In addition, this study applied the Léger prediction equation which may yield systematically different $\dot{V}O_{2peak}$

values in comparison with other prediction equations. This may impact the published criterion-referenced standards as well as the percentage estimates of healthy CRF across each sex by age group, indicating that the prediction equation used may artificially increase or decrease the percentage estimates in a systematic way.

Strategies for population health research

The availability of different criterion-referenced standards for CRF across the same sex by age groups may create confusion. A recent meta-analysis by Ruiz *et al*⁴⁵ combines seven published criterion-referenced standards^{35 36 39 41 47 50 53} on 9280 children and youth aged 8–19 years from 14 countries and provides the first international criterion-referenced standards for this age group. Their meta-analysis determined that fitness levels below 35 and 42 mL/kg/min for girls and boys, respectively, should raise a red flag. Internationally, 80% (95% CI 66% to 94%) of girls and 68% (95% CI 60% to 76%) of boys meet the international criterion-referenced standards, with younger children meeting the standards more often than older children. See online supplement 1 that provides an unweighted sex by age percentage for the Ruiz standards using the Tomkinson international pseudo-dataset. The Ruiz standards could be used as the interim overarching international standards to help compare percentages of apparently healthy children and youth across studies, countries and continents, providing clear international comparisons in this area of research. International standards could also be applied in a clinical and/or school setting by clinicians and practitioners, similar to the international body mass index growth curves⁴⁸ for children and youth, providing an effective screening instrument to help identify children and youth who need to increase their physical activity levels to improve overall health.

Our study also reports international referenced percentages for CRF in children and youth aged 9–17 years. These data are the first of their kind and carry specific utility for national and international population health research. These international referenced percentages can help researchers and practitioners (eg, physical educators, public health workers, health promotion practitioners, etc) interpret their percentage estimates by providing a reference to indicate whether their sample percentage is above or below the international percentage for a given sex by age group. This can help provide researchers and practitioners with context to their respective results. These international referenced percentages can also provide the first step towards establishing CRF and the 20mSRT as a tool for population health screening and surveillance.

Strengths and limitations

Our study compared 10 reported CRF criterion-referenced standards for children and youth, all of which were developed using the ROC curve method. We also describe percentages of apparently healthy children and youth using international normative centile values developed from a sample of over 1.1 million children and youth from 500 different countries.²³ These normative values were developed using rigorous data standardisation, post-stratification population weighting and LMS techniques.

A limitation of our study may be comparing criterion-referenced standards across different studies. For instance, study-specific standards were generated using different combinations of CRF field-based or laboratory-based assessment protocols, submaximal-effort or maximal-effort protocols, different exercise modes and $\dot{V}O_{2peak}$ prediction equations. This is of particular concern given that treadmill protocols tend to elicit

Original article

a $\dot{V}O_{2peak}$ roughly 9% higher than cycle ergometer protocols,⁴⁹ and the 20mSRT uses prediction equations to estimate $\dot{V}O_{2peak}$ that may introduce unintended error causing varying results.⁵⁰ Thus, each method used to estimate $\dot{V}O_{2peak}$ may yield different results due to different underlying levels of construct validity. Furthermore, each criterion-referenced standard used a different combination of clustered cardiometabolic health markers to establish the criterion domain. As a result, it is plausible that an individual is healthy given one set of cardiometabolic criteria, but unhealthy given a different set of criteria. This could explain some of the variability in our population percentage estimates. The existing criterion-referenced standards also fail to identify psychosocial health outcomes in children and youth, which remains an essential area of research given the importance of mental health.⁵¹ Furthermore, these standards do not incorporate a measure of maturation and the use of ratio scaling means that body size differences may not have been appropriately partitioned out, both of which may have an impact on identifying developmentally appropriate age-specific and sex-specific standards.⁵²

CONCLUSION

Our study provides a review of criterion-referenced standards for CRF and demonstrates the inconsistencies among different standards using pseudo-data derived from a large international sample of 20mSRT performances. This study also provides international referenced percentages of children and youth meeting each criterion-referenced standard. Although there are several limitations to ROC-generated criterion-referenced standards, these represent the best available data, and percentage estimates should be interpreted with caution. Nonetheless, these standards are the first step towards generating clinically relevant standards that are comparable with body mass index standards. Researchers, clinicians and the general public are encouraged to use the interim international criterion-referenced standards of 35 and 42 mL/kg/min for girls and boys, respectively, to identify children and youth at risk of poor health—raising a clinical red flag.

What are the new findings?

- Ten studies reported criterion-referenced standards for healthy cardiorespiratory fitness (CRF) among children and youth, with the percentage meeting each standard varying by sex and age.
- An age gradient exists where younger children are apparently healthier than older children, regardless of sex.
- There is a need to encourage the use of a single set of overarching criterion-referenced standards to improve international comparisons of healthy CRF among children and youth.

Acknowledgements The authors would like to acknowledge Kevin Belanger and Joel Barnes for their help with formatting and designing the figures for this manuscript.

Contributors JLL, GRT and MST developed the research questions and objectives. JLL led the data analysis and synthesis of results, and drafted the manuscript. All authors contributed to the interpretation of results, editing and reviewing of the final manuscript. All authors approved the final manuscript.

Funding FBO and JRR have received a grant from the Spanish Ministry of Science and Innovation (grant nos. RYC-2011-09011 and RYC-2010-05957, respectively), and FBO's current research activity is under the umbrella of the ActiveBrains project (Reference DEP2013-47540).

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- 1 Armstrong N, Tomkinson G, Ekelund U. Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during youth. *Br J Sports Med* 2011;45:849–58.
- 2 Ortega FB, Ruiz JR, Castillo MJ, et al. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes* 2008;32:1–11.
- 3 Jiménez-Pavón D, Castillo MJ, Moreno LA, et al; Helena Study Group. Fitness and fatness are independently associated with markers of insulin resistance in European adolescents; the HELENA study. *Int J Pediatr Obes* 2011;6:253–60.
- 4 Thomas NE, Baker JS, Graham MR, et al. C-reactive protein in schoolchildren and its relation to adiposity, physical activity, aerobic fitness and habitual diet. *Br J Sports Med* 2008;42:357–60.
- 5 Morales PF, Sánchez-López M, Moya-Martínez P, et al. Health-related quality of life, obesity, and fitness in schoolchildren: the Cuenca study. *Qual Life Res* 2013;22:1515–23.
- 6 Ortega FB, Labayen I, Ruiz JR, et al. Improvements in fitness reduce the risk of becoming overweight across puberty. *Med Sci Sports Exerc* 2011;43:1891–7.
- 7 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.
- 8 Höglström G, Nordström A, Nordström P. High aerobic fitness in late adolescence is associated with a reduced risk of myocardial infarction later in life: a nationwide cohort study in men. *Eur Heart J* 2014;35:3133–40.
- 9 Höglström G, Nordström A, Nordström P. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. *Int J Epidemiol* 2015;1159–68.
- 10 Léger L, Lambert J, Goulet A, et al. Capacité aérobie des québécois de 6 à 17 ans - Test navette de 20 mètres avec paliers de 1 minute [Aerobic capacity of 6 to 17 year old Québécois – 20 metre shuttle run test with 1 minute stages]. *Can J Appl Sport Sci* 1984;9:64–9.
- 11 Olds T, Tomkinson G, Léger L, et al. 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;24:1025–38.
- 12 Lang JJ, Tremblay MS, Léger L, et al. International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results. *Brit J Sports Med*. 2016.
- 13 Mayorga-Vega M, Aguilar-Soto P, Viciano J. Criterion-related validity of the 20-M shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *J Sports Sci Med* 2015;14:536–47.
- 14 Tomkinson GR, Olds TS. Field tests of fitness. In: Armstrong N, van Mechelen W, eds. *Paediatric Exercise Science and Medicine*. New York, NY: Oxford University Press, 2000:109–28.
- 15 Council of Europe. Eurofit: handbook for the EUROFIT tests of physical fitness. *Secretariat of the Committee for the Development of Sport within the Council of Europe*. Rome: Council of Europe, 1998:1–75.
- 16 In: Plowman SA, Meredith MD, eds. *Fitnessgram/Activitygram Reference Guide*. 4th Edn. Dallas, TX: The Cooper Institute, 2013:1–202.
- 17 Longmuir PE, Boyer C, Lloyd M, et al. The Canadian Assessment of physical Literacy: methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Health* 2015;15:767.
- 18 Australian Council for Health, Physical Education and Recreation. *Australian Fitness Education Award: User's Manual and Curriculum Ideas*. Physical Education and Recreation. Australia, SA: Australian Council for Health, 1996:1–59.
- 19 Ruiz JR, Castro-Piñero J, España-Romero V, et al. Field-based fitness assessment in young people: the ALPHA health-related fitness test battery for children and adolescents. *Br J Sports Med* 2011;45:518–24.
- 20 Cadenas-Sánchez C, Alcántara-Moral F, Sánchez-Delgado G, et al. Assessment of cardiorespiratory fitness in preschool children: Adaptation of the 20 metres shuttle run test. *Nutr Hosp* 2014;30:1333–43.
- 21 Cadenas-Sánchez C, Martine-Tellez B, Sánchez-Delgado G, et al. Assessing physical fitness in preschool children: Feasibility, reliability and practical recommendations for the PREFIT battery. *J Sci Med Sport* 2016;19:910–5.
- 22 Ortega FB, Cadenas-Sánchez C, Sánchez-Delgado G, et al. Systematic review and proposal of a field-based physical fitness-test battery in preschool children: the PREFIT battery. *Sports Med* 2015;45:533–55.
- 23 Tomkinson GR, Lang JJ, Tremblay MS, et al. International normative 20 M shuttle run values from 1142026 children and youth representing 50 countries. *Brit J Sports Med* 2016;0:1–14.

- 24 Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9-17-year-old Australians since 1985. *Br J Sports Med* 2013;47:98–108.
- 25 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;30:679–87.
- 26 Santos R, Mota J, Santos DA, et al. Physical fitness percentiles for Portuguese children and adolescents aged 10–18 years. *J Sports Sci* 2014;32:1510–8.
- 27 Sauka M, Priedite IS, Artjuhova L, et al. Physical fitness in northern European youth: reference values from the Latvian Physical Health in Youth Study. *Scand J Public Health* 2011;39:35–43.
- 28 Secchi JD, García GC, España-Romero V, et al. Physical fitness and future cardiovascular risk in Argentine children and adolescents: an introduction to the ALPHA test battery. *Arch Argent Pediatr* 2014;112:132–40.
- 29 Tambalis KD, Panagiotakos DB, Psarra G, et al. Physical fitness normative values for 6-18-year-old Greek boys and girls, using the empirical distribution and the lambda, mu, and sigma statistical method. *Eur J Sports Sci* Sept 24th 2015.
- 30 McNeil BJ, Keller E, Adelstein SJ. Primer on certain elements of medical decision making. *N Engl J Med* 1975;293:211–5.
- 31 Levy PS, Lemeshow S. Stratification random sampling: further issues. In: Levy PS, Lemeshow S, eds. *Sampling of Populations: Methods and Application*. Hoboken, NJ: John Wiley & Sons, Inc, 2008:143–88.
- 32 Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med* 1992;11:1305–19.
- 33 Tomkinson GR, Léger LA, Olds TS, et al. 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;33:285–300.
- 34 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.
- 35 Adegboye AR, Anderssen SA, Froberg K, et al. Recommended aerobic fitness level for metabolic health in children and adolescents: a study of diagnostic accuracy. *Br J Sports Med* 2011;45:722–8.
- 36 Welk GJ, Laurson KR, Eisenmann JC, et al. Development of youth aerobic-capacity standards using receiver operating characteristic curves. *Am J Prev Med* 2011;41:S111–S116.
- 37 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* 2009;21:7–20.
- 38 Danaei G, Finucane MM, Lu Y, et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *The Lancet* 2011;378:31–40.
- 39 Mesa JL, Ruiz JR, Ortega FB, et al. Aerobic physical fitness in relation to blood lipids and fasting glycaemia in adolescents: Influence of weight status. *Nutr Metab Cardiovasc Dis* 2006;16:285–93.
- 40 Silva DA, Tremblay M, Pelegrini A, et al. Association between aerobic fitness and high Blood Pressure in adolescents in Brazil: evidence for criterion-referenced cut-points. *Pediatr Exerc Sci* 2016;28:312–20.
- 41 Ruiz JR, Ortega FB, Rizzo NS, et al. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. *Pediatr Res* 2007;61:350–5.
- 42 Bai Y, Saint-Maurice PF, Welk GJ, et al. Prevalence of youth fitness in the United States: baseline results from the NFL PLAY 60 FITNESSGRAM partnership project. *J Pediatr* 2015;167:662–8.
- 43 Garber MD, Sajuria M, Lobelo F. Geographical variation in health-related physical fitness and body composition among Chilean 8th graders: a nationally representative cross sectional study. *PLoS One* 2014;9:e108053.
- 44 Welk GJ, Saint-Maurice PF, Csányi T. Health-related physical fitness in Hungarian youth: age, sex, and regional profiles. *Res Q Exerc Sport* 2015;86(Suppl 1): S45–S57.
- 45 Ruiz JR, Cavaero-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. *Brit J Sports Med*. 2015.
- 46 Ruiz JR, Silva G, Oliveira N, et al. Criterion-related validity of the 20-m shuttle run test in youths aged 13–19 years. *J Sports Sci* 2009;27:899–906.
- 47 Ruiz JR, Huybrechts I, Cuenca-Gracia M, et al. Cardiorespiratory fitness and ideal cardiovascular health in European adolescents. *Heart* 2014;0:1–8.
- 48 Cole TJ, Bellizzi MC, Flegal KM, et al. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1240–3.
- 49 Armstrong N, Davies B. An ergometric analysis of age group swimmers. *Br J Sports Med* 1981;15:20–6.
- 50 Moreira C, Santos R, Ruiz JR, et al. Comparison of different $\dot{V}O_{2max}$ equations in the ability to discriminate the metabolic risk in Portuguese adolescents. *J Sci Med Sport* 2011;14:79–84.
- 51 Tomkinson G. Aerobic fitness thresholds for cardio metabolic health in children and adolescents. *Br J Sports Med* 2011;45:686–7.
- 52 Armstrong N, Welsman JR. Aerobic fitness: What are we measuring? *Sports Med* 2003;33:5–25.
- 53 Lobelo F, Pate RR, Dowda M, et al. Validity of cardiorespiratory fitness criterion-referenced standards for adolescents. *Med Sci Sports Exerc* 2009;41:1222–9.
- 54 Boddy LM, Thomas NE, Fairclough SJ, et al. ROC generated thresholds for field-assessed aerobic fitness related to body size and cardiometabolic risk in schoolchildren. *PLoS One* 2012;7:e45755.
- 55 Silva G, Aires L, Mota J, et al. Normative and criterion-related standards for shuttle run performance in youth. *Pediatr Exerc Sci* 2012;24:157–69.

CHAPTER 5: MANUSCRIPT 3

This manuscript has been written and formatted following the publication standards of the peer-reviewed journal, *BMC Public Health*.

Title:

Cardiorespiratory fitness is associated with physical literacy among Canadian children aged 8 to 12 years.

Status:

Submitted

Cardiorespiratory fitness is associated with physical literacy among Canadian children aged 8 to 12 years

Justin J. Lang¹, Jean-Philippe Chaput¹, Patricia E. Longmuir¹, Joel D. Barnes¹, Kevin Belanger¹, Grant R. Tomkinson^{2,3}, Kristal D. Anderson⁴, Brenda Bruner⁵, Jennifer L. Copeland⁶, Melanie J. Gregg⁷, Nathan Hall⁷, Angela M. Kolen⁸, Kirstin N. Lane⁴, Barbi Law⁵, Dany J. MacDonald⁹, Luc J. Martin¹⁰, Travis J. Saunders⁹, Dwayne Sheehan¹¹, Michelle R. Stone¹², Sarah J. Woodruff¹³, Mark S. Tremblay¹

1. Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada.
2. Department of Kinesiology and Public Health Education, University of North Dakota, Grand Forks, ND, USA.
3. Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health Sciences & Sansom Institute for Health Research, University of South Australia, Adelaide, SA, Australia.
4. Camosun College, Victoria, BC, Canada.
5. Nipissing University, North Bay, ON, Canada.
6. University of Lethbridge, Lethbridge, AB, Canada.
7. University of Winnipeg, Winnipeg, MB, Canada.
8. St. Francis Xavier University, Antigonish, NS, Canada.
9. University of Prince Edward Island, Charlottetown, PE, Canada.
10. Queen's University, Kingston, ON, Canada.
11. Mount Royal University, Calgary, AB, Canada.
12. Dalhousie University, Halifax, NS, Canada.
13. University of Windsor, Windsor, ON, Canada.

Corresponding Author: Justin J. Lang

Keywords: physical fitness, physical activity, cognition, motivation, confidence, knowledge, understanding.

ABSTRACT

Background: The associations between cardiorespiratory fitness (CRF) and physical literacy in children are largely unknown. The aim of this study was to assess the relationships between CRF, measured using the 20-meter shuttle run test (20mSRT), and components of physical literacy among Canadian children aged 8–12 years. The secondary aim was to estimate the percentage of children with healthy CRF. **Methods:** A total of 9,393 (49.9% female) children, with a mean (SD) age of 10.1 (1.2) years, from a cross-sectional surveillance study were included for this analysis. The 20mSRT was assessed using a standardized protocol. The four domains of physical literacy (physical competence, daily behavior, motivation and confidence, and knowledge and understanding) were assessed using the Canadian Assessment of Physical Literacy. The Ruiz interim international standards of 42 and 35 ml/kg/min for boys and girls, respectively, were used to estimate the percentage of children with healthy CRF. **Results:** CRF was strongly and favorably associated with all components of physical literacy among school-aged Canadian children. The effect size between low and high CRF groups was large for the physical competence domain (Cohen's d range: 1.11–1.94) across sex and age groups, followed by moderate to large effect sizes for motivation and confidence (Cohen's d range: 0.54–1.18), small to moderate effect sizes for daily behavior (Cohen's d range: 0.25–0.81), and marginal to moderate effect sizes for knowledge and understanding (Cohen's d range: 0.08–0.70). Furthermore, in comparison with boys (mean percentage: 64.3[0.7]), there was a higher percentage of girls (mean percentage: 98.4[0.2]) with healthy CRF across all age groups. The percentage with healthy CRF declined with age, only in boys. **Conclusions:** This study identified strong associations

between CRF and physical literacy and its constituent components among 8-12 year-old children. The 20mSRT could be implemented as a screening assessment in the school-based setting to help physical educators identify those at risk of low physical literacy levels.

INTRODUCTION

Physical literacy (PL) has emerged as a core construct in the physical education paradigm, designed to support children's engagement in physical activity throughout the life course [1, 2]. To better understand the state of PL among Canadian children, an 11-site cross-sectional surveillance study called the RBC-Canadian Assessment of Physical Literacy Learn to Play project (RBC-CAPL) was conducted. The validated CAPL instrument was specifically chosen to measure each of the four domains that describe PL (i.e., physical competence, daily behavior, motivation and confidence, and knowledge and understanding) [3, 4]. The CAPL can be used as a surveillance tool to understand the overall PL levels of children and to identify the PL domains that may require targeted interventions. However, emerging issues with utilizing the CAPL in school-based settings are the amount of time required to assess large groups of children (approximately 90 minutes with five appraisers to test 25 children), and that teachers may not have the expertise or available resources to conduct a full CAPL assessment.

The 20-meter shuttle run test (20mSRT) [5, 6], a reliable and valid assessment of cardiorespiratory fitness (CRF) among children [7, 8], is a measure that informs the physical competence domain of the CAPL. The 20mSRT lends itself well to population-based surveillance because it is easy to administer, requires minimal equipment, and can be used to assess large groups of children simultaneously [9]. Furthermore, several studies have identified favorable associations between 20mSRT performance and aspects of PL among children and youth, including physical fitness [10], daily physical activity [11], cognitive ability [12, 13], and psychosocial health [14, 15].

To our knowledge, no study has assessed the associations between CRF and overall/domain-specific PL among children. Specifically, assessing the relationship between CRF and the four domains of PL identified in the consensus definition could provide further insight into the importance of CRF, and potential strategies for PL screening [2]. The RBC-CAPL provides an opportunity to assess these associations in a large sample of Canadian children, and to quantify the extent to which CRF relates to various components of PL.

The main objective of this study was to evaluate the associations between 20mSRT performance (i.e., CRF) and each component of PL assessed using the CAPL among Canadian school-aged children (8–12 years). The secondary objective was to determine the percentage of Canadian children who attained healthy levels of CRF using the Ruiz standards [16]. We hypothesized that children with higher levels of CRF would have higher overall PL scores across all for domains of the CAPL, in comparison to those with low levels of CRF.

METHODS

Study design

The RBC-CAPL was a cross-sectional surveillance study that took place between 2014 and 2016, and was designed to evaluate the PL levels of Canadian children using a standardized data collection protocol. The study design included 11 data collection sites from seven Canadian provinces, which included Victoria (British Columbia), Lethbridge

and Calgary (Alberta), Winnipeg (Manitoba), North Bay, Windsor, and Ottawa (Ontario), Trois-Rivières (Québec), Halifax and Antigonish (Nova Scotia), and Charlottetown (Prince Edward Island). The aim was to recruit up to 1,300 participants per site over a 3-year data collection period. Each site was also tasked with recruiting participants from urban (minimum of 50% of the sample) and rural (minimum of 20% of the sample) locations, while ensuring a balanced representation of high, medium, and low income communities. Ethics approval for this project was originally obtained from the Children's Hospital of Eastern Ontario Research Ethics Board (Ottawa, coordinating centre). Each site subsequently obtained approval from their respective research ethics board. Consent and ethics approval were also obtained from all participating school boards, summer camps, community centres, and/or sport leagues. Written informed consent was obtained from parents or legal guardians, and participating children provided assent.

Participants

Participant recruitment locations were selected across all sites using purposive, non-randomized sampling. Elementary schools across all sites were the primary participant recruitment locations for this study, whereas summer camps, community centres, and sport leagues were the secondary participant recruitment locations. Participants were considered eligible for this study if between the ages of 8.00–12.99 years (grades 4–6) and maximal effort exercise was not contraindicated. All eligible participants were invited to participate in this study, and potential participants were able to opt-out for any reason, without consequence. Of the 10,030 participants who took part in RBC-CAPL, a total of 9,393 remained in the present analysis after excluding participants without a

20mSRT score (n=637). Body mass index across sex and age groups of those excluded from the analysis did not differ significantly (range of p values: 0.13–0.84) from those included in the present analysis. Physical activity levels were significantly lower ($p<0.006$) for 8 year olds who were excluded from the analysis (n=30) due to missing data, but there were no differences (range of p values: 0.07–0.49) for children aged 9–12 years.

Data collection procedures

All data collection staff had a background in fitness or physical activity assessment, and each site's coordinators were subsequently trained by research staff from the coordinating centre (Ottawa). Data collection procedures followed the published CAPL protocol [3, 4], which provides standardized procedures to collect data across the four PL domains: physical competence, daily behavior, motivation and confidence, and knowledge and understanding.

Independent variable

CRF was assessed using the 15 m or 20mSRT protocols [5, 6]. All children were asked to run back and forth between two parallel lines, 15 m or 20 m apart, following the pace of an audio signal that began at a speed of 8.5 km/h and increased by 0.5 km/h at every 1-minute intervals. The total number of laps (shuttles) completed was recorded for each participant, and all data from the 15 m protocol were converted to the 20 m protocol using a conversion chart [17]. Researchers used indoor gymnasiums as the primary testing location, with outdoor locations used as a back-up location when necessary.

Participants were encouraged at all times to run a maximal effort test. Following the Tomkinson recommendations [18], 20mSRT performance for this study was reported as the running speed at the last completed stage and number of laps completed.

Dependent variables

Physical literacy

PL was assessed using the CAPL instrument which provides methods to assess the four domains of PL, as described below. The total CAPL score is an aggregate that combines all domains (physical competence [maximum of 32 points], daily behavior [maximum of 32 points], motivation and confidence [maximum of 18 points], and knowledge and understanding [maximum of 18 points]) and ranges from 0 (poor PL) to 100 (excellent PL) points [3]. The total CAPL score and each of the four domain scores were used to summarize the associations between CRF and PL.

Physical competence

The physical competence domain was modified from the original CAPL methods to provide an aggregate score that excluded the 20mSRT. Thus, the physical competence domain included three health-related fitness assessments, three anthropometric assessments, and one gross motor movement skill assessment. Grip strength was assessed using a handgrip dynamometer following established procedures [19]. The better score, measured to the nearest 0.5 kg, from each of the left and right hands were combined. The prone plank test was used to assess torso muscular endurance [20]. Participants were asked to hold a static prone position on their elbows and toes with a straight body

position from the ankles to the head for as long as possible with the time to exhaustion (nearest 0.1 second) recorded as the final score. Flexibility was assessed using the sit-and-reach protocol with a flexometer [19]. With shoes removed, participants were asked to sit with their legs stretched out in front of them and their knees flat on the floor. They were asked to extend their arms with their hands stacked while bending forward at the hips with legs remaining straight. The furthest distance attained while reaching forward toward their toes was recorded to the nearest 0.5 cm. Waist circumference was measured to the nearest 0.5 cm at the top of the iliac crest, using standardized procedures [19]. Standing height was assessed to the nearest 0.1 cm using a stadiometer, and body weight was recorded to the nearest 0.1 kg using a digital weighing scale. Body mass index was calculated from the measured height and weight values (kg/m^2).

Gross motor movement skills were assessed using the Canadian Agility and Movement Skill Assessment (CAMSA) protocol [21]. The CAMSA is a standardized agility course that provides a method to rapidly assess fundamental and complex movement skills (jumping, sliding, catching, throwing, skipping, hopping, and kicking) in a way that incorporates various ‘real-world’ movement capacities (coordination, balance, precision, acceleration, and deceleration). The overall CAMSA score combines movement quality scores with the obstacle course completion time to provide an overall score between 1.5 (low performer) and 42 (high performer) [21].

Daily behavior

The daily behavior domain assessed participants' engagement in physical activity and sedentary behaviors as three separate components: objective physical activity, self-report physical activity, and self-report screen time. Daily physical activity was assessed as the average number of steps taken each day using a SC-StepRx pedometer (StepsCount, Deep River, ON, Canada) [22]. Participants were asked to record their daily step counts on a pedometer tracking log before bedtime over the course of seven days. Participants were also asked to record the time that put on and took off the pedometer, the amount of missing time if the pedometer was removed during the day, and the reason why they may have removed the pedometer (e.g., forgot to wear, water-related activity). Pedometer data were considered valid if the total number of steps per day was between 1,000 and 30,000; at least 10 hours of wear time was accumulated per day; and at least three valid days were recorded [23]. Children were asked to self-report the average number of days per week that they performed at least 60 minutes of moderate to vigorous intensity physical activity. Children were also asked to self-report the time spent using screens (i.e., watching television, playing video games, computer games, or other screen-based devices) on a typical school day and weekend day as a measure of sedentary behavior.

Motivation and confidence

The motivation and confidence domain was derived from five components. A published scale was used to derive perceptions of physical activity benefits and barriers [24].

Children were asked how their activity levels compared with their peers, and how their skill level compared with their peers. Lastly, sub-scales of the Children's Self-Perception

and Adequacy in and Predilection for Physical Activity (CSAPPA) scale were used to assess participants' adequacy and predilection toward physical activity participation [24].

Knowledge and understanding

Knowledge and understanding was assessed using a standardized questionnaire (CAPL questionnaire) developed to reflect the grades 4, 5, and 6 Canadian curricula for physical and health education [3]. The questions were broadly related to the understanding of the physical activity and sedentary behavior guidelines, health-related fitness components, physical activity safety equipment, and methods for improving movement skills.

Participants completed the questionnaire using paper and pencil or an online website format.

Data analysis

All analyses were conducted using SAS 9.4 (SAS Institute Inc., Cary, North Carolina, USA). Descriptive characteristics for all variables were calculated as means and standard deviations. The independent variable, 20mSRT laps, was divided into tertiles across each age (i.e., age at the last birthday) by sex group, representing low, medium, and high CRF. To determine the analytical plan to assess the associations between CRF and each dependent variable, sex and age interaction terms were tested. First, significant sex interactions were identified between CRF and most components of PL (i.e., 16/20 significant interactions). Next, significant age interactions by sex were identified for more than half of the components of PL (i.e., 24/36 significant interactions). As a result of these significant interactions, a stratified analysis by sex and age was conducted. Thus,

mean scores for all components of PL were calculated for each sex by age CRF tertile group, and significant differences across groups were assessed using ANOVA. Post-hoc analyses (Dunn's test with Bonferroni correction) were used to identify significant differences between groups, using low CRF as the reference group. Effect sizes between low and high CRF groups were calculated using Cohen's d , and effect sizes were interpreted as small (d between 0.2 and 0.5), moderate (d between 0.5 and 0.8), and large ($d > 0.8$), with effect sizes below 0.2 considered trivial [27]. To estimate the percentage of children with healthy CRF, we used the interim international CRF standards of 42 and 35 ml/kg/min for boys and girls, respectively [16, 26].

RESULTS

Participants

A total of 9,393 (49.9% girls) children aged 8–12 years, from 11 Canadian sites, were included in this study. Nearly half (42.6%) of the sample was from Western Canada (i.e., Victoria, Lethbridge, Calgary, and Winnipeg). Table 1 provides the descriptive characteristics for all included participants.

<<< Insert Table 1 >>>

Associations across components of physical literacy

Tables 2–6 display the associations between CRF and all components of PL, with each table presenting the data for boys and girls of the same age. There were significant main effects across CRF tertile groups for most components of PL that were consistent and in the expected direction across all sex and age groups. The effect sizes (Cohen's d)

between low and high CRF groups ranged from small to large, with very few components of PL considered trivial. Generally, larger effect sizes were observed for boys than girls, and the effect sizes generally increased with age. The physical competence domain score consistently displayed the largest effect size (Cohen's d range: 1.11–1.94) across sex and age groups, followed by the total CAPL score (Cohen's d range: 0.92–1.60), the CAMSA (Cohen's d range: 0.97–1.52) and the plank (Cohen's d range: 0.86–1.36). The smallest effect size across sex and age groups was generally the knowledge and understanding domain (Cohen's d range: 0.08–0.70), followed by the sit and reach score for boys (Cohen's d range: 0.29–0.40), and the handgrip score for girls (Cohen's d range: 0.24–0.41).

<<< Insert Tables 2 to 6 >>>

Percentage of children with healthy CRF

The percentages of children with healthy CRF for each age group are presented in Figure 1. The mean percentage of children meeting the CRF standard was 64.3% (SD = 0.7) for boys and 98.4% (SD = 0.2) for girls. Although girls maintained an equivalent percentage of healthy CRF across age groups, the percentage of boys with healthy CRF decreased with age.

<<< Insert Figure 1 >>>

DISCUSSION

This study represents the largest effort to assess the associations between CRF and components of PL among school-aged children. Our findings suggest that there are clear

PL related benefits to having high CRF levels, supporting our hypothesis. For instance, those in the high CRF tertile consistently demonstrated better scores across all domains of PL in comparison with their peers in lower CRF tertile groups, regardless of sex and age. Of the PL components, the strongest associations were identified between CRF and physical competence (large effect size), followed by motivation and confidence (moderate to large effect size), daily behavior (small to moderate effect size), and knowledge and understanding (trivial to moderate effect size). In addition, findings from this study suggest that boys, in comparison with girls, have lower percentages of apparently healthy CRF and these percentages decreased as age increased.

Associations across components of physical literacy

Our study identifies consistent and favorable gradients across CRF tertiles for all components that describe PL. This finding suggests that CRF, measured using the 15mSRT or 20mSRT, is an important correlate of all PL domains and almost all components of the CAPL among Canadian children aged 8–12 years, supporting a growing body of literature that highlights the importance of CRF in these age groups [30].

This study identifies strong associations between CRF and most components of the physical competence domain, which supports the literature [10]. Of the components that describe physical competence, the largest effect sizes across sex and age groups were identified for the CAMSA (i.e., gross motor skills) and the plank test (i.e., muscular endurance). Although this study is the first to identify these strong associations, as the

CAMSA and the plank are novel assessments used in the CAPL, other studies using different tests support our findings. For instance, a previous study has identified strong positive associations between the Test of Gross Motor Development-2nd Edition and achieving the FITNESSGRAM® Healthy Fitness Zone for the 20mSRT [31, 32]. This further suggests that CRF is strongly related to gross motor skills. In addition, another study reported marginal but significant correlations between CRF and the ability to perform repeated push-ups and curl-ups [33], measures of muscular endurance, similar to the plank test. These associations suggest that future school-based interventions designed to improve CRF could incorporate skill development aimed at teaching fundamental motor skills (i.e., jumping, sliding, catching, throwing, skipping, hopping, and kicking), movement capacities (i.e., coordination, balance, precision, acceleration, and deceleration), and muscular endurance. These types of interventions may have an indirect or mediated effect by providing children with the abilities needed to further improve CRF. Indeed, many physical education programs incorporate these aspects of skill development in their respective curriculums.

In our study, the association between CRF and adiposity (body mass index and waist circumference) ranged from moderate to large. While the relationship between CRF and adiposity is certainly important, it may not be the most vital aspect associated with 20mSRT performance. Previous research has showed that adiposity explained between 40 and 60% of declines in distance running seen in children (aged 10–12 years) between 1985 and 1997 [34]. Indeed, 20mSRT performance in children likely results from a

combination of several physiological and psychological aspects, including motivation [35].

The importance of motivation is further supported by our study which identified the motivation and confidence domain as having moderate to large effect sizes across sex and age groups. This is noteworthy given that the association range for the motivation and confidence domain was stronger than the overall daily behavior domain across most sex by age groups. Thus, our findings highlight that creating a motivational climate (e.g., mastery climate – achievement goal theory, self-determination theory, etc.) that fosters increased effort and task motivation could positively impact CRF, and the context in which children are active (e.g., encouragement, reinforcement, etc.) may be more important than simply getting children moving.

Percentage of children with healthy CRF

The Ruiz standards of 42 and 35 ml/kg/min for boys and girls, respectively, have recently been proposed as an international interim standard to identify children at risk of poor health [16, 26]. Although international percentages of healthy CRF in children have been estimated elsewhere [26], no study to date has applied these standards to a large sample of children. Our results suggest that almost all Canadian girls meet the standard for healthy CRF. Although percentages obtained for girls could be correct, it is more likely that our results reflect a limitation of using one single standard for all age groups, and perhaps the standard for Canadian girls is too low. Previous research has identified that predicted relative $\text{VO}_{2\text{peak}}$ using the 20mSRT for boys increases slightly with age and

maturation, but these values tend to decrease substantially with age and maturation in girls [28]. For this reason, previous research has suggested that the health-related CRF standards for young girls should be higher (i.e., 38 ml/kg/min) [29]. Indeed, our results highlight the need to develop age specific standards for boys and girls, which could be informed using the recently published international normative-referenced standards, developed using a sample of 1,142,026 children and youth from 50 countries [18].

Strategies for CRF and PL surveillance

Our study identified strong associations between CRF and PL that generally increased with age. This finding suggests that 20mSRT performance could significantly differentiate children with high, medium, and low PL levels. Given this finding, the 20mSRT could be used as a simple screening assessment to help identify children with low PL and in need of a full CAPL assessment, providing a way to save time and resources in the school-based setting. This strategy is consistent with other studies that have advocated for the use of the 20mSRT as a population health surveillance instrument to help identify children and youth at risk of poor health outcomes [9]. This study also highlights the possibility of developing new PL-based standards for children that would incorporate a variety of cognitive, mental, and physical health indicators that may help screen children at increased risk of lifestyle-related disease. Future studies should aim to better understand the sensitivity and specificity of using CRF to predict PL.

Strengths and limitations

An important strength of this study is the large sample of Canadian children tested using a standardized data collection protocol, validated CAPL instruments, and trained assessors across all 11 sites. We collected data during a relatively small timeframe (i.e., 3 years), thereby limiting the possible effect of temporal trends. We also aimed to obtain a balanced sample across Eastern and Western Canada to diminish any potential cross-country differences in our estimates. However, our results need to be interpreted in light of the following limitations. First, this was a cross-sectional study and therefore causality cannot be inferred. Second, in light of the voluntary and purposive sampling strategy, the results may not be generalizable to the broader Canadian context, although our physical competence results are comparable to previous studies of nationally representative Canadian samples of similarly aged children [36]. Finally, the potential confounding effects of unmeasured variables (e.g., maturation, socioeconomic status, etc.) cannot be discounted.

CONCLUSION

Collectively, this study identified strong favourable associations between CRF and all components of PL among school-aged children. These findings suggest that CRF could be used as a simple screening assessment to identify children at risk of low PL levels. Furthermore, this study identified higher percentages of girls with healthy CRF, in comparison with boys. Future studies should aim to replicate these results in different countries and age groups and to further refine criterion standards for CRF among children.

Ethics approval and consent to participate: Obtained from the Children's Hospital of Eastern Ontario Research Ethics Board.

Availability of data and materials: The datasets used and/or analysed during the current study are available from Dr. Mark Tremblay (mtremblay@cheo.on.ca) on reasonable requests.

Competing interests: No competing interests declared.

Funding: This study was funded by RBC, the Public Health Agency of Canada, Mitacs, and was delivered in partnership with ParticipACTION.

Author's contributions: MST and PEL were responsible for the RBC-CAPL study design. JJL and MST developed the research questions and objectives for this study. JJL, MST and JPC led the data analysis and synthesis of results. JJL wrote the manuscript. All other authors contributed to data collection, editing, reviewing and approved of the final manuscript.

Acknowledgements: We would like to acknowledge all the research assistants that helped gather data for this study.

References

1. Lloyd M, Colley RC, Tremblay MS. Advancing the debate on ‘fitness testing’ for children: Perhaps we’re riding the wrong animal. *Pediatr Exerc Sci*. 2010;22(2):176-82.
2. Whitehead M. Definition of physical literacy and clarification of related issues. *ICSSPE Bulletin*. 2013;65(1.2).
3. Longmuir PE, Boyer C, Lloyd M, Yang Y, Boiarskaia E, Zhu W, et al. The Canadian Assessment of Physical Literacy: methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Health*. 2015;15:767.
4. Francis CE, Longmuir PE, Boyer C, Andersen LB, Barnes JD, Boiarkaia E, et al. The Canadian Assessment of Physical Literacy: Development of a Model of Children’s Capacity for a Healthy, Active Lifestyle through a Delphi Process. *J Phys Act Health*. 2016;13(2):214-22.
5. Léger L, Lambert J, Goulet A, Rowan C, Dinelle Y. Capacité aérobie des Québécois de 6 à 17 ans – test navette de 20 mètre avec paliers de 1 minute. *Can J Appl Sport Sci*. 1984;9:64-9.
6. Léger 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.
7. Mayorga-Vega M, Aguilar-Soto P, Viciano J. Criterion-related validity of the 20-M shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *J Sports Sci Med*. 2015; 14:536-47.
8. Ortega FB, Artero EG, Ruiz JR, Vicente-Rodriguez G, Bergman P, Hagströmer M, et al. Reliability of health-related physical fitness tests in European adolescents. The HELENA Study. *Int J Obes (Lond)*. 2008;32(Suppl 5):S49-57.
9. Lang JJ, Tremblay MS, Léger L, Olds TS, Tomkinson GR. International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results. *Br J Sports Med*. Published online first: 20 Sept 2016.
10. Dumith SC, van Dusen D, Kohl HW. Physical fitness measures among children and adolescents: Are they all necessary? *J Sports Med Phys Fitness*. 2012;52(2):181–9.
11. Kristensen PL, Moeller NC, Korsholm L, Kolle E, Wedderkopp N, Froberg K, et al. The association between aerobic fitness and physical activity in children and adolescents: The European Youth Heart Study. *Eur J Appl Physiol*. 2010;110:267–75.
12. Williams SM, Phelps D, Laurson KR, Thomas DQ, Brown DD. Fitness knowledge, cardiorespiratory endurance and body composition of high school students. *Biomed Hum Kinetics*. 2013;5(1):17-21.
13. Sardinha LB, Marques A, Martins S, Palmeira A, Minderico C. Fitness, fatness, and academic performance in seventh-grade elementary school students. *BMC Pediatr*. 2014;14:176.

14. Colquitt G, Walker A, Langdon JL, McCollum S, Pomazal M. Exploring student attitudes toward physical education and implications for policy. *Sport Scientific & Practical Aspects*. 2012;9(2):5-12.
15. Goa Z, Lodewyk KR, Zhang T. The role of ability beliefs and incentives in middle school students' intention, cardiovascular fitness, and effort. *J Teach Phys Educ*. 2009;28:3-20.
16. Ruiz JR, Caverio-Redendo I, Ortega FB, Welk GJ, Andersen LB, Martinez-Vizcaino V. 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*. Published online first: 26 September 2016.
17. McClain JJ, Welk GJ, Ihmels M, Schaben J. Comparison of two versions of the PACER aerobic fitness test. *J Phys Act Health*. 2006;3(Suppl 2):S47-S57.
18. Tomkinson GR, Lang JJ, Tremblay MS, Dale M, LeBlanc AG, Belanger K, et al. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. *Br J Sports Med*. Published Online First: 23 April 2016.
19. Tremblay MS, Shields M, Laviolette M, Craig CL, Janssen I, Gorber SC. Fitness of Canadian children and youth: results from the 2007–2009 Canadian Health Measures Survey. *Health Rep*. 2010;21(1):7-20.
20. Boyer C, Tremblay MS, Saunders TJ, McFarlane A, Borghese M, Lloyd M, Longmuir P. Feasibility, validity, and reliability of the plank isometric holds as a field-based assessment of torso muscular endurance for children 8 to 12 years of age. *Pediatr Exerc Sci*. 2013;25(3):407-22.
21. Longmuir PE, Boyer C, Lloyd M, Andersen LB, Barnes JD, Boiarskaia E, et al. Canadian Agility and Movement Skill Assessment (CAMSA): Validity, objectivity, and reliability evidence for children 8– 12 years of age. *J Sport Health Sci*. Available online first: 11 November 2015.
22. Saunders TJ, Gray CE, Borghese MM, McFarlane A, Mbonu A, Ferraro ZM, et al. Validity of SC-StepRx pedometer-derived moderate and vigorous physical activity during treadmill walking and running in a heterogeneous sample of children and youth. *BMC Public Health*. 2014;14(1):1–9.
23. Tudor-Locke C, McClain JJ, Hart TL, Sisson SB, Washington TL. Expected values for pedometer-determined physical activity in youth. *Res Q Exerc Sport*. 2009; 80(2):164–74.
24. Garcia AW, Broda MA, Frenn M, Coviak C, Pender NJ, Ronis DL. Gender and developmental differences in exercise beliefs among youth and prediction of their exercise behavior. *J Sch Health*. 1995; 65(6):213-9.
25. Hay JA. Adequacy in and predilection for physical activity in children. *Clin J Sports Med*. 1992;2:192-201.
26. Lang JJ, Tremblay MS, Ortega FB, Ruiz JR, Tomkinson GR. Review of criterion-referenced standards for cardiorespiratory fitness: what percentage of 1 142 026

international children and youth are apparently healthy? *Br J Sports Med*. Published Online First: 2 Jan 2017.

27. Cohen J. Statistical power analysis for the behavioral sciences. 2nd edn. New Jersey. Lawrence Erlbaum, 1988.
28. Mota J, Guerra S, Leandro C, Pinto A, Ribeiro JC, Duarte JA. Association of maturation, sex, and body fat in cardiorespiratory fitness. *Am J Hum Biol*. 2002;14:707-12.
29. Morrow JR, Falls HB, Kohl HW, editors. The prudential fitnessgram: technical reference manual. Dallas: Cooper Institute for Aerobic Research; 1994.
30. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes*. 2008;32(1):1-11.
31. Burns RD, Brusseau TA, Fu Y, Hannon JC. Predictors and trends of gross motor skill performance in at-risk elementary school-aged children. *Percept Mot Skills*. 2015;121(1):284-99.
32. Welk GJ, Laurson KR, Eisenmann JC, Cureton K. Development of youth aerobic-capacity standards using receiver operating characteristic curves. *Am J Prev Med*. 2011;41(4S2):S111-S116.
33. Coe DP, Peterson T, Blair C, Schutten MC, Peddie H. Physical fitness, academic achievement, and socioeconomic status in school-aged youth. *J School Health*. 2013;83:500-7.
34. Olds TS, Dollman J. Are changes in distance-run performance of Australian children between 1985 and 1997 explained by changes in fatness? *Pediatr Exerc Sci*. 2004;16:201-9.
35. Tomkinson GR, Olds TS. Field tests of fitness. In: Armstrong N, van Mechelen W, editors. *Paediatric exercise science and medicine*. New York, NY: Oxford University Press; 2000. p. 109-28.
36. Tremblay MS, Shields M, Laviolette M, Craig CL, Janssen I, Gorber SC. Fitness of Canadian children and youth: results from the 2007-2009 Canadian Health Measures Survey. *Health Reports*. 2010;21(1):82-003-XPE.

Table 1 Descriptive characteristics of participants.

	Boys (n=4710)	Girls (n=4683)	Total (n=9393)
<i>Age (years)</i>	10.1 (1.2)	10.1 (1.2)	10.1 (1.2)
<i>Site (n, %)</i>			
Victoria, British Colombia	258 (5.5)	220 (4.7)	478 (5.1)
Lethbridge, Alberta	532 (11.3)	531 (11.3)	1063 (11.3)
Calgary, Alberta	620 (13.2)	619 (13.2)	1239 (13.2)
Winnipeg, Manitoba	607 (12.9)	610 (13.0)	1217 (13.0)
North Bay, Ontario	529 (11.2)	578 (12.3)	1107 (11.8)
Windsor, Ontario	637 (13.5)	569 (12.2)	1206 (12.8)
Ottawa, Ontario	350 (7.4)	359 (7.7)	709 (7.5)
Trois-Rivières, Québec	56 (1.2)	34 (0.7)	90 (1.0)
Halifax, Nova Scotia	390 (8.3)	407 (8.7)	797 (8.5)
Antigonish, Nova Scotia	482 (10.2)	506 (10.8)	988 (10.5)
Charlottetown, P.E.I.	249 (5.3)	250 (5.3)	499 (5.3)
Independent variable			
20mSRT (# laps)	25.8 (15.8)	20.9 (11.6)	23.4 (14.1)
Running speed at the last completed stage (km/h) ^a			
8 year-olds	8.9 (2.1)	8.8 (1.6)	8.8 (1.8)
9 year-olds	9.1 (1.6)	8.9 (1.6)	9.0 (1.6)
10 year-olds	9.2 (1.5)	9.0 (1.3)	9.1 (1.4)
11 year-olds	9.4 (1.5)	9.2 (1.2)	9.3 (1.3)
12 year-olds	9.5 (1.6)	9.3 (1.2)	9.4 (1.4)
Dependent variables			
<i>Physical literacy</i>			
Total CAPL score (0-100)	63.9 (12.4)	63.7 (10.9)	63.8 (11.7)
<i>Physical competence domain</i>			
Total domain score (0-32)	20.8 (4.3)	20.8 (4.2)	20.8 (4.2)
Sit and reach score (cm)	25.4 (7.6)	30.9 (8.3)	28.1 (8.4)
Handgrip score (kg)	34.5 (9.6)	32.6 (9.2)	33.6 (9.5)
Plank score (sec)	62.0 (44.2)	61.3 (42.7)	61.5 (43.5)
Body mass index (kg/m ²)	18.9 (3.9)	19.0 (3.8)	19.0 (3.8)
Waist circumference (cm)	67.4 (11.0)	67.2 (10.6)	67.3 (10.8)
CAMSA score	31.5 (5.9)	30.4 (5.7)	31.0 (5.8)
<i>Daily behavior domain</i>			
Total domain score (0-32)	18.6 (7.9)	18.4 (7.4)	18.5 (7.6)
Average daily step counts	12405 (4081)	10793 (3508)	11530 (3865)
Self-reported screen time (h)	2.7 (2.1)	2.2 (1.8)	2.5 (1.9)
Self-reported physical activity (days meeting the guidelines)	5.0 (2.0)	4.9 (1.9)	5.0 (1.9)
<i>Motivation and confidence domain</i>			

Total domain score (0-18)	12.7 (2.8)	12.2 (2.6)	12.5 (2.7)
CSAPPA adequacy score	4.8 (0.9)	4.6 (0.9)	4.7 (0.9)
CSAPPA predilection score	4.8 (1.0)	4.8 (1.0)	4.8 (1.0)
Benefits and barriers ratio	1.6 (1.2)	1.5 (1.1)	1.6 (1.2)
Activity levels compared to peers	0.7 (0.2)	0.7 (0.2)	0.7 (0.2)
Skill level compared to peers	0.7 (0.2)	0.6 (0.2)	0.7 (0.2)
<i>Knowledge and understanding domain</i>			
Total domain score (0-18)	11.8 (2.8)	12.2 (2.6)	12.0 (2.7)
CAPL questionnaire score	11.8 (2.8)	12.2 (2.6)	12.0 (2.7)

All variables are presented as means (SD) unless otherwise stated.

20mSRT = 20-m shuttle run test; CRF = cardiorespiratory fitness; CAMSA = Canadian Agility and Movement Skill Assessment; CSAPPA = Children's Self-Perception and Adequacy in and Predilection for Physical Activity; CAPL = Canadian Assessment of Physical Literacy.

^a 20mSRT results are reported as running speed (km/h) at the last completed stage for each age group as recommended by Tomkinson et al. [18].

Table 2 Tertiles of cardiorespiratory fitness across components of physical literacy for 8 year-old boys and girls.

	Boys (n=522)				Girls (n=524)			
	1-low CRF 9.4 (2.7)	2-medium CRF 19.5 (3.6)	3-high CRF 39.5 (11.2)	Cohen's d	1-low CRF 9.5 (2.1)	2-medium CRF 15.9 (1.9)	3-high CRF 31.1 (10.6)	Cohen's d
20mSRT (# laps)								
Physical literacy								
Total CAPL score	55.1 (10.8)	62.1 (10.3) ^a	67.7 (9.8) ^{bc*}	1.22	58.7 (9.9)	62.7 (9.9) ^a	67.1 (8.3) ^{bc*}	0.92
Physical competence domain								
Total domain score	17.5 (3.4)	20.0 (3.2) ^a	22.4 (3.0) ^{bc*}	1.53	17.4 (3.8)	19.8 (3.2) ^a	21.3 (3.2) ^{bc*}	1.11
Sit and reach score (cm)	26.1 (6.7)	27.1 (7.0)	28.1 (6.8) ^{c*}	0.30	29.8 (7.0)	32.0 (6.5) ^a	32.1 (7.5) ^{c*}	0.32
Handgrip score (kg)	25.5 (6.8)	26.8 (5.6)	29.3 (5.6) ^{bc*}	0.61	24.5 (5.9)	24.9 (5.4)	26.8 (5.3) ^{bc*}	0.41
Plank score (sec)	37.6 (27.2)	51.9 (34.5) ^a	72.0 (41.5) ^{bc*}	0.98	39.3 (21.6)	54.9 (37.3) ^a	74.0 (50.6) ^{bc*}	0.89
BMI (kg/m ²)	18.1 (3.4)	16.9 (2.3) ^a	16.6 (1.9) ^{c*}	0.55	18.3 (3.5)	16.9 (2.5) ^a	16.7 (2.5) ^{c*}	0.53
Waist circumference (cm)	63.2 (9.0)	60.8 (6.9) ^a	59.0 (4.4) ^{bc*}	0.59	64.4 (9.8)	60.8 (7.7) ^a	59.2 (6.6) ^{c*}	0.62
CAMSA score	23.7 (6.0)	28.5 (5.3) ^a	31.2 (5.1) ^{bc*}	1.35	23.5 (5.7)	26.3 (5.7) ^a	28.9 (5.4) ^{bc*}	0.97
Daily behavior domain								
Total domain score	16.3 (7.9)	19.3 (7.3) ^a	21.6 (6.9) ^{bc*}	0.71	18.8 (7.4)	20.2 (7.1)	21.9 (6.4) ^{c*}	0.45
Average daily step counts	11369 (3786)	13589 (3907) ^a	14218 (3471) ^{c*}	0.78	11270 (3479)	11773 (3242)	11758 (3283)	0.14
Self-reported screen time (h/day)	3.2 (2.4)	2.4 (1.8) ^a	2.1 (2.0) ^{c*}	0.50	2.2 (2.0)	1.9 (1.8)	1.6 (1.5) ^{c*}	0.34
Average days meeting the guidelines	4.4 (2.3)	4.7 (2.4)	5.4 (1.9) ^{bc*}	0.47	4.5 (2.3)	4.8 (2.1)	5.2 (1.8) ^{c*}	0.34
Motivation and confidence domain								
Total domain score	11.5 (2.8)	12.3 (2.5) ^a	13.2 (2.3) ^{bc*}	0.66	12.0 (2.3)	11.9 (2.5)	13.2 (2.1) ^{bc*}	0.54
CSAPPA adequacy score	4.5 (0.9)	4.6 (0.8)	4.9 (0.8) ^{bc*}	0.47	4.5 (0.9)	4.5 (0.8)	4.9 (0.8) ^{bc*}	0.47
CSAPPA predilection score	4.3 (1.0)	4.8 (0.9) ^a	5.0 (0.9) ^{c*}	0.74	4.6 (0.9)	4.8 (0.9)	5.2 (0.7) ^{bc*}	0.74
Benefits and barriers ratio	1.3 (1.4)	1.5 (1.3)	1.8 (1.2) ^{c*}	0.38	1.4 (1.2)	1.3 (1.2)	1.6 (1.1) ^c	0.17
Activity levels compared to peers	0.7 (0.3)	0.7 (0.2)	0.8 (0.2) ^{c*}	0.39	0.7 (0.2)	0.7 (0.2)	0.8 (0.2) ^{b*}	0.50
Skill level compared to peers	0.7 (0.3)	0.7 (0.3)	0.8 (0.2) ^{bc*}	0.39	0.7 (0.2)	0.7 (0.2)	0.7 (0.2)	0.00
Knowledge and understanding domain								
Total domain score	9.8 (2.7)	10.5 (3.0)	10.4 (2.6)	0.23	10.5 (2.5)	10.7 (2.6)	10.7 (2.6)	0.08
CAPL questionnaire score	9.8 (2.7)	10.5 (3.0)	10.4 (2.6)	0.23	10.5 (2.5)	10.7 (2.6)	10.7 (2.6)	0.08

All variables are presented as means (SD). Statistical significance for the main effect was assessed using ANOVA. Bonferroni (Dunn) t-test was used to assess differences between groups. Cohen's *d* was used to calculate the effect size between the low and high CRF groups. CRF = cardiorespiratory fitness; 20mSRT = 20-m shuttle run test; BMI = body mass index; CSAPPA = Children's Self-Perception and Adequacy in and Predilection for Physical Activity; CAPL = Canadian Assessment of Physical Literacy.

* $p < 0.001$ for main effect; ^a $p < 0.05$ for 2 vs 1; ^b $p < 0.05$ for 3 vs 2; ^c $p < 0.05$ for 3 vs 1.

Table 3 Tertiles of cardiorespiratory fitness across components of physical literacy for 9 year-old boys and girls.

	Boys (n=946)				Girls (n=908)			
	1-low CRF 10.7 (3.0)	2-medium CRF 21.0 (3.5)	3-high CRF 40.9 (10.7)	Cohen's d	1-low CRF 9.6 (2.3)	2-medium CRF 17.2 (2.5)	3-high CRF 32.3 (8.5)	Cohen's d
20mSRT (# laps)								
Physical literacy								
Total CAPL score	57.3 (11.6)	63.9 (10.1) ^a	70.6 (9.7) ^{bc*}	1.24	59.4 (9.9)	63.4 (10.1) ^a	68.4 (8.4) ^{bc*}	0.98
Physical competence domain								
Total domain score	17.5 (3.7)	20.4 (3.4) ^a	22.9 (3.4) ^{bc*}	1.52	17.6 (3.8)	20.3 (3.5) ^a	22.7 (3.5) ^{bc*}	1.40
Sit and reach score (cm)	25.0 (7.9)	26.1 (7.3)	28.0 (7.0) ^{bc*}	0.40	29.5 (7.7)	31.4 (7.9) ^a	32.1 (7.7) ^{c*}	0.34
Handgrip score (kg)	29.9 (7.5)	30.3 (7.3)	32.1 (6.9) ^{bc*}	0.31	27.7 (6.8)	28.5 (6.6)	29.3 (6.3) ^{c*}	0.24
Plank score (sec)	41.8 (31.0)	55.8 (36.8) ^a	83.1 (53.7) ^{bc*}	0.94	44.4 (30.0)	60.1 (41.8) ^a	80.7 (51.3) ^{bc*}	0.86
BMI (kg/m ²)	20.1 (4.3)	17.7 (2.6) ^a	16.9 (2.0) ^{bc*}	0.95	19.9 (4.1)	18.2 (3.1) ^a	17.1 (2.6) ^{bc*}	0.82
Waist circumference (cm)	68.7 (12.0)	63.3 (7.5) ^a	60.6 (5.4) ^{bc*}	0.87	68.8 (10.7)	64.3 (9.4) ^a	60.4 (7.3) ^{bc*}	0.92
CAMSA score	26.8 (5.5)	29.9 (4.7) ^a	33.5 (4.2) ^{bc*}	1.37	25.9 (5.7)	29.4 (5.0) ^a	32.0 (4.5) ^{bc*}	1.19
Daily behavior domain								
Total domain score	17.3 (8.0)	19.8 (7.5) ^a	22.6 (6.7) ^{bc*}	0.72	18.9 (7.2)	19.2 (7.3)	20.9 (6.3) ^{bc*}	0.25
Average daily step counts	11745 (3751)	12697 (3805)	14632 (3982) ^{bc*}	0.75	10476 (3486)	11135 (3386)	11993 (3380) ^{bc*}	0.44
Self-reported screen time (h/day)	3.1 (2.6)	2.5 (1.8) ^a	2.0 (1.7) ^{bc*}	0.50	2.1 (1.9)	2.0 (1.7)	1.7 (1.5) ^{c*}	0.23
Average days meeting the guidelines	4.5 (2.2)	5.0 (2.1) ^a	5.5 (1.8) ^{bc*}	0.50	4.7 (2.1)	4.9 (2.0)	5.3 (1.6) ^{bc*}	0.32
Motivation and confidence domain								
Total domain score	11.9 (2.8)	12.6 (2.4) ^a	13.8 (2.1) ^{bc*}	0.77	11.7 (2.6)	12.5 (2.2) ^a	13.2 (2.0) ^{bc*}	0.65
CSAPPA adequacy score	4.5 (0.9)	4.8 (0.8) ^a	5.1 (0.7) ^{bc*}	0.74	4.3 (0.9)	4.6 (0.8) ^a	4.9 (0.8) ^{bc*}	0.71
CSAPPA predilection score	4.6 (1.0)	4.8 (0.9) ^a	5.2 (0.8) ^{bc*}	0.66	4.6 (1.0)	4.8 (0.9)	5.2 (0.8) ^{bc*}	0.66
Benefits and barriers ratio	1.5 (1.3)	1.6 (1.2)	1.8 (1.2) ^{c*}	0.24	1.5 (1.2)	1.6 (1.0)	1.7 (1.0)	0.18
Activity levels compared to peers	0.7 (0.2)	0.7 (0.2) ^a	0.8 (0.2) ^{bc*}	0.50	0.7 (0.2)	0.7 (0.2) ^a	0.8 (0.2) ^{bc*}	0.50
Skill level compared to peers	0.7 (0.3)	0.7 (0.2)	0.8 (0.2) ^{bc*}	0.39	0.6 (0.2)	0.7 (0.2) ^a	0.7 (0.2) ^{c*}	0.50
Knowledge and understanding domain								
Total domain score	10.6 (2.8)	11.0 (2.7)	11.3 (2.6) ^{c*}	0.26	11.1 (2.5)	11.4 (2.5)	11.6 (2.3) ^c	0.21
CAPL questionnaire score	10.6 (2.8)	11.0 (2.7)	11.3 (2.6) ^{c*}	0.26	11.1 (2.5)	11.4 (2.5)	11.6 (2.3) ^c	0.21

All variables are presented as means (SD). Statistical significance for the main effect was assessed using ANOVA. Bonferroni (Dunn) t-test was used to assess differences between groups. Cohen's *d* was used to calculate the effect size between the low and high CRF groups. CRF = cardiorespiratory fitness; 20mSRT = 20-m shuttle run test; BMI = body mass index; CSAPPA = Children's Self-Perception and Adequacy in and Predilection for Physical Activity; CAPL = Canadian Assessment of Physical Literacy.

* $p < 0.001$ for main effect; ^a $p < 0.05$ for 2 vs 1; ^b $p < 0.05$ for 3 vs 2; ^c $p < 0.05$ for 3 vs 1.

Table 4 Tertiles of cardiorespiratory fitness across components of physical literacy for 10 year-old boys and girls.

	Boys (n=1201)				Girls (n=1175)			
	1-low CRF 11.0 (3.0)	2-medium CRF 21.4 (3.9)	3-high CRF 42.6 (10.0)	Cohen's d	1-low CRF 10.0 (2.1)	2-medium CRF 16.7 (2.1)	3-high CRF 32.0 (9.9)	Cohen's d
20mSRT (# laps)								
Physical literacy								
Total CAPL score	56.4 (11.4)	63.6 (10.5) ^a	71.9 (9.2) ^{bc*}	1.50	57.3 (10.2)	63.4 (9.7) ^a	69.6 (9.9) ^{bc*}	1.22
Physical competence domain								
Total domain score	17.8 (3.8)	20.5 (3.5) ^a	23.7 (3.4) ^{bc*}	1.64	17.9 (3.8)	20.4 (3.4) ^a	23.1 (3.3) ^{bc*}	1.46
Sit and reach score (cm)	24.4 (7.5)	25.7 (7.4) ^a	26.4 (7.4) ^{c*}	0.27	28.6 (7.7)	30.1 (8.0) ^a	32.4 (8.5) ^{bc*}	0.47
Handgrip score (kg)	31.4 (7.9)	32.8 (8.2)	35.7 (7.6) ^{bc*}	0.56	30.1 (7.7)	31.2 (7.7)	32.8 (7.3) ^{bc*}	0.36
Plank score (sec)	40.7 (31.7)	56.8 (35.3) ^a	85.7 (51.5) ^{bc*}	1.05	40.5 (28.8)	55.7 (32.6) ^a	80.1 (46.2) ^{bc*}	1.03
BMI (kg/m ²)	20.6 (4.7)	18.7 (3.2) ^a	17.4 (2.2) ^{bc*}	0.87	20.7 (4.5)	18.8 (3.2) ^a	17.6 (2.8) ^{bc*}	0.83
Waist circumference (cm)	72.4 (12.9)	66.6 (9.1) ^a	62.8 (6.5) ^{bc*}	0.94	72.2 (12.3)	66.5 (8.8) ^a	63.5 (7.7) ^{bc*}	0.85
CAMSA score	27.8 (5.8)	31.5 (4.5) ^a	34.7 (4.0) ^{bc*}	1.39	27.4 (5.3)	29.7 (4.8) ^a	32.9 (4.6) ^{bc*}	1.11
Daily behavior domain								
Total domain score	15.7 (7.6)	18.3 (7.4) ^a	21.5 (6.7) ^{bc*}	0.81	16.7 (7.1)	18.4 (6.9) ^a	20.5 (7.1) ^{bc*}	0.54
Average daily step counts	10923 (3696)	12327 (3935) ^a	13907 (4026) ^{bc*}	0.77	10065 (3126)	10771 (3825)	11569 (3668) ^{bc*}	0.44
Self-reported screen time (h/day)	3.3 (2.3)	2.7 (1.9) ^a	2.2 (1.6) ^{bc*}	0.56	2.4 (1.9)	1.9 (1.5) ^a	1.9 (1.5) ^{c*}	0.29
Average days meeting the guidelines	4.5 (2.1)	5.1 (1.9) ^a	5.8 (1.5) ^{bc*}	0.71	4.5 (1.9)	5.0 (1.9) ^a	5.3 (1.5) ^{c*}	0.48
Motivation and confidence domain								
Total domain score	11.8 (2.9)	12.8 (2.6) ^a	14.2 (2.1) ^{bc*}	0.95	11.2 (2.9)	12.2 (2.5) ^a	13.2 (2.2) ^{bc*}	0.78
CSAPPA adequacy score	4.5 (1.0)	4.9 (0.9) ^a	5.2 (0.7) ^{bc*}	0.81	4.2 (1.0)	4.6 (0.9) ^a	4.9 (0.8) ^{bc*}	0.77
CSAPPA predilection score	4.5 (1.1)	4.9 (0.9) ^a	5.3 (0.7) ^{bc*}	0.87	4.5 (1.1)	4.8 (1.0) ^a	5.2 (0.8) ^{bc*}	0.83
Benefits and barriers ratio	1.4 (1.3)	1.6 (1.1) ^a	2.1 (1.1) ^{bc*}	0.58	1.3 (1.2)	1.4 (1.1)	1.7 (1.0) ^{bc*}	0.36
Activity levels compared to peers	0.7 (0.2)	0.7 (0.2) ^a	0.8 (0.2) ^{bc*}	0.50	0.6 (0.2)	0.7 (0.2) ^a	0.7 (0.2) ^{bc*}	0.50
Skill level compared to peers	0.6 (0.3)	0.7 (0.2) ^a	0.8 (0.2) ^{bc*}	0.78	0.6 (0.2)	0.6 (0.2) ^a	0.7 (0.2) ^{bc*}	0.50
Knowledge and understanding domain								
Total domain score	11.4 (2.8)	12.0 (2.6) ^a	12.5 (2.6) ^{bc*}	0.41	11.6 (2.7)	12.4 (2.4) ^a	12.8 (2.4) ^{c*}	0.47
CAPL questionnaire score	11.4 (2.8)	12.0 (2.6) ^a	12.5 (2.6) ^{bc*}	0.41	11.6 (2.7)	12.4 (2.4) ^a	12.8 (2.4) ^{c*}	0.47

All variables are presented as means (SD). Statistical significance for the main effect was assessed using ANOVA. Bonferroni (Dunn) t-test was used to assess differences between groups. Cohen's *d* was used to calculate the effect size between the low and high CRF groups. CRF = cardiorespiratory fitness; 20mSRT = 20-m shuttle run test; BMI = body mass index; CSAPPA = Children's Self-Perception and Adequacy in and Predilection for Physical Activity; CAPL = Canadian Assessment of Physical Literacy.

* $p < 0.001$ for main effect; ^a $p < 0.05$ for 2 vs 1; ^b $p < 0.05$ for 3 vs 2; ^c $p < 0.05$ for 3 vs 1.

Table 5 Tertiles of cardiorespiratory fitness across components of physical literacy for 11 year-old boys and girls.

	Boys (n=1510)				Girls (n=1550)			
	1-low CRF 11.5 (3.3)	2-medium CRF 24.0 (4.6)	3-high CRF 47.6 (12.2)	Cohen's d	1-low CRF 11.6 (2.8)	2-medium CRF 19.6 (2.6)	3-high CRF 35.6 (9.2)	Cohen's d
20mSRT (# laps)								
Physical literacy								
Total CAPL score	56.4 (11.4)	65.1 (11.0) ^a	72.2 (10.3) ^{bc*}	1.45	58.0 (10.3)	63.5 (9.7) ^a	70.3 (9.9) ^{bc*}	1.22
Physical competence domain								
Total domain score	18.1 (3.8)	21.6 (3.5) ^a	24.4 (3.3) ^{bc*}	1.77	19.0 (3.6)	21.6 (3.3) ^a	24.0 (3.4) ^{bc*}	1.43
Sit and reach score (cm)	23.0 (7.6)	25.2 (7.6) ^a	25.8 (7.3) ^{c*}	0.38	28.6 (8.6)	30.4 (8.7) ^a	33.1 (8.6) ^{bc*}	0.52
Handgrip score (kg)	36.0 (9.8)	37.0 (9.1)	39.7 (8.5) ^{bc*}	0.40	34.6 (9.2)	35.5 (9.0)	37.2 (8.6) ^{bc*}	0.29
Plank score (sec)	42.2 (31.9)	63.0 (39.1) ^a	93.2 (53.1) ^{bc*}	1.16	43.3 (27.9)	60.0 (37.5) ^a	85.5 (48.5) ^{bc*}	1.07
BMI (kg/m ²)	21.8 (5.0)	19.2 (3.3) ^a	17.8 (2.4) ^{bc*}	1.01	21.1 (4.5)	19.3 (3.3) ^a	18.0 (2.7) ^{bc*}	0.84
Waist circumference (cm)	76.7 (14.1)	69.1 (9.4) ^a	65.3 (6.7) ^{bc*}	1.31	74.1 (12.3)	68.5 (9.3) ^a	64.8 (7.1) ^{bc*}	0.93
CAMSA score	29.7 (5.4)	33.5 (4.5) ^a	36.1 (3.6) ^{bc*}	1.40	29.5 (5.1)	32.0 (4.4) ^a	34.7 (4.1) ^{bc*}	1.12
Daily behavior domain								
Total domain score	15.2 (7.8)	18.2 (7.5) ^a	20.5 (7.6) ^{bc*}	0.69	15.6 (7.7)	17.2 (7.0) ^a	19.6 (7.1) ^{bc*}	0.54
Average daily step counts	10984 (3740)	11825 (3865) ^a	13209 (4118) ^{bc*}	0.57	9679 (3282)	10176 (3266)	11587 (3777) ^{bc*}	0.54
Self-reported screen time (h/day)	3.4 (2.3)	2.8 (1.9) ^a	2.4 (1.7) ^{bc*}	0.50	2.7 (2.0)	2.4 (1.7) ^a	2.0 (1.5) ^{bc*}	0.40
Average days meeting the guidelines	4.4 (2.1)	5.1 (1.8) ^a	5.7 (1.6) ^{bc*}	0.70	4.4 (1.9)	4.9 (1.7) ^a	5.3 (1.6) ^{bc*}	0.51
Motivation and confidence domain								
Total domain score	11.3 (3.1)	12.7 (2.8) ^a	14.1 (2.1) ^{bc*}	1.06	11.1 (2.6)	11.9 (2.4) ^a	13.3 (2.4) ^{bc*}	0.88
CSAPPA adequacy score	4.4 (1.0)	4.8 (0.9) ^a	5.2 (0.7) ^{bc*}	0.93	4.2 (0.9)	4.5 (0.8) ^a	4.9 (0.8) ^{bc*}	0.82
CSAPPA predilection score	4.4 (1.1)	4.8 (1.0) ^a	5.3 (0.8) ^{bc*}	0.94	4.4 (1.0)	4.6 (0.9) ^a	5.1 (0.8) ^{bc*}	0.77
Benefits and barriers ratio	1.3 (1.2)	1.7 (1.2) ^a	2.0 (1.1) ^{bc*}	0.61	1.3 (1.1)	1.5 (1.0) ^a	1.8 (1.0) ^{bc*}	0.48
Activity levels compared to peers	0.6 (0.2)	0.7 (0.2) ^a	0.8 (0.2) ^{bc*}	1.00	0.6 (0.2)	0.7 (0.2) ^a	0.8 (0.2) ^{bc*}	1.00
Skill level compared to peers	0.6 (0.2)	0.7 (0.2) ^a	0.8 (0.2) ^{bc*}	1.00	0.6 (0.2)	0.6 (0.2) ^a	0.7 (0.2) ^{bc*}	0.50
Knowledge and understanding domain								
Total domain score	11.8 (2.8)	12.7 (2.6) ^a	13.1 (2.5) ^{c*}	0.49	12.2 (2.6)	12.8 (2.4) ^a	13.4 (2.4) ^{bc*}	0.48
CAPL questionnaire score	11.8 (2.8)	12.7 (2.6) ^a	13.1 (2.5) ^{c*}	0.49	12.2 (2.6)	12.8 (2.4) ^a	13.4 (2.4) ^{bc*}	0.48

All variables are presented as means (SD). Statistical significance for the main effect was assessed using ANOVA. Bonferroni (Dunn) t-test was used to assess differences between groups. Cohen's *d* was used to calculate the effect size between the low and high CRF groups. CRF = cardiorespiratory fitness; 20mSRT = 20-m shuttle run test; BMI = body mass index; CSAPPA = Children's Self-Perception and Adequacy in and Predilection for Physical Activity; CAPL = Canadian Assessment of Physical Literacy.

* $p < 0.001$ for main effect; ^a $p < 0.05$ for 2 vs 1; ^b $p < 0.05$ for 3 vs 2; ^c $p < 0.05$ for 3 vs 1.

Table 6 Tertiles of cardiorespiratory fitness across components of physical literacy for 12 year-old boys and girls.

	Boys (n=531)				Girls (n=526)			
	1-low CRF 12.5 (3.9)	2-medium CRF 26.6 (4.9)	3-high CRF 51.9 (13.1)	Cohen's d	1-low CRF 12.5 (3.0)	2-medium CRF 21.0 (2.9)	3-high CRF 39.0 (10.9)	Cohen's d
20mSRT (# laps)								
Physical literacy								
Total CAPL score	55.5 (12.0)	66.7 (11.3) ^a	73.3 (10.2) ^{bc*}	1.60	58.5 (10.7)	64.7 (10.0) ^a	71.7 (10.3) ^{bc*}	1.26
Physical competence domain								
Total domain score	18.4 (4.0)	22.2 (3.5) ^a	25.1 (2.8) ^{bc*}	1.94	19.5 (3.9)	21.9 (3.8) ^a	24.8 (3.4) ^{bc*}	1.45
Sit and reach score (cm)	22.3 (8.1)	24.3 (7.7)	25.2 (7.6) ^{c*}	0.37	31.3 (9.0)	31.1 (8.6)	33.4 (8.8) [*]	0.24
Handgrip score (kg)	39.5 (11.6)	42.9 (9.7) ^a	46.0 (11.1) ^{bc*}	0.57	38.6 (11.4)	40.0 (8.9)	42.5 (8.7) [*]	0.39
Plank score (sec)	42.3 (21.4)	65.9 (34.1) ^a	87.5 (41.9) ^{bc*}	1.36	44.3 (23.6)	57.1 (35.5) ^a	93.9 (58.7) ^{bc*}	1.11
BMI (kg/m ²)	21.8 (5.1)	19.9 (3.9) ^a	18.1 (2.4) ^{bc*}	0.93	21.6 (4.8)	20.0 (3.8) ^a	18.6 (2.5) ^{bc*}	0.78
Waist circumference (cm)	76.5 (14.2)	70.8 (9.9) ^a	66.0 (6.4) ^{bc*}	0.95	74.9 (11.5)	71.2 (10.8) ^a	67.9 (7.6) ^{bc*}	0.72
CAMSA score	29.8 (5.8)	34.3 (4.3) ^a	37.0 (3.4) ^{bc*}	1.52	30.4 (5.3)	33.3 (4.1) ^a	35.6 (4.4) ^{bc*}	1.07
Daily behavior domain								
Total domain score	14.4 (7.8)	18.3 (8.0) ^a	20.3 (8.0) ^{bc*}	0.75	15.1 (7.6)	17.1 (7.6)	19.6 (7.4) ^{bc*}	0.60
Average daily step counts	10080 (3441)	12132 (4337) ^a	13837 (4457) ^{bc*}	0.94	9048 (2683)	9851 (2948)	11132 (3251) ^{bc*}	0.70
Self-reported screen time (h/day)	3.6 (2.3)	2.8 (1.8) ^a	2.6 (1.9) [*]	0.47	3.0 (1.9)	2.5 (1.9) ^a	2.1 (1.5) ^{c*}	0.52
Average days meeting the guidelines	4.4 (2.1)	4.9 (1.8) ^a	5.6 (1.7) ^{bc*}	0.63	4.4 (1.7)	4.8 (1.9) ^a	5.4 (1.5) ^{c*}	0.62
Motivation and confidence domain								
Total domain score	11.0 (3.5)	12.7 (2.8) ^a	14.4 (2.1) ^{bc*}	1.18	11.2 (2.9)	12.2 (2.6) ^a	13.5 (2.3) ^{bc*}	0.88
CSAPPA adequacy score	4.3 (1.1)	4.9 (0.9) ^a	5.3 (0.6) ^{bc*}	1.13	4.3 (0.9)	4.6 (0.9) ^a	5.1 (0.8) ^{bc*}	0.94
CSAPPA predilection score	4.2 (1.2)	4.9 (1.0) ^a	5.3 (0.7) ^{bc*}	1.12	4.4 (1.1)	4.8 (1.0) ^a	5.1 (0.8) ^{bc*}	0.73
Benefits and barriers ratio	1.3 (1.4)	1.5 (1.2)	2.1 (1.0) ^{bc*}	0.66	1.2 (1.2)	1.5 (1.1)	1.8 (1.0) ^{c*}	0.54
Activity levels compared to peers	0.6 (0.2)	0.7 (0.2) ^a	0.8 (0.1) ^{bc*}	1.27	0.6 (0.2)	0.7 (0.2)	0.8 (0.2) ^{bc*}	1.00
Skill level compared to peers	0.6 (0.3)	0.7 (0.2) ^a	0.8 (0.2) ^{bc*}	0.79	0.6 (0.2)	0.6 (0.2) ^a	0.7 (0.2) ^{bc*}	0.50
Knowledge and understanding domain								
Total domain score	11.6 (2.8)	13.1 (2.4) ^a	13.4 (2.3) ^{c*}	0.70	12.4 (2.7)	13.2 (2.4) ^a	13.6 (2.2) ^{c*}	0.49
CAPL questionnaire score	11.6 (2.8)	13.1 (2.4) ^a	13.4 (2.3) ^{c*}	0.70	12.4 (2.7)	13.2 (2.4) ^a	13.6 (2.2) ^{c*}	0.49

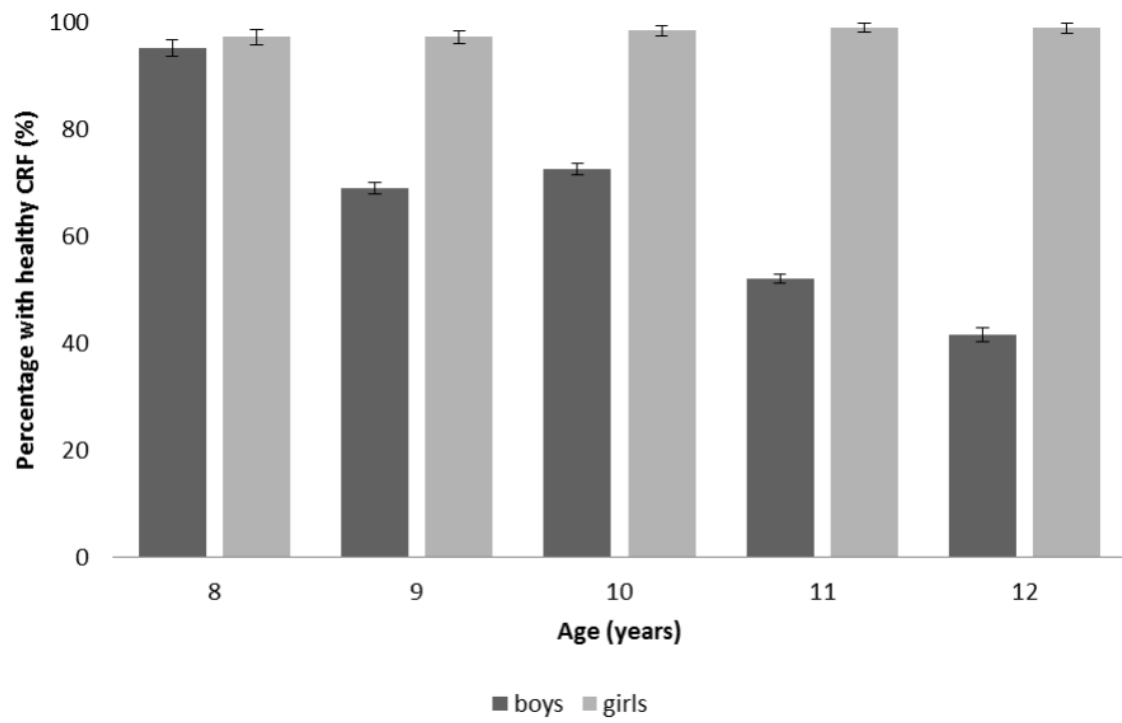
All variables are presented as means (SD). Statistical significance for the main effect was assessed using ANOVA. Bonferroni (Dunn) t-test was used to assess differences between groups. Cohen's *d* was used to calculate the effect size between the low and high CRF groups. CRF = cardiorespiratory fitness; 20mSRT = 20-m shuttle run test; BMI = body mass index; CSAPPA = Children's Self-Perception and Adequacy in and Predilection for Physical Activity; CAPL = Canadian Assessment of Physical Literacy.

* $p < 0.001$ for main effect; ^a $p < 0.05$ for 2 vs 1; ^b $p < 0.05$ for 3 vs 2; ^c $p < 0.05$ for 3 vs 1.

FIGURE CAPTION

Figure 1 Percentages of Canadian boys and girls with healthy cardiorespiratory fitness using the interim international standards [16]. Data are presented as % and standard deviations.

Figure 1



CHAPTER 6: MANUSCRIPT 4

This manuscript has been written and formatted following the publication standards of the peer-reviewed journal, *British Journal of Sport Medicine*. It appears in its published format.

Title:

International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results

Citation:

Lang JJ, Tremblay MS, Léger L, Olds T, Tomkinson GR. International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results. *British Journal of Sports Medicine*. Published Online First: 20 September 2016.
doi:10.1136/bjsports-2016-096224

Online link:

<http://bjsm.bmj.com/content/early/2016/09/20/bjsports-2016-096224.abstract>

International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results

Justin J Lang,¹ Mark S Tremblay,¹ Luc Léger,² Tim Olds,³ Grant R Tomkinson^{3,4}

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bjsports-2016-096224>).

¹Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

²Département de kinésiologie, Université de Montréal, Montréal, Québec, Canada

³Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health Sciences & Sansom Institute for Health Research, University of South Australia, Adelaide, South Australia, Australia

⁴Department of Kinesiology and Public Health Education, University of North Dakota, Grand Forks, North Dakota, USA

Correspondence to

Justin J Lang, Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Rd, Ottawa, Ontario, Canada K1H 8L1; jjlang@cheo.on.ca

Accepted 29 August 2016

ABSTRACT

Objectives To describe and compare 20 m shuttle run test (20mSRT) performance among children and youth across 50 countries; to explore broad socioeconomic indicators that correlate with 20mSRT performance in children and youth across countries and to evaluate the utility of the 20mSRT as an international population health indicator for children and youth.

Methods A systematic review was undertaken to identify papers that explicitly reported descriptive 20mSRT (with 1-min stages) data on apparently healthy 9–17 year-olds. Descriptive data were standardised to running speed (km/h) at the last completed stage. Country-specific 20mSRT performance indices were calculated as population-weighted mean z-scores relative to all children of the same age and sex from all countries. Countries were categorised into developed and developing groups based on the Human Development Index, and a correlational analysis was performed to describe the association between country-specific performance indices and broad socioeconomic indicators using Spearman's rank correlation coefficient.

Results Performance indices were calculated for 50 countries using collated data on 1 142 026 children and youth aged 9–17 years. The best performing countries were from Africa and Central-Northern Europe. Countries from South America were consistently among the worst performing countries. Country-specific income inequality (Gini index) was a strong negative correlate of the performance index across all 50 countries.

Conclusions The pattern of variability in the performance index broadly supports the theory of a physical activity transition and income inequality as the strongest structural determinant of health in children and youth. This simple and cost-effective assessment would be a powerful tool for international population health surveillance.

Background

Cardiorespiratory endurance (CRE) is described as a trait that is partly determined by genetic factors¹ and modifiable by physical activity.² There is growing evidence of an inverse relationship between CRE and a variety of important health markers in children and youth.^{3–6} Being a trait, CRE is less labile than physical activity,⁷ and thus tracks moderately well from childhood to adulthood.^{8–10} In addition, longitudinal studies have identified a significant graded association between low CRE in late adolescence and an increase in the risk of cardiovascular disease¹¹ and early death¹² in

adulthood. Taken together, these findings indicate that the population status of CRE in children and youth may help predict future non-communicable disease burden. As a result, there have been recent calls in North America,^{13–14} South America,¹⁵ Europe¹⁶ and Oceania¹⁷ to monitor CRE in children and youth through continuous surveillance efforts.

The 20 m shuttle run test (20mSRT)^{18–19} is considered the best and most popular field-based measurement of CRE in children and youth because of its low cost, simplicity and ability to test large groups of children simultaneously.^{20–22} In 9–17 years, it demonstrates moderate-to-high criterion-related validity, indicating that it is a good surrogate measure of maximal oxygen uptake,²³ the underlying construct of CRE. Novel data standardisation techniques also allow for meaningful international comparisons of 20mSRT performance in youth.²⁴ Recent studies have identified large variability in children's 20mSRT performance within and between countries,^{6–15–25–27} which may be associated with broad population-level socioeconomic, cultural and policy factors (henceforth referred to as socioeconomic indicators). Consequently, as a surveillance instrument, the 20mSRT could help identify populations with high CRE and provide insights into broad socioeconomic indicators that may inform intervention, programme and/or policy targets to help improve CRE in other populations, especially those with low CRE.

In 2006, Olds *et al*²⁷ published a seminal study describing the worldwide variability in 20mSRT performance in children and youth across 37 countries. That study represented an effort to describe and rank the international state of CRE in children and youth, serving as a comparison standard for countries. As a result, some countries were motivated to assess and verify their international 20mSRT performance rank²⁸ and others to establish country-specific normative data to determine their international rank compared to the standard.^{17–29} However, since 2006 it is possible that the variability in 20mSRT performance has changed due to temporal trends in children's CRE that have been reported to differ in magnitude and direction at the country level.^{22–24} These findings support the need for a 10-year update of Olds *et al*'s initial study.

Thus, the objectives of this paper were threefold: (1) to update, describe and compare the 20mSRT

To cite: Lang JJ, Tremblay MS, Léger L, *et al*. *Br J Sports Med* Published Online First: [please include Day Month Year] doi:10.1136/bjsports-2016-096224

Original article

performance of children and youth across countries; (2) to explore broad, country-level, socioeconomic indicators that correlate with 20mSRT performance in children and youth and (3) to evaluate the utility of the 20mSRT as an international population health indicator for children and youth.

METHODS

Data source

A systematic review of literature was conducted to locate studies that reported descriptive 20mSRT data on apparently healthy (free from known disease/injury) 9–17 year-olds. The search strategy for this study is reported in detail by Tomkinson *et al.*³⁰ Briefly, researchers conducted a search of five databases, including MEDLINE (1946–2015), PsycINFO (1806–2014), EMBASE (1947–2014), SPORTDiscus (1949–2014) and Cochrane Central Register of Controlled Trials (2005–2014). The search strategy included the following terms: shuttle run*, OR beep test, OR multi-stage, OR aerobic, OR cardio*, OR endurance; with child*, adolescen*, pubescen* boy, girl, young and youth as search term modifiers. No search restriction was placed on language. Studies were excluded if they reported data on special interest groups (eg, elite athletes and mentally/physically disabled) or unhealthy/injured groups. Consequently, studies were included if participants were broadly representative of their source population and if descriptive data were provided for 20mSRT performance (using the 1-min stage protocol) at the sex by age by country level. The reference lists for all included studies were reviewed for additional manuscripts not identified through the database search. When needed, attempts were made to contact study authors to clarify study details, to request raw data and/or to ask if they knew of other relevant studies. In addition, the personal libraries of the study authors were reviewed for relevant studies.

Inclusion/exclusion

A total of 979 studies were identified as relevant. Of these, 802 studies were excluded for one or more of the following reasons: (1) an unknown version of the 20mSRT; (2) descriptive data were combined for girls and boys into a single group; (3) study participants were atypical of the general population (eg, elite athletes or physically delayed); (4) the study reported duplicate data already reported in another included study and (5) sample size at each sex by age by study group was <40 participants as the means and SDs for smaller samples were too labile. All 20mSRT performance data from the remaining 177 studies were extracted into Excel (Microsoft Office 2010, USA) using a standardised template (see online supplementary material 1 for a description of the studies).

Standardisation of data and statistical analysis

The goal of data standardisation was to express all 20mSRT performance scores as a common protocol and result metric to facilitate cross-country comparisons. Figure 1 describes the general methodological approach used in this study. In brief, three common protocol variants for the 20mSRT were identified,²⁴ with all performances standardised to the Léger protocol.¹⁹ This protocol begins at a speed of 8.5 km/h for the first stage (min) and increases by 0.5 km/h every stage thereafter. Furthermore, 20mSRT performances were reported in one of the six different result metrics (speed at the last completed stage, completed stage or minutes, distance, completed laps within final stage, total laps or estimated $\text{VO}_{2\text{peak}}$) with all results standardised to running speed (km/h) at the last

completed stage. Details on data standardisation procedures are described by Tomkinson *et al.*^{24 30}

The main objective of this study was to compare 20mSRT performances across countries. To achieve this, sample-weighted mean and SD 20mSRT performance values were calculated for each sex by age by country group and used to calculate sex- by age- by country-specific mean z-scores using the international 20mSRT performance norms (table 1) reported by Tomkinson *et al.*³⁰

$$\text{Mean z-score} = \frac{\bar{X} - \mu}{\sigma}$$

where $\bar{X}$ is the country sample-weighted mean sex- by age-specific performance score, μ is the international mean performance score for the given sex- by age-specific group and σ is the international SD for the given sex- by age-specific group (see table 1).

Positive z-scores indicated that 20mSRT performances were better than the international mean and negative z-scores that

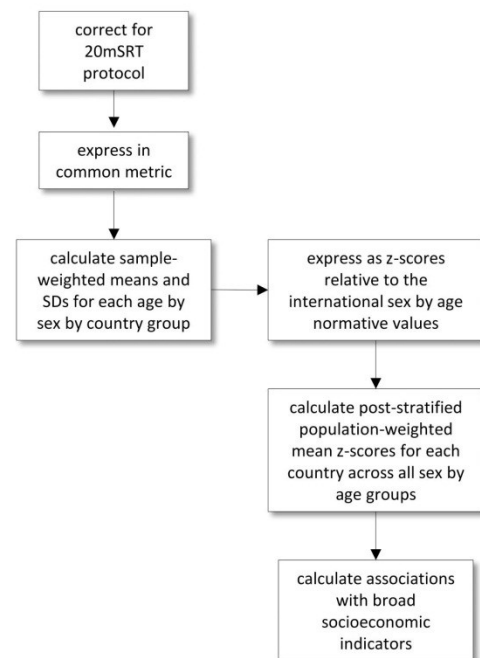


Figure 1 Flow chart describing the general methodological procedure for this study. All protocol results (Léger, Eurofit or Queen's University of Belfast) were corrected to reflect the Léger protocol. All results were then corrected to reflect speed (km/h) at the last completed stage as a common result metric. Next, sample-weighted mean 20mSRT performance score and SD were calculated for each sex by age by country group. Then z-scores were generated relative to the international norms for each sex by age by country group to calculate the mean country specific z-score, described as the 'performance index.' Each country specific performance index was corrected using a poststratified population-weighted technique. Finally, a correlation analysis took place to analyse associations between the performance index and broad socioeconomic indicators. 20mSRT, 20 m shuttle run test.

Table 1 International 20mSRT norms for running speed (km/h) at the last completed stage³⁰

Age	Boys			Girls		
	N	Mean	SD	N	Mean	SD
9	101 532	10.03	0.85	98 166	9.71	0.70
10	81 538	10.13	0.92	81 415	9.75	0.77
11	79 660	10.25	1.00	78 029	9.79	0.83
12	70 266	10.47	1.11	69 497	9.83	0.88
13	72 913	10.73	1.21	62 491	9.86	0.91
14	65 317	10.97	1.28	55 035	9.89	0.93
15	53 287	11.12	1.33	43 730	9.91	0.94
16	35 490	11.27	1.37	34 978	9.93	0.95
17	29 123	11.41	1.42	29 559	9.96	0.97
Total	589 126			552 900		

20mSRT, 20 m shuttle run test.

20mSRT performances were worse than the international mean. Population-weighted mean z-scores and corresponding 95% CIs were calculated for each country, across all sex by age groups for which data were available, using a poststratification population-weighting procedure described by Levy and Lemeshow.³¹ Population estimates standardised to the mean measurement year of 2000 were extracted from the United Nations World Population Prospective report, that's what we mean by mean measurement year.³² These population-weighted z-scores represented the country-specific 'performance index', that is, the overall standardised deviation of a country's children 20mSRT performance from the international age- and sex-specific means. The index served as a method to rank and compare the overall performance of each country.

The secondary objective was to explore correlates between the 20mSRT performance index and broad socioeconomic indicators. An internet search of publicly available databases was conducted to identify socioeconomic indicators that, based on the literature, would be meaningful to explore. In total, six indicators were identified (table 2), and data were extracted at the country level. The association between the 20mSRT performance index and each socioeconomic indicator was quantified using Spearman's rank correlation coefficient with the 95% confidence limits calculated using Fisher's z-transformation. Holm-Bonferroni sequential correction was used to adjust for α slippage.³⁹ Correlations of 0.1, 0.3 and 0.5 were used as thresholds for weak, moderate and strong, respectively, with correlations <0.1 considered to be trivial.⁴⁰ In addition, countries were grouped by the human development index (HDI) into developed (HDI of 0.741–0.943) and developing (HDI of 0.286–0.740) nations using the 2011 inequality-adjusted United Nations Human Development Report.³⁶ The association between the 20mSRT performance index and each indicator was further explored by HDI status to identify potential interactions. Combined, the first two objectives of this study helped inform and evaluate the utility of the 20mSRT as an international population health indicator for children and youth.

RESULTS

Cross-country variability

Standardised 20mSRT performance indices are presented across 50 countries, representing 1 142 026 children and youth aged 9–17 years. The 50 countries represented six of the seven continents (North and South America, Europe, Africa, Asia and Oceania), and a wide range of HDI (0.400 (Côte d'Ivoire) to 0.943

Table 2 Potential correlates of 20mSRT performance in children and youth

Variable	Data source	Description
Childhood obesity (%)	Ng <i>et al</i> ³³ Data available for 49 countries	Calculated as the mean country-specific proportion of boys and girls aged <20 years that classified as obese based on the International Obesity Task Force definition. Countries with a higher proportion received a higher rank.
Physical activity (%)	Kalman <i>et al</i> ³⁴ Data originally obtained from HSBC study Data available for 22 developed countries	Calculated as the mean country-specific percentage of boys and girls aged 11, 13 and 15 years that achieved at least 1 hour of moderate-to-vigorous physical activity per day. Countries with a higher proportion received a higher rank.
Mean climate (°C)	Weatherbase ³⁵ Data available for 50 countries	Calculated across the number of years and cities where data were available (ie, Canada had 26 years of data across 2287 cities). Colder countries received a higher rank.
HDI	The United Nations ³⁶ Data available for 50 countries	Summarises country-specific achievement in a variety of indicators related to standards of living through a scale from 0.456 (low HDI) to 0.889 (very high HDI). High HDI received a higher rank.
Gini index	The World Bank ³⁷ Data available for 48 countries	Summarises the distribution of wealth among individuals in a country where 0 represents perfect equality and 100 implies perfect inequality. Countries closest to inequality received a higher rank.
Urbanisation (%)	The World Bank ³⁸ Data available for 46 countries	Proportion of the population that reside in an urban area within the country. Countries with a higher percentage of urban areas received a higher rank.

20mSRT, 20 m shuttle run test; HDI, human development index.

(Norway)).³⁶ The performance index values for each of the 50 countries are presented in table 3 along with their respective 95% CI and centile rank. Confidence in our performance index scores varied substantially between countries based on country-specific sample size. Performance indices ranged from +1.85 for Tanzania (the best performing country) to −0.96 for Mexico (the worst performing country). Countries from Central and Northern Europe and Africa accounted for 9 of the top 10 performing countries. South American countries were consistently among the worst performing countries, with all countries except Colombia and Suriname, falling below the international mean (mean±95% CI: centile 40±11; z-score −0.27±0.32). Collectively, Africa (mean±95% CI: centile 68±8; z-score 0.53±0.32) was typically the best performing continent, closely beating Europe (mean±95% CI: centile 63±7; zscore 0.37±0.21). However, when dividing Europe into two segments, similar to that performed by Ortega *et al*,¹⁹³ a performance gradient existed where Central-Northern European countries (mean±95% CI: centile 66±7; z-score 0.48±0.23) consistently outperformed countries from Southern Europe (mean±95% CI: centile 45±7; z-score −0.12±0.19).

Correlates of 20mSRT performance

The Gini index was a moderate to strong negative correlate of 20mSRT performance, regardless of country development status (table 4). Other correlations across the sample ranged from

Original article

Table 3 Descriptive data for countries presented in rank order of by 2016 performance index

Country	N	Performance index $\pm$ 95% CI	Centile rank
Tanzania ⁴¹	274	1.85 $\pm$ 0.08	97
Iceland ⁴²	6130	1.45 $\pm$ 0.03	93
Estonia ^{43–48}	8013	1.42 $\pm$ 0.02	92
Norway ⁴⁹	2305	1.15 $\pm$ 0.04	87
Japan ⁴⁹	279 732	0.98 $\pm$ 0.01	84
Denmark ⁵⁰	9342	0.85 $\pm$ 0.03	80
Czech Republic ^{51 52}	1953	0.75 $\pm$ 0.05	77
Finland ^{51 53 54}	5679	0.74 $\pm$ 0.03	77
Côte d'Ivoire ⁵⁵	3273	0.68 $\pm$ 0.04	75
Morocco ⁵⁶	2522	0.65 $\pm$ 0.03	74
France ^{16 57–62} (G Cazorla, JP Colin, L Léger. Condition physique des jeunes collégiens et lycéens de la région aquitaine (France). Unpublished manuscript, 1999).	7244	0.59 $\pm$ 0.03	72
Lithuania ^{43 45}	6479	0.52 $\pm$ 0.02	70
Benin ⁶³	4141	0.51 $\pm$ 0.03	70
Djibouti ⁶⁴	1524	0.48 $\pm$ 0.04	68
Mauritius ⁶⁵	2695	0.47 $\pm$ 0.03	68
Slovakia ^{66–71}	4827	0.39 $\pm$ 0.03	65
Germany ^{16 51}	1146	0.38 $\pm$ 0.06	65
Sénégal ⁷²	2243	0.37 $\pm$ 0.04	64
Canada ^{18 73–77}	29 009	0.32 $\pm$ 0.01	63
Turkey ⁷⁸	45	0.32 $\pm$ 0.20	63
UK ^{79–91} (RM Ranson, S Taylor. Unpublished 'F' Factor Report for Denbighshire County Council. Wrexham, UK: Glyndwr University, 2010)	49 140	0.31 $\pm$ 0.01	62
Poland ^{92–95}	58 681	0.30 $\pm$ 0.01	62
Suriname ⁹⁶	270	0.30 $\pm$ 0.08	62
Belgium ^{16 51 97–105}	23 358	0.28 $\pm$ 0.01	61
China ¹⁰⁶	548	0.25 $\pm$ 0.07	60
Sweden ¹⁶	45	0.20 $\pm$ 0.37	58
Austria ¹⁶	70	0.19 $\pm$ 0.23	58
Netherlands ^{107–109}	3457	0.18 $\pm$ 0.03	57
Spain ^{16 110–122}	25 950	0.14 $\pm$ 0.01	56
South Africa ^{123–130}	2970	0.09 $\pm$ 0.04	54
Switzerland ^{131 132}	3224	0.08 $\pm$ 0.03	53
Uganda ¹³³	1446	0.08 $\pm$ 0.05	53
Seychelles ^{134 135}	9174	0.07 $\pm$ 0.02	53
Colombia ^{136–138}	36 117	0.05 $\pm$ 0.01	52
Australia ^{139–149} (S Mulkearns, G Naughton, J Carlson, et al. The effects of additional physical education on health-related fitness measures of circumpubertal females. Unpublished manuscript, 2000.)	26 719	–0.01 $\pm$ 0.01	50
Hungary ^{16 51 150 151}	3294	–0.08 $\pm$ 0.04	47
Italy ^{16 152–157}	5900	–0.09 $\pm$ 0.02	46
Hong Kong ¹⁵⁸	1214	–0.11 $\pm$ 0.05	46
Chile ^{15 159}	17 057	–0.13 $\pm$ 0.01	45
Argentina ¹⁶⁰	1259	–0.16 $\pm$ 0.05	44
Philippines ¹⁶¹	340	–0.19 $\pm$ 0.11	42
Greece ^{51 162–166}	268 382	–0.26 $\pm$ 0.00	40
Portugal ^{167–172} (Oliveira personal communication)	30 764	–0.26 $\pm$ 0.01	40
Brazil ^{173–175}	4528	–0.34 $\pm$ 0.03	37
Cyprus ^{176 177}	6902	–0.34 $\pm$ 0.03	37

Continued

Table 3 Continued

Country	N	Performance index $\pm$ 95% CI	Centile rank
Republic of Korea ¹⁷⁸	82	–0.38 $\pm$ 0.12	35
USA ^{179–189}	171 058	–0.47 $\pm$ 0.01	32
Latvia ¹⁹⁰	3400	–0.64 $\pm$ 0.03	26
Peru ¹⁹¹	560	–0.67 $\pm$ 0.06	25
Mexico ¹⁹² (Esteve-Lanao personal communication)	7541	–0.96 $\pm$ 0.02	17

trivial to moderate. Mean climate was a strong positive correlate for only developed countries. In contrast, HDI was a strong negative correlate for only developing countries. Furthermore, a potential interaction existed where some socioeconomic indicators were positively correlated with 20mSRT performance in developed countries, but negatively associated in developing countries.

DISCUSSION

This study represents the largest international surveillance effort to describe the status of CRE in children and youth by presenting standardised 20mSRT performance indices across 50 countries. The aim of the present study was to update, describe and compare 20mSRT performance in children and youth across countries; to explore associations between 20mSRT performance and broad country-level socioeconomic indicators and to assess the utility of the 20mSRT as an international population health indicator for children and youth.

International variability in 20mSRT performance

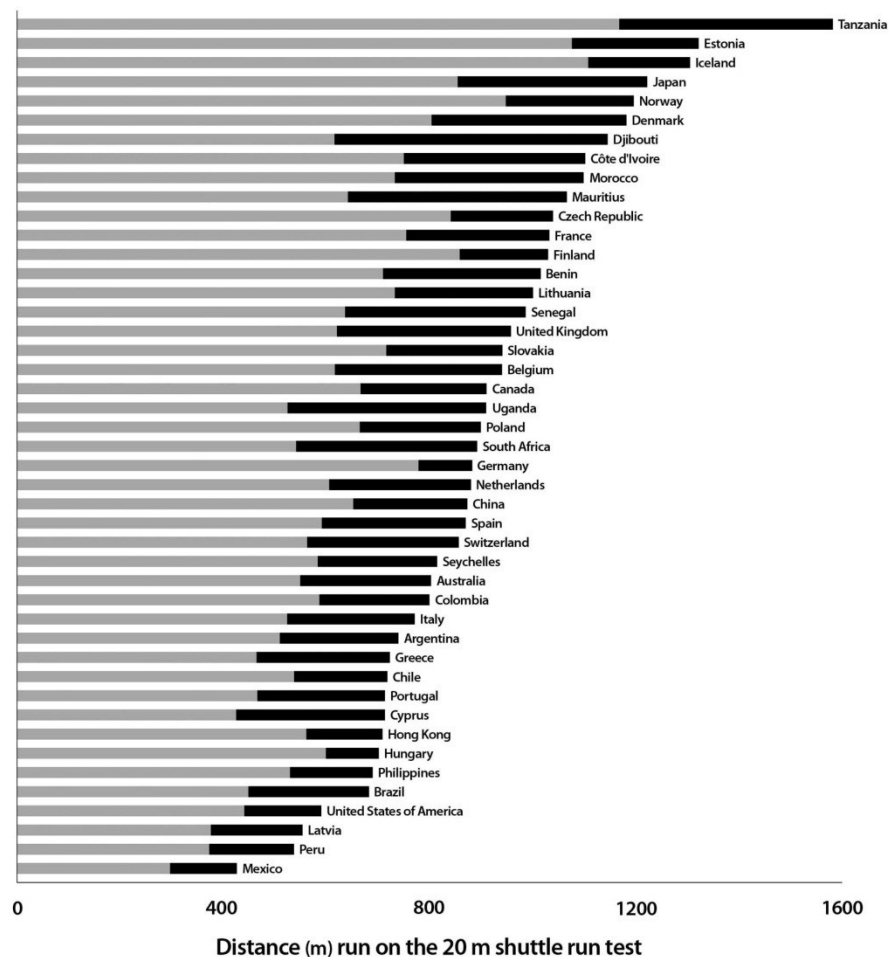
There was a very high correlation ($r=0.90$) between the 33 common country performance indices reported in this study and in Olds *et al.*²⁷ Furthermore, our results suggest that 20mSRT performances in children and youth vary substantially across countries, spanning 2.8 SDs from the best performing country (Tanzania) to the worst performing country (Mexico). To illustrate this finding, figure 2 depicts the country-specific mean distance (m) run on the 20mSRT by boys and girls, which indicates that boys and girls from Tanzania would cover a distance approximately four times greater than their Mexican counterparts (equivalent to a maximal running speed of ~ 1.3 times faster). The magnitude of variability in CRE between countries is concerning and suggests that low CRE may be a public health issue in some countries. Indeed, this substantial variability provides some evidence supporting CRE as a population-level indicator of health in children and youth. Future research should further explore the direct impact of low CRE in childhood and the future disease outcomes in adulthood. However, following a quartile framework, countries that fall below the 25th centile should take action by implementing policies and programmes targeted at increasing CRE in children and youth. Likewise, countries that fall between the 25th and 50th centile should take precautionary measures.

A total of 23 European countries are represented in this study. Our findings suggest a clear latitudinal gradient where Central-Northern countries perform better on the 20mSRT than their Southern counterparts by ~ 0.6 SDs. Following the same Central-Northern versus Southern division of Europe, Ortega *et al.*¹⁹³ also found a difference in CRE where Central-Northern countries performed significantly better on the 20mSRT than

Table 4 Spearman's rank correlation coefficients (95% CI) between 20mSRT performance index and socioeconomic indicators

	All countries (n=50)	Developed (n=34)	Developing (n=16)
Childhood obesity (%)	-0.19 (-0.44 to 0.09)	-0.27 (-0.56 to 0.08)	-0.00 (-0.50 to 0.50)
Physical activity (%)		-0.12 (-0.52 to 0.32)	
Mean climate (°C)	0.26 (-0.02 to 0.50)	0.63** (0.37 to 0.80)	-0.21 (-0.64 to 0.32)
HDI	0.01 (-0.27 to 0.29)	0.29 (-0.05 to 0.57)	-0.56 (-0.83 to -0.09)
Gini index	-0.44* (-0.64 to -0.18)	-0.70** (-0.84 to -0.47)	-0.49 (-0.79 to 0.01)
Urbanisation (%)	-0.01 (-0.30 to 0.28)	0.15 (-0.20 to 0.47)	-0.45 (-0.80 to 0.13)

Developed=high and very high (0.741–0.943) HDI; developing=very low, low and medium (0.286–0.740) HDI.

* indicates significant result at $p < 0.05$; ** indicates significant result at $p < 0.01$; adjusted using Holm–Bonferroni sequential correction. 20mSRT, 20 m shuttle run test; HDI, Human development index.**Figure 2** Bar graph describing the distance (m) run by boys and girls on the 20mSRT based on country-specific and age-specific performance indices. Only countries (n=45) with 20mSRT performance indices for boys (represented by the black bars) and girls (represented by the grey bars) are shown. 20mSRT, 20 m shuttle run test.

their Southern counterparts. These findings have been attributed to differences in physical activity levels in children and youth across countries,^{194–195} which may be further exacerbated by disparities in socioeconomic indicators.¹⁹⁶

Within countries we found sex differences, where boys consistently outperformed girls on the 20mSRT across all countries (figure 2). Sex differences in CRE are consistently reported in the literature,^{27–197} with our study providing a unique

Original article

perspective that suggests the magnitude of the sex differences may be country specific. These sex-related differences may be due to age-related and maturation-related differences between countries. Further analysis of cross-country sex differences is beyond the scope of this paper. Nevertheless, there is a need for further research on this topic.

Possible socioeconomic correlates

Following an ecological framework,¹⁹⁸ highlighting correlates of 20mSRT performance may provide country-level target areas for policies and programmes that may ultimately help increase CRE in children and youth. There are many possible broad country-level correlates of 20mSRT performance in children and youth. These may include HDI, the distribution of wealth within a country, urbanisation, climate and country-specific physical activity and childhood obesity levels.

Childhood obesity

Internationally, the measurement and surveillance of body composition is a public health norm with prevalence data available for ~188 countries.³³ Studies indicate that, similar to CRE, the prevalence of overweight and obesity in children and youth varies substantially across countries.³³ Our study supports childhood obesity prevalence as a weak negative correlate of 20mSRT performance in developed countries ($r_s = -0.27$; 95% CI: -0.56 to 0.08). In developing countries, the correlation between childhood obesity prevalence and 20mSRT performance was trivial ($r_s = -0.00$; 95% CI: -0.50 to 0.50), suggesting that childhood obesity prevalence is a stronger negative correlate of CRE in more developed countries.

Physical activity

Physical activity is considered a modifier of CRE. Thus, we would anticipate that higher population levels of physical activity result in higher CRE levels in children and youth.¹⁹⁹ However, physical activity is a behaviour that varies substantially across time and is difficult to measure at the population level.^{7 200 201} Validated self-report surveys are the most popular way to estimate physical activity across countries in children and youth, such as the Global School-based Student Health Survey²⁰² and the Health Behaviour in School-aged Children (HBSC) survey.²⁰³ Data from the HBSC study suggest that moderate-to-vigorous physical activity is a weak negative correlate of 20mSRT performance in children and youth in developed countries ($r_s = -0.12$; 95% CI: -0.52 to 0.32). This result may reflect limitations of the HBSC study which used self-report measures for selected age groups across less than half of the countries for which performance index data were available. In contrast, these results may also reflect the fact that moderate-to-vigorous physical activity is not a good correlate of maximal-effort CRE in children and youth.

More recently, efforts have been made to use objective measures of physical activity (pedometers and accelerometers),¹⁹⁴ but these efforts have been predominantly adopted in developed countries.²⁰⁴ However, a recent international study comparing objectively measured physical activity suggests that children in Northern Europe and Africa are more active than children in South America, broadly consistent with our findings.²⁰⁵

Urbanisation

Worldwide, there is a major population demographic shift where now, more than ever, the majority of the world's population lives in urban and/or suburban settings.²⁰⁶ These demographic shifts are synonymous with a variety of lifestyle

transitions as a result of mechanisation replacing laborious jobs, and an increase in transportation choices and food availability. A population shift towards cities is described as a component of the epidemiological²⁰⁷ and physical activity²⁰⁸ transitions due to the resulting decrease in habitual physical activity in these populations. The physical activity transition is a theoretical perspective that can be described as:

'Overall, socioeconomic and demographic changes occurring in concert with the epidemiological transition appear to be associated with the adoption of a lifestyle characterized by lower energy requirements at work and for the procurement of food, less dependence on walking as a primary means of transportation, and the popularity of sedentary activities to fill a growing availability of leisure time no longer needed to carry out activities related to securing food and shelter. (ref. 208, p. 277)'

Since physical activity is a modifier of CRE, there are plausible associations between high urbanisation and low CRE levels in children and youth.

In our study, urbanisation was a weak correlate of 20mSRT performance for developed countries ($r_s = 0.15$; 95% CI: -0.20 to 0.47). This result is most likely due to developed countries already attaining a high level of urbanisation. In contrast, urbanisation is a moderate negative correlate of 20mSRT performance in developing countries ($r_s = -0.45$; 95% CI: -0.80 to 0.13). Although this finding is not statistically significant, it describes the potential negative impact of rapid urbanisation with regard to CRE. This pattern is especially apparent in some South American countries (Brazil, Peru and Colombia) with high urbanisation and low CRE levels. Therefore, these countries might be in the midst of a physical activity transition. This is a drastically different demographic situation compared to countries with relatively low urbanisation rates (Uganda, Sénégal, Mauritius and Benin) that demonstrate high CRE levels in children and youth.

Human Development Index

The HDI is a composite measure combining country-specific prosperity and various indicators of quality of life. Our study indicates that HDI is a weak positive correlate of 20mSRT performance in developed countries ($r_s = 0.29$; 95% CI: -0.05 to 0.57). In contrast, developing countries show an opposite gradient where HDI is a strong negative correlate of 20mSRT performance in children and youth ($r_s = -0.56$; 95% CI: -0.83 to -0.09). A likely explanation of these results is that physical activity in developing countries is often obligatory because food is relatively scarce, and demanding chores required for survival. However, when these countries become more developed, food becomes readily available and the presence of mechanisation and automation makes physical activity no longer obligatory. In developed countries, it is likely that leisure time and wealth provides alternative opportunities for physical activity pursuits. These results lend further support to the presence of a physical activity transition in developing countries.²⁰⁸

Gini index

The Gini index is a population measure of the distribution of wealth. Our study indicates that a large population distribution in wealth (income inequality) is a moderate to strong negative correlate of CRE in children and youth, regardless of country development status. This finding suggests that policies aimed at reducing the gap between rich and poor might be a suitable population approach to increase CRE. Our results align with previous research that highlighted income inequality as the

strongest structural determinant of health in adolescents.²⁰⁹ Furthermore, Elgar *et al*²¹⁰ found that high national income inequality was associated with low population levels of physical activity, high BMI and more psychological and physical health symptoms in youth.

Climate

Our study indicates that climate is a strong positive correlate of 20mSRT performance ($r_s=0.63$; 95% CI: 0.37 to 0.80) for developed countries, where colder countries have higher CRE than warmer countries. In developing countries, climate is a weak negative correlate ($r_s=-0.21$; 95% CI: -0.64 to 0.32), where warmer countries tend to perform better than their colder counterparts. These results are conflicting and difficult to interpret. It is possible that the strong positive correlation in developed countries is due to the negative physiological effect of exercising in warm and humid temperatures. However, in contrast to previous findings, the strong correlation in developed countries remained significant after removing Northern European countries from the analysis ($r_s=0.53$; 95% CI: 0.21 to 0.75), suggesting that the result was not an artefact of the superior performance of children and youth in Northern Europe.²⁷

Strengths and limitations

Our study represents the largest and most comprehensive cross-country comparison of children's CRE. Although this study is not the first of its kind, it does provide a significant update to the study by Olds *et al*²⁷ by extending the coverage date from 2003 to 2015, and by adding an additional 13 countries and 724 000 data points to the analysis. As a result of the broad scope of this study, we acknowledge several limitations: (1) globally, CRE has declined by 2.2% per decade from 1981 to 2015 (for all children), with little change over the past 15 years. This decline may have artificially increased the rankings of countries with measurements that were made earlier (ie, Benin in 1987, Côte d'Ivoire in 1985, Turkey in 1991 and Suriname in 1991) and decreased those that were made more recently; (2) we combined a variety of studies with different sampling strategies (probability and non-probability) and sampling frames (national, state/provincial and regional); (3) some country-specific performance indices were calculated using small samples and limited sex by age groups which reduced the confidence in some of our country estimates; (4) the testing conditions may have varied between included studies (ie, indoor/outdoor, altitude and temperature); (5) measurement drift may have occurred in administering the 20mSRT between and within countries and (6) individual physiological (eg, mechanical efficiency,²¹¹ fractional utilisation²¹² and lactate threshold²¹³), psychosocial (eg, motivation and self-efficacy²⁰) and cultural factors may have also contributed to differences in performance and may have changed in some countries over time.

To strengthen our study, we combined data from 177 studies, representing the best available data on 20mSRT performance in children and youth, using rigorous and standardised data treatment techniques. These techniques control for variability in reporting 20mSRT results and provide us with the ability to compare testing results meaningfully between studies. We used a poststratification population-weighting procedure, so that our performance index reflects the country-specific population demographics. Furthermore, we believe that the size of our study, with over 1.1 million test performances, will help decrease any study specific limitations.

Surveillance of CRE as a powerful population health indicator in children and youth

Convincing evidence indicates that the precursors for future disease have their origin in childhood and adolescence, making these age groups very important for current and future health.²¹⁴ Thus, childhood and adolescent health promotion efforts in the form of interventions, programmes and policies are important public and population health initiatives. However, it is impossible to know the impact of such efforts without the consistent surveillance and monitoring of population health indicators. Currently, there are two comparable population health indicators in children and youth—body weight status and physical activity levels. The surveillance of body weight status at a population level is essential, but changes in body weight do not occur fast enough to understand the short- and medium-term impact of newly implemented policies and programmes. Evidence also suggests that individuals can classify as obese but still have healthy CRE and a healthy phenotype, known as the obesity paradox.²¹⁵ Ideally, monitoring physical activity levels in a population would provide important feedback on health promotion efforts, but there remain several limitations in the effectiveness and feasibility of measuring physical activity in populations of children and youth.^{200–201} Although physical activity and CRE are independently associated with health markers,²¹⁶ conceptually, CRE is a possible proxy for overall physical activity levels within a population. Thus, the measurement of CRE, via the 20mSRT, can be an objective surrogate measure that can be used to study the link between physical activity and health. As a population health indicator, the 20mSRT is an effective and feasible tool that demonstrates minimal barriers with regard to culture and language, is cost-effective, results are easy to score and interpret, and it can be easily administered with large groups of children.²⁰ We believe that the population-wide and international measurement of CRE is important to understand the health of children and youth, the impact of health promotion policies and programmes as well as the potential future population burden of non-communicable disease.

Recommendation

We recommend the development of an international 20mSRT internet portal that provides resources to correctly administer and record 20mSRT performances following the recommendations of Tomkinson *et al*.³⁰ This easy-to-access portal would also include an international 20mSRT data repository where researchers, physical educators, coaches and individuals could input their performance data for children and youth free of charge. In return, the website would compute a results sheet comparing the group/individual performance score with international,³⁰ regional and country normative centile values, providing the user with general performance feedback. All data would be housed on a secure server and used to continuously update international normative values, cross-country comparisons and temporal trends in 20mSRT performance. This data repository would serve as an international surveillance system to monitor 20mSRT performance in children and youth. The authors are currently working to launch such a repository.

CONCLUSION

Our study demonstrates the broad international utility of the 20mSRT in children and youth. This field-based assessment is suitable for international surveillance because of its low cost, ease in administering, very good test-retest reliability and

Original article

validity as well as ability to test large groups of children simultaneously. This study used a systematic review method, rigorous data treatment and a poststratification population-weighting procedures to obtain and compare 20mSRT performance data across 50 countries. Our data suggest that children and youth in Central-Northern Europe and Africa are among the world's best 20mSRT performers, whereas those from South America are among the worst performers. Worldwide patterns of variability in 20mSRT performance broadly mirror socioeconomic indicators in a way that support the theory of a physical activity transition and the importance of wealth distribution as a strong structural determinant of health. These data can help inform the development of an international 20mSRT surveillance system to monitor the variability and temporal trends in 20mSRT performance for children and youth across countries, providing insights into the general health and well-being of populations.

What are the findings?

- This study provides a 50-country comparison of 20 m shuttle run performance using rigorous data standardisation techniques.
- There are broad country-specific socioeconomic indicators that correlate with 20 m shuttle run performance in children and youth.
- 20 m shuttle run test data can be used as an international population health indicator for surveillance in children and youth.

Acknowledgements The authors acknowledge all individuals who provided details of their study and/or generously offered their raw data.

Contributors JLL, GRT and MST developed the research questions and objectives. JLL and GRT led the data analysis and synthesis of results. JLL wrote the manuscript. All authors contributed to interpreting results, editing, reviewing and approved the final manuscript.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Tucker R, Collins M. What makes champions? A review of the relative contribution of genes and training to sporting success. *Br J Sports Med* 2012;46:555–61.
- 2 Gutin B, Yin Z, Humphries MC, et al. Relationships of moderate and vigorous physical activity to fitness and fatness in adolescents. *Am J Clin Nutr* 2005;81:746–50.
- 3 Ortega FB, Ruiz JR, Castillo MJ, et al. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes (Lond)* 2008;32:1–11.
- 4 IOM (Institute of Medicine). *Fitness measures and health outcomes in youth*. Washington DC: The National Academies Press, 2012:111–51.
- 5 Ruiz JR, Ortega FB, Rizzo NS, et al. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. *Pediatr Res* 2007;61:350–5.
- 6 dos Santos FK, Prista A, Gomes TNQF, et al. Body mass index, cardiorespiratory fitness and cardiometabolic risk factors in youth from Portugal and Mozambique. *Int J Obes (Lond)* 2015;39:1467–74.
- 7 Ridley K, Olds T, Hands B, et al. Intra-individual variations in children's physical activity patterns: Implications for measurement. *J Sci Med Sport* 2009;12:568–72.
- 8 Malina RM. Physical activity and fitness: pathways from childhood to adulthood. *Am J Hum Biol* 2001;13:162–72.
- 9 Ortega FB, Ruiz JR, Labayen I, et al. Role of socio-cultural factors on changes in fitness and adiposity in youth: a 6-year follow-up study. *Nutr Metab Cardiovasc Dis* 2013;23:883–90.
- 10 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.
- 11 Höglström G, Nordström A, Nordström P. High aerobic fitness in late adolescence is associated with a reduced risk of myocardial infarction later in life: a nationwide cohort study in men. *Eur Heart J* 2014;35:3133–40.
- 12 Höglström G, Nordström A, Nordström P. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. *Int J Epidemiol*. Published Online First: 20 December 2015.
- 13 Kaminsky LA, Arena R, Beckie TM, et al. The importance of cardiorespiratory fitness in the United States: the need for a national registry: a policy statement from the American Heart Association. *Circulation* 2013;127:652–62.
- 14 Morrow JR Jr and Ede A. Research Quarterly for Exercise and Sport lecture. Statewide physical fitness testing: a big waist or a big waste? *Res Q Exerc Sport* 2009;80:696–701.
- 15 Garber MD, Sajuria M, Lobelo F. Geographic variation in health-related physical fitness and body composition among Chilean 8th graders: a nationally representative cross-sectional study. *PLoS ONE* 2014;9:e108053.
- 16 Ortega FB, Artero EG, Ruiz JR, et al. Physical fitness levels among European adolescents: the HELENA study. *Br J Sports Med* 2011;45:20–9.
- 17 Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9–17 year-old Australians since 1985. *Br J Sports Med* 2013;47:98–108.
- 18 Léger L, Lambert J, Goulet A, et al. [Aerobic capacity of 6 to 17-year-old Quebecois—20 meter shuttle run test with 1 minute stages]. *Can J Appl Sport Sci* 1984;9:64–9.
- 19 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.
- 20 Tomkinson GR, Olds TS. Field tests of fitness. In: Armstrong N, van Mechelen W, eds. *Paediatric exercise science and medicine*. New York, NY: Oxford University Press, 2008:109–28.
- 21 Castro-Piñero J, Artero EG, España-Romero V, et al. Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med* 2010;44:934–43.
- 22 Tomkinson G, Olds T. Secular changes in pediatric aerobic fitness test performance: the global picture. *Med Sport Sci* 2007;50:46–66.
- 23 Mayorga-Vega D, Aguilar-Soto P, Viciana J. Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: a meta-analysis. *J Sports Sci Med* 2015;14:536–47.
- 24 Tomkinson GR, Léger LA, Olds TS, et al. 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;33:285–300.
- 25 Adamo KB, Sheel AW, Onywera V, et al. Child obesity and fitness levels among Kenyan and Canadian children from urban and rural environments: a KIDS-CAN Research Alliance Study. *Int J Pediatr Obes* 2011;6:e225–32.
- 26 Kalabiska I, Uvacek M, Petrekantits M, et al. Comparison of running performance and prevalence of overweight and obesity in Hungarian and Ukrainian adolescents. *Acta Physiol Hung* 2010;97:393–400.
- 27 Olds TS, Tomkinson GR, Léger L, et al. 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;24:1025–38.
- 28 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* 2008;26:953–7.
- 29 Haugen T, Høigaard R, Seiler S. Normative data of BMI and physical fitness in a Norwegian sample of early adolescents. *Scand J Public Health* 2014;42:67–73.
- 30 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*. Published Online First: 20 May 2016.
- 31 Levy PS, Lemeshow S. Stratification random sampling: further issues. In Levy PS, Lemeshow S, eds. *Sampling of populations: methods and application*. Hoboken, NJ: John Wiley & Sons, 2008:168–75.
- 32 United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects: The 2012 Revision*, DVD Edition. 2013.
- 33 Ng M, Fleming T, Robinson M, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2014;384:766–81.
- 34 Kalman M, Inchley J, Sigmundova D, et al. Secular trends in moderate-to-vigorous physical activity in 32 countries from 2002 to 2010: a cross-national perspective. *Eur J Pub Health* 2015;25(Suppl 2):37–40.
- 35 Weatherbase. <http://www.weatherbase.com/weather/countryall.php3>
- 36 United Nations Development Programme. Human Development Reports. 2011. http://www.undp.org/content/dam/undp/library/corporate/HDR/2011%20Global%20HDR/English/HDR_2011_EN_Complete.pdf
- 37 The World Bank. Gini index (World Bank estimate). <http://data.worldbank.org/indicator/SI.POV.GINI/countries?display=default>
- 38 The World Bank. Urban population (% of total). <http://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?display=default>
- 39 Holm S. A simple sequentially rejective multiple test procedure. *Scand J Statist* 1979;6:65–70.
- 40 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd edn. New Jersey: Lawrence Erlbaum, 1988.

- 41 Aandstad A, Berntsen S, Hageberg R, *et al.* A comparison of estimated maximal oxygen uptake in 9 and 10-year-old schoolchildren in Tanzania and Norway. *Br J Sports Med* 2006;40:278–92.
- 42 Gunnarsson HG, Sigriksson R. *Eurofit: physical fitness of Icelandic pupils at age of 6-15 years old*. Reykjavik, Iceland, 1999.
- 43 Jürimäe T, Volbekiene V, Jürimäe J. Changes in Eurofit test performance of Estonian and Lithuanian children and adolescents (1992-2002). In: Tomkinson GR, Olds TS, eds. *Pediatric fitness: secular trends and geographic variability*. Basel: Karger, 2007:129–42.
- 44 Jürimäe T, Saar M. Self-perceived and actual indicators of motor abilities in children and adolescents. *Percept Mot Skills* 2003;97:862–6.
- 45 Jürimäe T, Volbekiene V. Eurofit test results in Estonian and Lithuanian 11 to 17-year-old children: a comparative study. *Eur J Phys Educ* 1998;3:178–84.
- 46 Raudsepp L, Jürimäe T. Relationships of physical activity and somatic characteristics with physical fitness and motor skill in prepubertal girls. *Am J Hum Biol* 1997;9:513–21.
- 47 Raudsepp L, Jürimäe T. Relationships between somatic variables, physical activity, fitness and fundamental motor skills in prepubertal boys. *Biol Sport* 1996;13:279–89.
- 48 Kull M, Jürimäe T. Using the Eurofit test battery in Estonian 16-18 years old adolescents. *Acta Comment Univ Tartu* 1994;96:49–52.
- 49 Ministry of Education, Culture, Sports, Science and Technology. Report book on the survey of physical fitness and athletic ability. Tokyo: Ministry of Education, Culture, Sports, Science and Technology, 1999–2015.
- 50 Nielsen GA, Andersen LB. The association between high blood pressure, physical fitness, and body mass index in adolescents. *Prev Med* 2003;36:229–34.
- 51 Telama R, Naul R, Nupponen H, *et al.* *Physical fitness, sporting lifestyles and olympic ideals: cross-cultural studies on youth sport in Europe*. Schorndorf, Germany: Verlag Karl Hofmann, 2002.
- 52 Gajda V. *The applications of the UNIFITTEST battery at basic schools (age groups 7-14): summary report*. Ostrava: Pdf OU, 1994.
- 53 Palomäki S, Heikinaro-Johansson P, Huotari P. Cardiorespiratory performance and physical activity in normal weight and overweight Finnish adolescents from 2003 to 2010. *J Sports Sci* 2015;33:588–96.
- 54 Viljanen T, Taimela S, Kujala U. Koululaisten fyysinen aktiivisuus, kestävyyskunto ja painostuskorkeus. *Liikunta ja Tiede* 2000;37:23–6.
- 55 Cazorla G, Dudal J, Dieu JL. État de développement des capacités physiques des jeunes ivoiriens 7-18 ans et plus. Rapport pour le Ministère Ivoirien de la Jeunesse et Sports et pour le Ministère Français de la Coopération. Paris, France: Ministère Ivoirien de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1985.
- 56 Dadouchi MF, Boudiab K, Yahia J, *et al.* De la détermination du profil athlétique marocain au contenu et à l'évaluation. Ministère de l'Éducation nationale et de la jeunesse. École normale supérieure de Marrakech. 2003.
- 57 Baquet G, Twisk JWR, Kemper HCG, *et al.* Longitudinal follow-up of fitness during childhood: interaction with physical activity. *Am J Hum Biol* 2006;18:51–8.
- 58 Baquet G, Berthoin S, Gerbeaux M, *et al.* High-intensity aerobic training during a 10 week one-hour physical education cycle: effect on physical fitness of adolescents aged 11 to 16. *Int J Sports Exerc Med* 2001;22:295–300.
- 59 Baquet G, Berthoin S, Gerbeaux M, *et al.* Assessment of the maximal aerobic speed with the incremental running field tests in children. *Biol Sport* 1999;16:23–30.
- 60 Cazorla G, Portes A, James F. Opération Martinique-Eval. Centre d'Evaluation Sport Santé, Fort de France (Martinique). Rapport pour l'Inspection d'Académie de la Martinique. Fort de France, Martinique: L'Éducation d'Académie de la Martinique, 1997.
- 61 Cazorla G. Batterie France-Éval: Mesures, épreuves et barèmes: Évaluation des qualités physiques des jeunes Français d'âge scolaire: 7-11 ans. Rapport pour le Secrétariat d'État auprès du Premier Ministre Chargé de la Jeunesse et de Sports. Paris: Ministère de la Jeunesse et de Sports, 1987.
- 62 Brunet J, Van Praagh E. *Batterie expérimentale de tests moteurs Eurofit: rapport d'Activité de la région auvergne — 1984–1985*. Clermont-Ferrand, France: Université de Clermont-Ferrand, 1985.
- 63 Cazorla G, Gauthon P, Arémou R, *et al.* État de développement des capacités physiques des jeunes béninois 8-18 ans et plus. Rapport pour le Ministère Béninois de la Jeunesse et Sports et pour le Ministère Français de la Coopération. Paris, France: Ministère Béninois de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1987.
- 64 Cazorla G, Dudal J, Lefrancq L. État de développement des capacités physiques des jeunes Djiboutiens 11-18 ans et plus. Rapport pour le Ministère Djiboutien de la Jeunesse et Sports et pour le Ministère Français de la Coopération. Paris, France: Ministère Djiboutien de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1986.
- 65 Cazorla G, Dudal J, Garinet P. État de développement des capacités physiques des jeunes mauriciens 7-18 ans et plus. Rapport pour le Ministère Mauricien de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1986.
- 66 Krška P, Sedláček J, Hubinák A, *et al.* General motor fitness and somatic parameters comparison between former population and present primary school girls in Ruzomberok. In press.
- 67 Krška P, Sedláček J, Hubinák A, *et al.* Physical development and general motor performance of present primary school boys population in Ruzomberok. In press.
- 68 Kyselovicová O. Programy aerobiku z aspektu rozvoja telesnej zdatnosti dievcat na gymnáziu. *Acta Facult Educ Phys Univers Comen* 2000;41:50–61.
- 69 Kasa J, Majherová M. Physical and motor development of children by Eurofit. *Stud Psychol (Bratisl)* 1997;39:270–4.
- 70 Moravec R. Eurofit—Physique and motor fitness of the Slovak school youth. In: Moravec R, Kampmiller T, Sedláček J, eds. *Eurofit—Physique and motor fitness of the Slovak school youth*. Bratislava: Slovak Scientific Society for Physical Education and Sports, 1996:9–22.
- 71 Belej M, Junger J, Mikus M. Vykonnost ziaikov prijatych na stredne školy vo východoslovenskom regióne zisťovaná systémom EUROFIT. *Tel Vych Sport* 1995;5:12–15.
- 72 Cazorla G, Dudal J, Faye J. État de développement des capacités physiques des jeunes sénégalais 7-18 ans et plus. Rapport pour le Ministère Sénégalais de la Jeunesse et Sports et pour le Ministère Français de la Coopération. Paris, France: Ministère Sénégalais de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1988.
- 73 Veldhuizen S, Cairney J, Hay J, *et al.* Relative age effects in fitness testing in a general school sample: how relative are they? *J Sports Sci* 2015;33:109–15.
- 74 Voss C, Sandercock G, Higgins JW, *et al.* A cross-cultural comparison of body composition, physical fitness, and physical activity between regional samples of Canadian and English children and adolescents. *Can J Public Health* 2014;104:e245–50.
- 75 Leone M, Kalinova E, Comtois AS. *Global motor skill assessment from the UQAC-UQAM test battery: Canadian normative values by age and gender*. Québec: Université du Québec à Chicoutimi, 2011.
- 76 Reed KE, Warburton DER, Whitney CL, *et al.* Secular changes in shuttle-run performance: a 23-year retrospective comparison of 9- to 11-year-old children. *Pediatr Exerc Sci* 2006;18:364–73.
- 77 Massicotte D. Partial curl-ups, push-ups and multistage 20 meter shuttle run, national norms for 6 to 17 year-olds. Final report submitted to Canadian Association for Health, Physical Education and Recreation (CAHPER) and Fitness and Amateur Sport Canada. Montréal: University of Quebec at Montréal, 1990.
- 78 Galiş M, Ergen E, Turnagöl H, *et al.* Beden eğitimi derslerinin bir öğretim yılı boyunca 15-16 yaş grubu öğrenciler üzerindeki fizyolojik etkilerinin Eurofit test bataryası ile izlenmesi. In: 3. Ulusal Spor Bilimleri Kongresi. Ankara, Turkey: Hacettepe Üniversitesi, 1992:367–72.
- 79 Sandercock G, Ogunleye A, Voss C. Six-year changes in body mass index and cardiorespiratory fitness of English schoolchildren from an affluent area. *Int J Obes (Lond)* 2015;39:1504–7.
- 80 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;30:679–87.
- 81 Boddy LM, Fairclough SJ, Atkinson G, *et al.* Changes in cardiorespiratory fitness in 9- to 10.9-year-old children: SportsLinX 1998-2010. *Med Sci Sports Exerc* 2012;44:481–6.
- 82 Sandercock G, Voss C, McConnell D. Ten year secular declines in the cardiorespiratory fitness of affluent English children are largely independent of changes in body mass index. *Arch Dis Child* 2010;95:46–7.
- 83 Liverpool City Council. Liverpool sports linx project 01-03: report on the health and fitness of Liverpool primary and secondary school children. Liverpool, UK: Liverpool City Council, 2003.
- 84 Boreham C, Twisk J, Murray L, *et al.* Fitness, fatness, and coronary heart disease risk in adolescents: the Northern Ireland Young Hearts Project. *Med Sci Sports Exerc* 2001;33:270–4.
- 85 Watkins DC. *Ten year trends (1990-2000) in biological and behavioural risk factors for coronary heart disease in northern Irish adolescents* [thesis]. Belfast, UK: The Queen's University of Belfast, 2001.
- 86 Twisk JWR, Boreham C, Cran G, *et al.* Clustering of biological risk factors for cardiovascular disease and the longitudinal relationship with lifestyle of an adolescent population: the Northern Ireland Young Hearts Project. *J Cardiovasc Risk* 1999;6:355–62.
- 87 Mahoney C. 20-MST and PWC₁₇₀ validity in non-Caucasian children in the UK. *Br J Sports Med* 1992;26:45–7.
- 88 Mahoney CA, Boreham CAG. Eurofit in Belfast primary schools. *Scot J Phys Educ* 1991;19:1–4.
- 89 Riddoch C, Savage JM, Murphy N, *et al.* Long term health implications of fitness and physical activity patterns. *Arch Dis Child* 1991;66:1426–33.
- 90 Nichols AK, Riddoch CJ. The development of fitness test batteries for use in higher education. In: *Trends and developments in physical education: Proceedings of the VIII Commonwealth and International Conference on Sport, Physical Education, Dance, Recreation and Health*; London: E & FN Spon, 1986:378–84.
- 91 Lewitt MS, Baker JS, Mooney GP, *et al.* Pubertal stage and measures of adiposity in British schoolchildren. *Ann Hum Biol* 2012;39:440–7.
- 92 Bronikowski M, Bronikowska M. Salutogenesis as a framework for improving health resources of adolescent boys. *Scand J Public Health* 2009;37:525–31.

Original article

- 93 Pilicz S, Przeweda R, Dobosz J, et al. Punktacja sprawności fizycznej młodzieży Polskiej wg międzynarodowego testu sprawności fizycznej: Kryteria pomiaru wydolności organizmu testem Coopera. Warszawa: Akademia Wychowania Fizycznego Józefa Piłsudskiego, 2003.
- 94 Maciaszek J, Osinski W. Poziom sprawności fizycznej u chłopców i dziewcząt Poznańskich w wieku 10-14 lat. *Rocz Naukowe AWF Poznaniu* 2001;50:3-17.
- 95 Mleczo E, Ozimek M. *Rozwój somatyczny i motoryczny młodzieży Krakowskiej między 15 a 19 rokiem życia z uwzględnieniem czynników środowiskowych*. Kraków: Akademia Wychowania Fizycznego, 2000.
- 96 Lieveid J, Vrijens J, Bouckaert J. Ethnic differences in body structure and physical fitness in Surinamese boys aged 14 years. In: Claessens AL, Lefevre J, Vanden Eynde B, eds. *World-wide variation in physical fitness*. Leuven: Katholieke Universiteit Leuven, 1993:87-92.
- 97 Vandendriessche JB, Vandenborgh BFR, Vaeyens R, et al. Variation in sport participation, fitness and motor coordination with socioeconomic status among Flemish children. *Pediatr Exerc Sci* 2012;24:113-28.
- 98 Seghers J, Rutten C. Clustering of multiple lifestyle behaviours and its relationship with weight status and cardiorespiratory fitness in a sample of Flemish 11- to 12-year olds. *Public Health Nutr* 2010;13:1838-46.
- 99 Verstraete SJ, Cardon GM, De Clercq DL, et al. A comprehensive physical activity promotion programme at elementary school: the effects on physical activity, physical fitness and psychosocial correlates of physical activity. *Public Health Nutr* 2007;10:477-84.
- 100 Cardon G, De Bourdeaudhuij I, De Clercq D, et al. Physical fitness, physical activity, and self-reported back and neck pain in elementary schoolchildren. *Pediatr Exerc Sci* 2004;16:147-57.
- 101 Heyters C, Marique T. *Le baromètre de la condition physique*. Bruxelles: Ministère de la Communauté française, Direction générale du sport, 2004.
- 102 Baquet G, Berthoin S, Padovano C, et al. Effects d'un cycle de course de durée de type intermittent (court-court) sur la condition physique des adolescents. *Rev Educ Phys* 2000;40:51-60.
- 103 Lefevre J, Bouckaert J, Duquet W, et al. De barometer van de fysieke fitheid van de Vlaamse jeugd 1997. De resultaten. *Sport (Blosso Brussel)* 1998;4:16-22.
- 104 Beunen G, Borms J, Vrijens J, et al. Fysieke fitheid en sportbeoefening van de Vlaamse jeugd. Volumen 1: Fysieke fitheid van de jeugd van 6 tot 18 jaar. Brussels: Blosso, 1991.
- 105 Poortmans J, Vlaeminck M, Collin M, et al. [Indirect estimation of the maximal aerobic power of a Male and female population from Brussels aged 6 to 23 years. Comparison with a direct technique for measuring maximal oxygen consumption]. *J Physiol (Paris)* 1986;81:195-201.
- 106 Wang PG, Gong G, Wang SQ, et al. Relationship of body fat and cardiorespiratory fitness with cardiovascular risk in Chinese children. *PLoS ONE* 2011;6:e27896.
- 107 Brouwer SJ, Stol RP, Liem ET, et al. The role of fitness in the association between fatness and cardiometabolic risk from children to adolescence. *Pediatr Diabetes* 2013;14:57-65.
- 108 Slinger J, van Breda E, Kuipers H. Aerobic fitness data for Dutch adolescents (2002-2005). *Pediatr Exerc Sci* 2009;21:10-18.
- 109 van Mechelen W, van Lier WH, Hobbil H, et al. *Eurofit: handleiding met referentietabelen voor 12- tot en met 16-jarige jongens en meisjes in Nederland*. Haarlem: Uitgeverij de Vrieseborch, 1991.
- 110 Gullas-González R, Martínez-Vizcaino V, García-Prieto JC, et al. Excess of weight, but not underweight, is associated with poor physical fitness in children and adolescents from Castilla-La Mancha, Spain. *Eur J Pediatr* 2014;173:727-35.
- 111 Torrijos-Niño C, Martínez-Vizcaino V, Pardo-Guijarro MJ, et al. Physical fitness, obesity, and academic achievement in schoolchildren. *J Pediatr* 2014;165:104-9.
- 112 Chillón P, Ortega FB, Ferrando JA, et al. Physical fitness in rural and urban children and adolescents from Spain. *J Sci Med Sport* 2011;14:417-23.
- 113 Castro-Pinero J, González-Montesinos JL, Mora J, et al. Percentile values for muscular strength field tests in children aged 6 to 17 years: influence of weight status. *J Strength Cond Res* 2009;23:2295-310.
- 114 Ortega FB, Ruiz JR, Castillo MJ, et al. Low level of physical fitness in Spanish adolescents. Relevance for future cardiovascular health (AVENA Study). *Rev Esp Cardiol* 2005;58:898-909.
- 115 García Baena J. *La condición física en la educación secundaria. Trabajo de investigación* [thesis]. Madrid, Spain: Universidad Nacional de Educación a Distancia, 1999.
- 116 Prat JA, Casamort J, Balagué N, et al. *Eurofit: la batería Eurofit en Catalunya*. Barcelona: Secretaria General de l'Esport, 1998.
- 117 Tercedor P, Delgado-Fernandez M. Condición física relacionada con la salud en escolares de 10 años de edad de Granada. *Proceedings of the II congreso internacional sobre la enseñanza de la educación física y el deporte escolar (second international congress about teaching physical education and school sport)*; 1998.
- 118 Sainz RM. *La batería Eurofit en Euskadi*. Vitoria-Gasteiz: Instituto Vasco de Educación Física, 1996.
- 119 Ureña F. *Valoración y baremación de la aptitud física en el alumnado de segundo ciclo de educación secundaria obligatoria de la comunidad autónoma de Murcia. Su utilización según los postulados de la reforma* [dissertation]. Murcia, Spain: Universidad de Murcia, 1996.
- 120 Brito Ojeda EM, Navarro Valdivielso M, García Afonso D, et al. *La condición física en la población escolar de gran Canaria (10-19 años)*. Las Palmas de Gran Canaria, Spain: Excmo. Cabildo Insular de Gran Canaria, 1995.
- 121 Sainz RM. *Aptitudes psíquicas y físicas: Estudio de la aptitud física de los adolescentes de la provincia de Vizcaya y su relación con la personalidad* [dissertation]. Bilbao, Spain: Universidad de Deusto, 1992.
- 122 Rivas FJ. Valoración de la evolución antropométrica y de las características motrices en la población escolar de un centro de E.G., B. mediante un estudio transversal. In II Congreso Galego da Educación Física e o Deporte. La Coruña: Escola Galega do Deporte de la Xunta de Galicia, 1987.
- 123 Pienaar C, Coetzee B, Monyeki AM. The use of anthropometric measurements and the influence of demographic factors on the prediction of VO2max in a cohort of adolescents: the PAHL study. *Ann Hum Biol* 2015;42:135-43.
- 124 Du Toit D, Pienaar AE, Truter L. Relationship between physical fitness and academic performance in South African children. *S Afr J Res Sport PH* 2011;33:23-35.
- 125 Pienaar AE, Viljoen A. Physical and motor ability, anthropometrical and growth characteristics of boys in the northwest province of South Africa: a sport talent perspective. *S Afr J Res Sport Phys Educ Recreation* 2010;32:71-93.
- 126 Du Preez SM. *The effect of physical activity on the body composition and health related fitness of 9 to 13-year-old boys* [thesis]. Potchefstroom, South Africa: North-West University, 2008.
- 127 Stadler MC. *The influence of a physical activity intervention program (PAI) on the physical fitness levels, body composition and health risk behaviour of 9 to 13-year-old girls* [dissertation]. Potchefstroom, South Africa: North-West University, 2007.
- 128 Du Toit L, Venter RE, Potgieter JR. The relationship between cardiorespiratory fitness, body composition and physical self-perception of adolescent girls. *J Hum Mov Stud* 2005;48:353-64.
- 129 Van Gent M, Malan DDJ, Pienaar AE. A comparison of the anthropometric, physical and motor growth characteristics of 12-15 year old girls in the North West province with Australian girls. *Afr J Phys Health Educ Recr Dance* 2002;8:309-20.
- 130 Du Randt R. The 1996 South African sport talent identification project: report prepared for the Sports Information and Science Agency. Pretoria: South African Sports Commission, 1996.
- 131 Schmid M, Romann M, Kriemler S, et al. Wie kann die Fitness von Schulkindern gemessen werden? *Sportmed Sporttraumatol* 2007;55:52-61.
- 132 Cauderay M, Narring F, Michaud P-A. A cross-sectional survey assessing physical fitness of 9- to 19-year-old girls and boys in Switzerland. *Pediatr Exerc Sci* 2000;12:398-412.
- 133 Richards J, Foster C, Townsend N. Physical fitness and mental health impact of a sport-for-development intervention in a post-conflict setting: Randomised controlled trial nested within an observational study of adolescents in Gulu, Uganda. *BMC Public Health* 2014;14:619.
- 134 Bovet P, Auguste R, Burdette H. Strong inverse association between physical fitness and overweight in adolescents: a large school-based survey. *Int J Behav Nutr Phys Act* 2007;4:24.
- 135 Cazorla G, Rousseau G, Dudal J, et al. Évaluation des capacités motrices de l'enfant, de l'adolescent et du jeune sénégalais: 7-18 ans et plus. Rapport pour le Ministère Sénégalais de l'Éducation et pour le Ministère Français de la Coopération. Paris, France: Ministère Sénégalais de l'Éducation et Ministère Français de la Coopération, 1990.
- 136 Cubides RC, Alarcón LGA, Galvi ARG. Diferencias en la actividad física y la condición física entre los escolares de secundaria de dos programas curriculares oficiales de Bogotá, Colombia. *Nutr Hosp* 2015;32:2228-34.
- 137 Fernandez-Ortega JA. Estudio transversal de las cualidades funcionales de los escolares bogotanos: Valores de potencia aeróbica, potencia muscular, velocidad de desplazamiento y velocidad de reacción, de los siete a los dieciocho años. *Rev Edu Fis Deportes* 2013;32:1151-70.
- 138 Mojica GT, Poveda JG, Pinilla MI, et al. Sobre peso, inactividad física y baja condición física en un colegio de Bogotá, Colombia. *Arch Latinoam Nutr* 2008;58:265-73.
- 139 Hardy LL, King L, Espinel P, et al. NSW Schools Physical Activity and Nutrition Survey (SPANS) 2010: full report. Sydney: NSW Ministry of Health, 2010.
- 140 McIntyre S. *Trends in aerobic fitness from 1994 to 2009 in 10- and 11-year-old Australian children* [thesis]. Notre Dame, IN: University of Notre Dame, 2009.
- 141 Booth M, Okely AD, Denney-Wilson E, et al. NSW Schools Physical Activity and Nutrition Survey (SPANS) 2004: full report. Sydney: NSW Department of Health, 2006.
- 142 Lloyd KC, Antonas KN. Nutritional habits and fitness levels of schoolchildren. *Proceedings of the Nutrition Society of Australia—Twenty-fourth annual scientific meeting, Fremantle, WA, 3-6 December*; Adelaide, SA: Nutrition Society of Australia, 2000:138.
- 143 Cooley D, McNaughton L. Aerobic fitness of Tasmanian secondary school children using the 20-m shuttle run test. *Percept Mot Skills* 1999;88:188-98.
- 144 Booth M, Macaskill P, McLellan L, et al. *NSW schools fitness and physical activity survey 1997*. Sydney: NSW Department of Education and Training, 1997.

- 145 Okely AD, Gray T, Cotton WG. Effect of an extended stay outdoor education program on aerobic fitness. In Gray T, Haylar B, eds. *Catalysts for change: proceedings from the 10th national outdoor education conference*. Sydney, Australia, 1997:206–10.
- 146 Australian Council for Health, Physical Education and Recreation. *Australian fitness education award: user's manual and curriculum ideas*. Adelaide, SA: Australian Council for Health, Physical Education and Recreation, 1996.
- 147 Vandongen R, Jenner DA, Thompson C, et al. A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10- to 12-year-old children. *Prev Med* 1995;24:9–22.
- 148 Australian Sports Commission. *Norms for sport related fitness tests in Australian students aged 12-17 years*. Belconnen, ACT: Australian Sports Commission, 1994.
- 149 Jenner DA, Vandongen R, Beilin L. Relationships between blood pressure and measures of dietary energy intake, physical fitness, and physical activity in Australian children aged 11-12 years. *J Epidemiol Community Health* 1992;46:108–13.
- 150 Welk GJ, Saint-Maurice PF, Csányi T. Health-related physical fitness in Hungarian youth: age, sex, and regional profiles. *Res Q Exerc Sport* 2015;86(Suppl 1):S45–57.
- 151 Barabás, A. Measurement of aerobic power by field tests. In: Coudert J, Van Praagh E, eds. *Pediatric work physiology: children and exercise XVI*. Paris: Masson, 1992:39–41.
- 152 Grassi GP, Turci M, Sforza C. Aerobic fitness and somatic growth in adolescents: a cross sectional investigation in a high school context. *J Sports Med Phys Fitness* 2006;46:412–18.
- 153 Cilia G, Bazzano C, Bellucci M, et al. I risultati dei test Eurofit nella scuola Matteucci di Roma. *Alcmeone* 1998;2:16–20.
- 154 Cilia G, Bellucci M, Bazzano C, et al. Eurofit 1997: Banche dati per la scuola. *Alcmeone* 1997;3:13–32.
- 155 Cilia G, Bellucci M, Riva M. *Eurofit 1995*. Roma: Istituto Superiore Statale di Educazione Fisica, 1996.
- 156 Cilia G, Bellucci M. *Eurofit: tests Europei di attitudine fisica*. Roma: Istituto Superiore Statale di Educazione Fisica, 1993.
- 157 Council of Europe. *Évaluation de l'aptitude physique: Eurofit batterie expérimentale*. Rome: Council of Europe, 1986.
- 158 Wong TW, Yu TS, Wang XR, et al. Predicted maximal oxygen uptake in normal Hong Kong Chinese schoolchildren and those with respiratory diseases. *Pediatr Pulmonol* 2001;31:126–32.
- 159 Kain J, Uauy R, Albala, et al. School-based obesity prevention in Chilean primary school children: methodology and evaluation of a controlled study. *Int J Obes Relat Metab Disord* 2004;28:483–93.
- 160 Secchi JD, García GC, España-Romero V, et al. Physical fitness and future cardiovascular risk in Argentine children and adolescents: an introduction to the ALPHA test battery. *Arch Argent Pediatr* 2014;112:132–40.
- 161 Gonzalez-Suarez CB, Grimmer-Sommers K. The association of physical activity and physical fitness with pre-adolescent obesity: an observational study in Metromanila, Philippines. *J Phy Act Health* 8:804–10.
- 162 Tambalis KD, Panagiotakos DB, Psarra G, et al. Physical fitness normative values for 6-18-year-old Greek boys and girls, using the empirical distribution and the lambda, mu, and sigma statistical method. *Eur J Sport Sci* 2016;16:736–46.
- 163 Tambalis K, Panagiotakos D, Sidossis L. Greek children living in rural areas are heavier but fitter compared to their urban counterparts: a comparative, time-series (1997-2008) analysis. *J Rural Health* 2011;27:270–7.
- 164 Tokmakidis SP, Kasambalis A, Christodoulos AD. Fitness levels of Greek primary schoolchildren in relationship to overweight and obesity. *Eur J Pediatr* 2006;165:867–74.
- 165 Manios Y, Yiannakouris N, Papoutsakis C. Behavioral and physiological indices related to BMI in a cohort of primary schoolchildren in Greece. *Am J Hum Biol* 2004;16:639–47.
- 166 Georgiadis G. Evaluation of physical fitness of Greek youth aged 6-18 years [dissertation]. Athens, Greece: University of Athens, 1993.
- 167 Santos R, Mota J, Santos DA, et al. Physical fitness percentiles for Portuguese children and adolescents aged 10-18 years. *J Sports Sci* 2014;32:1510–81.
- 168 Coelho-Silva MJ, Ronque ERV, Cyrino ES, et al. Nutritional status, biological maturation and cardiorespiratory fitness in Azorean youth aged 11-15 years. *BMC Public Health* 2013;13:495.
- 169 Rodrigues LP, Leitao R, Lopes VP. Physical fitness predicts adiposity longitudinal changes over childhood and adolescence. *J Sci Med Sport* 2013;16:118–23.
- 170 Marta CC, Marinho DA, Barbosa TM. Physical fitness differences between prepubescent boys and girls. *J Strength Cond Res* 2012;26:1756–66.
- 171 Silva G, Aires L, Mota J, et al. Normative and criterion-related standards for shuttle run performance in youth. *Pediatr Exerc Sci* 2012;24:157–69.
- 172 Mota J, Guerra S, Leandro C. Association of maturation, sex, and body fat in cardiorespiratory fitness. *Am J Hum Biol* 2002;14:707–12.
- 173 Ribeiro RR, Santos KD, Carvalho WCG, et al. Aerobic fitness and biological and sociodemographic indicators in female school children. *Rev Bras Cineantropom Desempenho Hum* 2013;15:448–57.
- 174 Hobold E. *Indicadores de aptidão física relacionada à saúde de crianças e adolescentes do município de Marechal Cândido Rondon* [dissertation]. Paraná, Brazil: Universidade Federal de Santa Catarina, 2003.
- 175 Pieta S. *Estudio de la aptitud física de una muestra de la población escolar del estado de Paraná mediante la batería Eurofit* [dissertation]. León, Brazil: Universidad de León, 2000.
- 176 Shahin A. Ying and yang of body composition assessment. *Chin J Integr Med* 2011;17:675–9.
- 177 Tinazci C, Emiroglu O. Assessment of physical fitness levels, gender and age differences of rural and urban elementary school children. *Turkiye Klinikleri J Med Sci* 2010;30:1–7.
- 178 Kim H-B, Stebbins CL, Chai J-H, et al. Taekwondo training and fitness in female adolescents. *J Sports Sci* 2011;29:133–8.
- 179 Welk GJ, Saint-Maurice PF, Allums-Featherston K, et al. Design and Evaluation of the NFL PLAY 60 FITNESSGRAM Partnership Project. *Res Q Exerc Sport* 2016;87:1–13.
- 180 Carrel AL, Bowser J, White D, et al. Standardized childhood fitness percentiles derived from school-based testing. *J Pediatr* 2012;161:120–4.
- 181 Liu W, Zillifro TD, Nichols RA. Tracking of health-related physical fitness for middle school boys and girls. *Pediatr Exerc Sci* 2012;24:549–62.
- 182 Welk GJ, De Saint-Maurice Maduro PF, Laurson KR, et al. Field evaluation of the new FITNESSGRAM® criterion-referenced standards. *Am J Prev Med* 2011;41 (Suppl 2):S131–42.
- 183 Beets MW, Pitetti KH, Cardinal BJ. Progressive aerobic cardiovascular endurance run and body mass index among an ethnically diverse sample of 10-15-year-olds. *Res Q Exerc Sport* 2005;76:389–97.
- 184 Beets MW, Pitetti KH. A comparison of shuttle-run performance between mid-western youth and their national and international counterparts. *Pediatr Exerc Sci* 2004;16:94–112.
- 185 Welk GJ, Schaben JA, Shelley M. Physical activity and physical fitness in children schooled at home and children attending public schools. *Pediatr Exerc Sci* 2004;16:310–23.
- 186 Lloyd LK, Bishop PA, Walker JL, et al. The influence of body size and composition on FITNESSGRAM® test performance and the adjustment of FITNESSGRAM® test scores for skinfold thickness in youth. *Meas Phys Educ Exerc Sci* 2003;7:205–26.
- 187 Chun DM, Corbin CB, Pangrazi RP. Validation of criterion-referenced standards for the mile run and progressive aerobic cardiovascular endurance tests. *Res Q Exerc Sport* 2000;71:125–34.
- 188 Wolford N. *The difference in physical fitness levels of fifth graders according to socioeconomic groups and genders* [dissertation]. Ann Arbor, MI: University Microforms International, 1998.
- 189 Mahar MT, Rowe DA, Parker CR, et al. Criterion-referenced and norm-referenced agreement between the mile run/walk and PACER. *Meas Phys Educ Exerc Sci* 1997;1:245–58.
- 190 Sauka M, Priedite IS, Artjuhova L, et al. Physical fitness in Northern European youth: reference values from the Latvian physical health in youth study. *Scand J Public Health* 2011;39:35–43.
- 191 Cossio-Bolaños MA, Arruda M. Propuesta de valores normativos para la evaluación de la aptitud física en niños de 6 a 12 años de Arequipa, Perú. *Rev Med Herediana* 2009;20:206–12.
- 192 Galaviz KI, Tremblay MS, Colley R, et al. Associations between physical activity, cardiorespiratory fitness, and obesity in Mexican children. *Salud Pública Méx* 2012;54:463–9.
- 193 Ortega FB, Ruiz JR, Labayen I, et al. Health inequalities in urban adolescents: role of physical activity, diet, and genetics. *Pediatrics* 2014;133:e884–95.
- 194 Cooper AR, Goodman A, Page AS, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometer database (ICAD). *Int J Behav Nutr* 2015;15:113.
- 195 Guthold R, Cowan MJ, Autenrieth CS, et al. Physical activity and sedentary behavior among schoolchildren: a 34-country comparison. *J Pediatrics* 2010;157:43–9.
- 196 Currie C, Zanotti C, Morgan D, et al, eds. Social determinants of health and well-being among young people. Health behaviour in school-aged children (HSBC) study: international report from the 2009/2010 survey. Copenhagen, WHO Regional Office for Europe, 2012: Health Policy for Children and Adolescents, No. 6.
- 197 Malina R, Bouchard C, Bar-Or O, eds. *Growth, maturation, and physical activity*. 2nd edn. Champaign, IL: Human Kinetics, 2004.
- 198 Sallis JF, Owen N. Ecological models of health behaviour. In: Glanz K, Lewis FM, Rimer BK, eds. *Health behaviour and health education: theory, research, and practice*. 3rd edn. San Francisco, CA: Jossey-Bass, 2002:462–84.
- 199 Parikh T, Stratton G. Influence of intensity of physical activity on adiposity and cardiorespiratory fitness in 5-18-year-olds. *Sports Med* 2011;41:477–88.
- 200 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.
- 201 Ekelund U, Tomkinson GR, Armstrong N. What proportion of youth are physically active? Measurement issues, levels and recent time trends. *Br J Sports Med* 2011;45:859–65.

Original article

- 202 World Health Organization. Global School-based Student Health Survey (GSHS) website. <http://www.who.int/chp/gshs/en>
- 203 Roberts C, Currie C, Samdal O, *et al.* Measuring the health and health behaviours of adolescents through cross-national survey research: recent developments in the Health Behaviour in School-aged Children (HBSC) study. *J Pub Health* 2007;15:179–86.
- 204 Hallal PC, Andersen LB, Bull FC, *et al.* Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet* 2012;380:247–57.
- 205 Katzmarzyk PT, Barreira TV, Broyles ST, *et al.* Physical activity, sedentary time, and obesity in an international sample of children. *Med Sci Sports Exerc* 2015;47:2062–9.
- 206 United Nations. United Nations Expert Group Meeting on Population Distribution, Urbanization, Internal Migration & Development, New York, 21–23 January 2008. New York: United Nations, 2008.
- 207 Omran AR. The epidemiological transition. A theory of the epidemiology of population change. *Milbank Mem Fund Q* 1971;49:509–38.
- 208 Katzmarzyk PT, Mason C. The physical activity transition. *J Phys Act Health* 2009;6:269–80.
- 209 Viner RM, Ozer EM, Denny S, *et al.* Adolescence and the social determinants of health. *Lancet* 2012;379:1641–52.
- 210 Elgar FJ, Pfortner TK, Moor I, *et al.* Socioeconomic inequalities in adolescent health 2002–2010: a time-series analysis of 34 countries participating in the Health Behaviour in School-aged Children study. *Lancet* 2015;385:2088–95.
- 211 Lussier L, Buskirk ER. Effects of an endurance training regimen of assessment of work capacity in prepubertal children. *Ann N Y Acad Sci* 1977;301:734–47.
- 212 Krahenbuhl GS, Pangrazi RP, Chomokos EA. Aerobic responses of young boys to submaximal running. *Res Q* 1979;50:413–21.
- 213 Sjödin B. The relationship among running economy, aerobic power, and onset of blood lactate accumulation in young boys (11–15 years). In: Komi PV, ed. *Exercise and sport biology*. Champaign, IL: Human Kinetics, 1982:57–60.
- 214 Berenson GS, Srinivasan SR, Boa W, *et al.* Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The Bogalusa Heart Study. *N Engl J Med* 1998;338:1650–6.
- 215 Lavie CJ, De Schutter A, Milani RV. Health obese versus unhealthy lean: the obesity paradox. *Nat Rev Endocrinol* 2015;11:55–62.
- 216 Myers J, McAuley P, Lavie CJ, *et al.* Physical activity and cardiorespiratory fitness as major markers of cardiovascular risk: their independent and interwoven importance to health status. *Prog Cardiovasc Dis* 2015;57:306–14.

CHAPTER 7: MANUSCRIPT 5

This manuscript has been written and formatted following the publication standards of the peer-reviewed journal, *Exercise and Sport Sciences Reviews*.

Title:

Making a case for cardiorespiratory fitness surveillance among children and youth

Status:

Submitted as an invited review under the ‘Perspectives for Progress’ category

Title

Making a case for cardiorespiratory fitness surveillance among children and youth

Authors' names

Justin J. LANG¹, Grant R. TOMKINSON^{2,3}, Ian JANSSEN⁴, Jonatan R. RUIZ^{5,6}, Francisco B. ORTEGA^{5,6}, Luc Léger⁷, & Mark S. TREMBLAY^{1*}

Institutional affiliations

¹Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada.

²Department of Kinesiology and Public Health Education, University of North Dakota, Grand Forks, ND, USA.

³Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health Sciences & Sansom Institute for Health Research, University of South Australia, Adelaide, SA, Australia.

⁴School of Kinesiology and Health Studies, Queen's University, Kingston, ON, Canada

⁵PROMoting FITness and Health through physical activity research group (PROFITH), Department of Physical Education and Sports, Faculty of Sport Sciences, University of Granada, Granada, Spain.

⁶Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden.

⁷Département de kinésiologie, Université de Montréal, Montreal, Quebec, Canada.

Short title

Cardiorespiratory Fitness Surveillance

Corresponding author

Mark S. TREMBLAY*

✉ Healthy Active Living and Obesity Research Group
Children's Hospital of Eastern Ontario Research Institute
401 Smyth Rd
Ottawa, ON, K1H 8L1, Canada
☎ +1 613 737-7600 x4114
📧 mtremblay@cheo.on.ca

Disclosure of funding: None received for this work.

Conflicts of interest: None reported.

Total character count: 62,286

Abstract

We review the evidence that supports cardiorespiratory fitness (CRF) as an important indicator of current and future health among school-aged children and youth, independent of physical activity levels. We discuss the merit of CRF measurement for population health surveillance and propose the development of CRF guidelines to help support regional, national, and international surveillance efforts.

Summary

The assessment of cardiorespiratory fitness among children and youth is an important and feasible measure for population health surveillance.

Key Points

- *The measurement of cardiorespiratory fitness (CRF) provides a feasible way to study the link between physical activity and health among children and youth at the population level.*
- *The surveillance of CRF is a favourable approach to understand and monitor the health of children and youth within and between countries.*
- *The measurement of CRF could provide novel surveillance opportunities in scientific research, physical education and clinical settings.*
- *The development of CRF guidelines could help support future surveillance efforts at the national and international levels.*

Key Words

Monitor, Physical Activity, Aerobic, Pediatric, Health, Populations, Public Health

INTRODUCTION

Worldwide, the average life expectancy at birth increased from 61.7 years in 1980 to 71.8 years in 2015 (1). This improvement has not been matched with equivalent gains in healthy life expectancy (i.e., life expectancy in good health without non-communicable disease [NCD] and/or disability) creating a situation where individuals are now living longer with comorbidities and disabilities (2). This reality and resulting need for costly treatments has created financial strain on health care systems and economies across the world. The World Economic Forum has estimated that current NCD projections will result in a cumulative US\$47 trillion in lost economic production and health care costs worldwide between 2011 and 2030 (3). Physical activity (PA) and cardiorespiratory fitness (CRF) represent important modifiable risk factors that have been linked to preventing the global burden of NCD from coronary heart disease, type 2 diabetes, breast and colon cancer (4–6). As a result, there is a need to shift toward a cost effective, population-based, NCD prevention paradigm through effective interventions and policies that promote healthy active living.

Research demonstrates that many NCD risk factors manifest in childhood (7), suggesting that NCD prevention initiatives should begin during these years. In several countries there is evidence of temporal declines in certain PA domain (i.e., active transport, physical education, and organized sports [8–10]), PA across the life course (11), CRF levels (12), strength (13), and increases in sedentary behaviors and screen time (14) among children and youth over recent decades — supporting the need for health promoting policies and interventions to help reverse these deleterious trends. It has been argued that a key component of a national strategy to deal with NCD burden is a sustainable population health surveillance system that can quantify and monitor NCD risk and the associated determinants (15). Surveillance systems are especially important in guiding and evaluating health promoting policies and interventions.

Among children and youth, the surveillance of PA levels has emerged as a primary measure to track and compare pediatric health across regions, countries and cities (16). It is possible that the elevation of PA to a primary surveillance indicator was partly attributable to the development and distribution of PA guidelines that provided recommendations on the amount of PA needed for health promotion. While the importance of PA has been recognized in surveillance efforts, the same cannot be said for CRF, despite the fact that the CRF level of children and youth is an important indicator of current (17) and future health (18), independent of PA (19). Indeed, it has been suggested that CRF should be included as a measure for national and international health surveillance efforts in young individuals, similar to the surveillance of PA (20–22). Thus, the objective of this review was to examine the utility of the 20m shuttle run test (20mSRT) — a feasible, valid (23–25) and reliable (26–28) field-based measure of CRF in children and youth — as a population health surveillance indicator. We start by discussing the importance of CRF as a measure of health among children and youth. We then discuss novel strategies for CRF surveillance, including the importance of establishing guidelines as an effective knowledge translation tool at the population level. Finally, we provide recommendations that may help progress the international utility of CRF as a population health surveillance indicator.

CARDIORESPIRATORY FITNESS AND PHYSICAL ACTIVITY

In 1994, Bouchard and Shephard published a conceptual model that described the relationships among health-related fitness, PA and health for adults (29). Our modified model for children and youth, as illustrated in Figure 1, describes PA as a behavior that takes place during active play, organized sport, active transportation, and school-based activities. PA behaviors are closely interconnected with a variety of health-related fitness traits that include but is not limited to CRF (i.e., morphological, muscular, motor, and metabolic). The strength of the relationship between

habitual PA and CRF has often been characterized as small to moderate, although the strength of the relationship improves when higher PA intensities are isolated (30, 31). Both PA and CRF are influenced by heredity, other lifestyle behaviors such as diet, personal attributes such as age and gender, and exposures to physical and social environments.

<<< Insert Figure 1 here >>>

Among adults, the surveillance of wellness, morbidity and mortality provide important health outcomes, and allow for vital data to help inform and guide healthy public policy. However, among children and youth, the surveillance of lifestyle-related NCDs (e.g., type 2 diabetes, ischemic heart disease) and related deaths are not common practice because these are rare conditions (32). Rather, surveillance in the pediatric population focuses on risk factors for NCDs, such as obesity, and their upstream determinants, such as PA and to a lesser extent, CRF.

Note: the remainder of this section is not meant to be an in-depth review of measurement, but rather a broad overview of how measurement relates to population-based surveillance.

Cardiorespiratory Fitness Measurement

CRF is described as a trait that represents the ability to deliver oxygen to the muscles and to utilise it to produce energy to support muscle activity during physical activity and exercise (20,33). The measurement of CRF has an extensive history and is likely one of the most widely reported physiological variables in the field of pediatric exercise science (34–36), yet our understanding of the measurement of CRF remains incomplete. Most measures of CRF are objective for children and youth, with only one known subjective measure of CRF (i.e., International Fitness Scale [IFIS] [37, 38]). These objective CRF measures can be broadly categorized into lab- and field-based assessments (26).

Lab-based assessments of CRF using indirect calorimetry provide a valid measure of peak oxygen uptake ($\dot{V}O_{2\text{peak}}$), which is considered the gold standard measure of CRF in children and youth. In young populations $\dot{V}O_{2\text{peak}}$ is often used instead of $\dot{V}O_{\text{max}}$ to describe a maximal effort because these populations don't often attain the same plateau in oxygen uptake seen adults (39). The measurement of $\dot{V}O_{2\text{peak}}$ is an important physiological variable that can be calculated in absolute terms as $\text{L} \cdot \text{min}^{-1}$ values, or in relative terms, typically controlled for body mass and calculated as $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ values (40). Both calculations produce varying results, however, controlling for body mass seems to be preferred given that this method is a stronger predictor of health, although fat mass may confound the interpretation of CRF (41,42). Among children and youth the use of allometric scaling has also been proposed as a favourable method (43). Since lab-based assessments have many limitations, such as the need for expensive equipment, specialized human resources, high participant burden, and the time consuming nature of the assessments, they are not typically feasible for population-based surveillance.

A variety of field-based assessments have been developed to predict $\dot{V}O_{2\text{peak}}$ (e.g., distance runs, endurance shuttle runs), and these field-based assessments provide reasonably accurate estimates (25, 26). Of the field-based CRF assessments, the 20mSRT is the most common assessment used worldwide. Originally developed in the 1980s by Léger et al. (44,45), the 20mSRT has since been used to assess CRF in children and youth across more than 50 countries (12, 26, 46–48), and over 3 million kids in the US alone (49). In the 20mSRT participants are asked to run back and forth between shuttles (i.e., two parallel lines 20 m apart) following a repeated audio signal beep that begins at pace of $8.5 \text{ km} \cdot \text{h}^{-1}$ and increases by $0.5 \text{ km} \cdot \text{h}^{-1}$ every stage (i.e., 1-minute duration) (44,45). The number of shuttles a child is able to complete provides an estimate of their CRF. The 20mSRT assessment is appropriate for population-based surveillance for a number of reasons

including the flexibility of testing environments (indoor, outdoor, small testing area), simplicity of testing procedures, easily interpretable results, no need for specialized human resources, low cost, and ability to test large groups of children simultaneously with relatively low participant burden (18, 26). The 20mSRT also demonstrates a moderate-to-high criterion-related validity (23–25) and high reliability (26–28), suggesting it is a good field-based measure of CRF. The major scientific concern of the 20mSRT is that it does not provide a direct and isolated assessment of $\dot{V}O_{2\text{peak}}$, but rather, performance on the test is likely a combination of multiple factors including running economy, $\dot{V}O_2$ kinetics, fractional utilization of oxygen, fat mass, lactate threshold, anaerobic capacity, and psychosocial factors such as motivation, self-efficacy, and pacing (26). This suggests that lab- and field-based assessments measure slightly different aspects of the CRF construct. Thus, researchers have debated whether 20mSRT performance values should be reported as raw values (i.e., laps, stages), or converted to $\dot{V}O_{2\text{peak}}$ ($\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) given that the prediction equation applied introduces additional error to the estimate (24, 50, 51). To our knowledge, there are at least five published equations available to predict $\dot{V}O_{2\text{peak}}$ using 20mSRT scores (45, 52–55; Table 1), all of which differ in validity and would produce different $\dot{V}O_{2\text{peak}}$ estimates, creating a need for consensus on a standardized equation in this field. However, as discussed below, the decision to use raw values instead of predicted $\dot{V}O_{2\text{peak}}$ values has implications related to health- and performance-related fitness.

<<< Insert Table 1 here >>>

The use of raw 20mSRT values (i.e., shuttles, stages, running speed [$\text{km}\cdot\text{h}^{-1}$] at the last completed stage) provides the best estimate of CRF-performance, and raw values have been widely used in the development of normative-referenced standards (47, 56–63). The use of normative-referenced standards to provide an interpretation of 20mSRT scores is a reflection of performance-related fitness. Normative-referenced standards provide a relative interpretation of 20mSRT performance

by determining centile ranks in comparison with peers of the same age and sex (64), making it is possible to identify high, average, and/or low 20mSRT performance. The primary limitation of this approach is that normative-referenced standards are not linked to a health outcome. In contrast, criterion-referenced standards identify the minimum amount of CRF needed for good health, with several different standards emerging in the literature over the past 10 years (64–69). Although criterion-referenced standards represent a way to identify the percentage of healthy children and youth in a population, a major limitation is that these standards do not readily allow for the use of raw 20mSRT performance scores. The reason for this limitation is that the majority of criterion-referenced standards have been developed using a variety of lab-based assessments of direct and indirect measured $\dot{V}O_{2\text{peak}}$, resulting in the standards almost always being reported as $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ values. Thus, there is a need to convert 20mSRT scores to a common metric, $\dot{V}O_{2\text{peak}}$, or the standards themselves must be converted to 20mSRT raw values. In essence, health- and performance-related fitness have emerged as two separate ideologies on how to interpret CRF results. However, the use of normative- and criterion-referenced standards is not mutually exclusive, and in fact, these approaches complement each other well. Thus, both methods could potentially co-exist, and should be reported, in population-based CRF surveillance efforts.

HEALTH INDICATORS AMONG CHILDREN AND YOUTH

The most important feature of a health indicator is that it is favourably and substantially associated with health variables of interest (e.g., mental, emotional, cognitive, psychosocial, physiological, and physical health). Ideally, a single measure at the population level could provide an aggregate measure of pediatric population health. In addition to current health status, it is important for health indicators to be predictive of future health. This is especially important in

pediatric populations because it could help inform and prepare jurisdictions for the potential future burden of disease.

Cardiorespiratory Fitness and Health

CRF levels are regarded as an important health indicator (17). It is noteworthy that CRF is meaningfully associated with health, independent of PA levels, in children and youth (19, 32), suggesting that the measurement of CRF provides additional insight into the health of kids above which can be provided by PA alone. Recent research has identified that high levels of CRF are favourably associated with healthy cardiovascular and metabolic profiles among children and youth (66, 70–72). The strength of the associations appears to be stronger for CRF when compared with objective measures of PA levels in children and youth (73, 74). Similarly, several studies have shown considerably stronger inverse relationships between CRF and mortality than between PA and mortality among adult populations (75, 76). Overall, the measurement of CRF seems to provide a good summative indication of the ability of numerous bodily systems and organs (e.g., cardiorespiratory, musculoskeletal, endocrine-metabolic, psycho-neurological, hematocirculatory) involved in human movement.

The findings discussed in the previous paragraph used a variety of lab- and field-based methods to measure CRF. A recent systematic review of literature focused on the associations between 20mSRT performance and health indicators among school-aged children and youth (77). The review identified 142 studies, representing over 300,000 children and youth from 32 countries. The results indicated that 20mSRT performance was favorably associated with indicators of cardiometabolic, anthropometric, cognitive, and psychosocial health. Furthermore, there was limited, but consistent evidence supporting favorable associations between 20mSRT performance and self-esteem, quality of life, and mental health (77). The systematic review concluded that the

20mSRT is an important indicator of holistic health among children and youth. Those who perform better on the 20mSRT have better overall health as compared to those who do not perform well.

Cardiorespiratory fitness and future health

A systematic review of 42 longitudinal studies that assessed the relationship between fitness levels in children and future health has provided further insight into the importance of CRF (18). The review concluded that there was strong evidence suggesting that high levels of CRF in childhood predicts a healthy cardiovascular risk profile (i.e., lower levels of blood lipids, blood pressure, and overall and central adiposity) in adulthood. There was moderate evidence supporting CRF levels in childhood as a predictor of future metabolic syndrome and arterial stiffness. Lastly, moderate evidence supported increases in CRF being inversely associated with changes in blood lipids and lipoproteins in adulthood. Furthermore, these results have been supported by a large Swedish cohort study of over 700,000 adolescents with a median follow-up of 29 to 34 years, which identified that low CRF in adolescence was significantly associated with an increased risk of myocardial infarction (76) and early mortality (77) in adulthood. Together, these findings support the notion that CRF levels among children and youth are not only an indicator of current health, but also a potential indicator of future population disease burden.

Cardiorespiratory fitness standards

As described above, there are two different types of standards used to interpret CRF results — criterion-referenced and normative-referenced. Of these, criterion-referenced standards provide a way to determine the amount of CRF associated with good health (e.g., healthy cardiometabolic profile). Recently, a review of criterion-referenced standards for CRF identified 10 different sets

of standards for children and youth, all of which produced different percentage estimates of children and youth with healthy CRF (64). Consequently, it was recommended that the results from a meta-analysis of seven criterion-referenced standards could be used as interim international criterion-referenced standards. The findings highlighted that boys and girls (aged 8–18 years) should attain the minimum standards of 42 and 35 $\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$, respectively, as anything below these standards should raise a red flag (67). Although a recent Canadian study applied these international standards in children aged 8–12 years and identified that nearly all Canadian girls have healthy CRF, whereas the percentage of boys with healthy CRF declined substantially with age (80). Thus, it is possible that these results highlight a limitation of the standards, which may be a result of using a single sex-specific standard for all age groups with only cardiometabolic criteria. As a result, there is a need for more research in this area.

International normative-referenced standards have been developed using a sample of 1,142,026 children and youth, representing 50 different countries (47). These standards provide a way to determine how children and youth perform on the 20mSRT compared to their peers by identifying their percentile rank. These types of standards, whether international (47) or national (56–63), could help complement criterion-referenced standards, especially in younger children of whom most probably have a healthy cardiometabolic profile, equating to high percentages with healthy CRF (64, 80). Together, criterion- and normative-reference standards could be used in research, physical education and clinical settings to screen for pediatric populations at risk of poor health and to compare the health of children and youth across various jurisdictions.

POPULATION-BASED SURVEILLANCE UTILITY

The implementation of effective population-based surveillance at the national and international level could provide insight into the effectiveness of broad public policy and interventions aimed

at improving health. There are two analytical strategies available to help optimise the use of surveillance data. First, data could be compared across jurisdictions at similar time points to help identify healthy and unhealthy populations. As a result, policy-makers in regions with unhealthy populations could learn and take similar policy-based approaches as comparable healthy populations. Second, repeated measures in the same population could identify temporal trends in health. Indeed, temporal trends may help monitor the impact of implemented policies aimed at improving health in a population, and they could help predict the health of future generations. It is also beneficial to compare similar health temporal trends across jurisdictions, although these analyses are complex and not common.

Comparisons of Surveillance Measures across Countries

Recently, a published study provided a comparison of 20mSRT performance across 50 countries, which currently represents the largest international effort to compare CRF levels among children and youth. Results from this study indicated that children and youth from countries in Northern Europe and Africa were the best 20mSRT performers, whereas those from South American countries were generally the worst 20mSRT performers (81). These findings were based on a systematic review and analysis of pooled aggregate 20mSRT results from 177 published studies, representing a fairly comprehensive surveillance effort. By comparing these results with more complex PA surveillance efforts, it is possible to identify the accuracy of these results given the interconnected relationship between CRF and PA (30).

An approach used to help understand differences in the PA levels of children and youth across countries is through the Active Healthy Kids Global Alliance's report cards on the physical activity of children and youth global matrix project (82). Through this project, researchers across countries cooperate to review and synthesize recent data on country-based PA levels among

children and youth using harmonized methods. As a result, this project has produced report cards for 38 countries that have helped identify patterns of low PA levels in some Asian, North American and South American countries, and higher PA levels in some Northern European, Oceanic and African countries (82). These results are consistent with findings from the recent 50-country comparison of 20mSRT performance scores among children and youth (81).

Nationally representative objective measures of PA levels among children and youth are rare and only available for a few high-income countries (i.e., parts of North America, Europe, and Oceania). The International Children's Accelerometry Database (ICAD) is a data sharing project that has generated the largest dataset of objective estimates of PA among children and youth (83). Results from ICAD across 10 countries support findings that PA levels are higher among children and youth in Northern European countries when compared with their counterparts in North America (84). Objective measures of PA are also available for over 3,000 youth from 10 European countries through the Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) study. The HELENA study demonstrated a European gradient where youth from Northern European countries were more active than their Southern European counterparts (85, 86). This result was replicated for CRF using 20mSRT scores available from 23 European countries (81). Another analysis using HELENA data found that differences in CRF across European countries were attenuated when controlling for PA levels, highlighting the interconnected relationship between CRF and objectively measured PA (85). Together, these findings suggest that 20mSRT scores provide a fairly accurate reflection of objectively measured PA at the population level. Thus, if measuring PA using objective methods is not possible due to financial, economic or human resource constraints, the measurement of CRF using the 20mSRT and analyzed with health- and performance-related methods could be a feasible alternative. Using the 20mSRT for surveillance of population health would be particularly important in low- and

middle-income countries given that these countries may be in the midst of a physical activity transition, which may forecast future health consequences (81, 87). Furthermore, in countries with established objective PA surveillance systems (e.g., Canada, USA, UK), the addition of CRF surveillance would be valuable because CRF predicts health independent of PA, and this would therefore, provide additional information on health above and beyond what can be obtained by measuring PA alone.

Temporal Trends in Surveillance Measures

Temporal trends using consistent objective measures of PA over several years are only available in some high-income countries (e.g., Japan [88], Czech Republic [89], Canada [90], Sweden [11, 91]), but not internationally. This is partly because the methods used to measure PA have changed considerably over the years, and the use of accelerometers and other objective measures in the surveillance of pediatric PA levels is relatively new. Furthermore, temporal trends using subjective PA measures could be problematic because the bias associated with these measures may have systematically changed over time. Given that CRF assesses an individual's physiological response to their total PA profile, the best approach to fully understand trends in PA in most countries and globally is probably through the assessment of CRF using the widely adopted 20mSRT field-based assessment test. The 20mSRT has been conducted using similar protocols since its inception, providing a way to calculate consistently reported temporal trend data. Thus, trends in CRF using the 20mSRT can be dated back to the late 1980s, with a study published in 2003 having reported an average global decline in performance of 0.43% per annum across 129,882 children and youth from 11 countries (48). A follow-up study using an updated dataset to the 2003 study compared 20mSRT performance of 965,264 children and youth from 19 countries between 1981 and 2015 (92). The included countries represented high-income and upper-middle-income countries supporting the need for more data in low- and middle-income

countries. The study results suggested that time trends in CRF were not uniform across countries with all but two countries (Japan and Brazil) demonstrating declines in recent decades. However, a general slowing of the decline existed amongst most countries where temporal trends in CRF seem to have stabilized, at least in high-income and upper-middle-income countries, in recent years (92). This finding is encouraging and may suggest that the plethora of interventions and policies implemented in these countries might be having an impact on slowing the decline in childhood CRF levels. Nevertheless, there remains a need to explore new policies and interventions aimed at improving CRF and PA levels in these populations — insight that Japan and Brazil might be able to provide.

NOVEL POPULATION HEALTH SURVEILLANCE STRATEGIES

The measurement of CRF (i.e., 20mSRT) is cost-effective and would be easy to administer at the population level. As a result, CRF surveillance provides several opportunities in the scientific research, physical education and clinical settings.

The Scientific Research Setting

An important issue in 20mSRT research is that three known protocol variants are extant, and results can be reported in multiple performance metrics (48). Although data conversion techniques are available (47), they are time consuming and problematic. It is important that researchers report 20mSRT results following the Tomkinson recommendations to improve data commensurability (47). To achieve this goal, a recent publication has recommended the development of an online 20mSRT data repository where researchers around the world can share their 20mSRT data in a consistent way in order to help improve surveillance efforts (81). In return, the website (www.activehealthykids.org/kids-fit-guide/) would compute a results sheet

comparing the participants' performance to national and international normative-referenced and criterion-referenced standards, providing the researcher with valuable analytical support. These data would be securely stored and used to improve CRF surveillance efforts by updating regional, national and international CRF trends, cross-country comparisons and normative-referenced data. In addition, the database would provide rich data for researchers worldwide to analyze and use in their respective research projects. Being free and simple to use, this data repository could help bridge the gap in surveillance efforts in low- and middle-income countries. The website is currently under construction in preparation for its official launch in early 2018.

The Physical Education Setting

The school setting remains one of the most promising locations to implement consistent national and international surveillance for school-aged children and youth (93). School-based surveillance efforts could also be considered an effective population-based intervention. For instance, the commitment to consistent periodic surveillance surveys places importance and develops a 'conversation' around particular measures. This is evident in school settings where national standardized testing in some countries reinforces the importance of certain academic skills over others. Although the 20mSRT has been used successfully in the school setting for years, only a few countries and jurisdictions have committed to CRF surveillance in schools (e.g., Hungary [9], certain States in the US [49], South Korea [95], Japan [96], China [96]). Of these countries Slovenia provides a good example of how fitness surveillance can lead to improved health.

Slovenia collects national CRF and other fitness data for all children and youth (aged 6–19 years) through a surveillance initiative called SLOfit (97). The SLOfit initiative has been mandatory for all primary and secondary schools for more than 30 years. Through mandatory annual surveillance of standardized fitness measures researchers and policy-makers were able to identify

a declining trend in youth fitness from 1990 to 2010. To mitigate this trend, Slovenia implemented a national health promoting PA intervention called Healthy Lifestyle in the school year (2010/2011), which provided children and youth with two optional additional hours of PA per week. As a result, the fitness levels of Slovenian children and youth have steadily improved since 2011. Furthermore, a recent report compared PA levels across 38 countries (representing six continents) and reported that Slovenia was ranked number one in overall national PA levels among youth (81). The Slovenia case-study highlights the importance of surveillance in guiding and monitoring the impact of health promoting interventions and policies — an example that other countries could adopt.

The Clinical Setting

In adults, the use of CRF assessments in clinical settings have been proposed by the American Heart Association as an important vital sign that could help improve patient risk classification for adverse health outcomes and provide a better way to inform patient management (5). Similar to adults, the use of CRF assessments in children and youth could be used as a clinical vital sign in at-risk patients (e.g., obesity, diabetes mellitus, heart conditions, etc.) to help improve risk classification, monitor changes in health status, and help prescribe evidence-informed lifestyle management. The 20mSRT is a maximal assessment that is not recommended for at-risk patients, thus, less valid submaximal tests the YWCA submaximal bicycle test (98), PWC170 (99), child adapted Astrand-Ryhming test (100), mCAFT (101), and HALO submaximal treadmill test (102) could be used for these populations to provide a measurement of CRF. The consistent collection and reporting of CRF data for at-risk children and youth, in addition to other health measures, could provide valuable surveillance data for scientists to research the link between changes in CRF and health outcomes in these populations. This approach also lends itself well to the current

ACSM *Exercise is Medicine* movement (103) by providing clinicians with a way to objectively monitor PA levels in their patients through the measurement of CRF.

IMPORTANCE OF GUIDELINES FOR PUBLIC HEALTH

The development and distribution of guidelines provides a method to set a benchmark that is widely accepted and easily communicated across all sectors of the population. Given that CRF is a holistic indicator of health among children and youth that has strong utility for population-based surveillance, developing public health guidelines may be a beneficial approach to help advance CRF surveillance.

Cardiorespiratory Fitness Guidelines

The interim international criterion-referenced standards of 42 and 35 ml·kg⁻¹·min⁻¹ for boys and girls aged 8–18 years (67), respectively, represent the best available CRF standards with which to identify children and youth with good health despite there being a need for more research. The interim standards have been converted to 20mSRT laps and the corresponding international percentages of children and youth that meet the standards have been estimated (Table 2). In addition, the international normative-referenced standards could be used to establish cut-points for low, medium, and high 20mSRT performance to complement the criterion-referenced standards (47; Table 3). Combined, these 20mSRT values currently represent the best available evidence from which to build CRF guidelines among children and youth.

<<< Insert Tables 2 and 3 here >>>

The development of CRF guidelines could provide many benefits. For instance, CRF guidelines could be an effective knowledge translation tool to communicate the importance of CRF for

current and future health. The guidelines would help support surveillance efforts by communicating the minimum level of fitness needed for good health, similar to the PA guidelines. These guidelines would be effective for screening initiatives in the school and clinical settings to identify those at risk of poor health. Screening efforts could help identify those who need support in increasing PA levels through exercise prescription and/or lifestyle management. Furthermore, the establishment of an online 20mSRT CRF data repository would allow for the guidelines to be progressively assessed and modified as emerging evidence is collected.

How to Meet the Cardiorespiratory Fitness Guidelines

Although CRF has an important genetic component (30), it has been described as a reflection of an individual's PA participation in recent weeks or months (104). Research supports the notion that regular PA, especially at moderate and vigorous intensities, could produce a significant response in CRF levels. More specifically, a 12-week training program can induce an 8–9% (~20 normative-referenced percentile points) increase in CRF among children and youth independent of sex, age and maturation (33). Thus, it is likely that those who do not meet the CRF guidelines are not physically active enough to support a healthy CRF level. As a result, when screening pediatric populations using the CRF guidelines, it would be reasonable to prescribe PA guidelines to those with low CRF as a way to improve health, and to encourage continued PA in those with good CRF to maintain health. In addition, it is important to encourage healthy lifestyle choices to maintain healthy body weight and avoid excess body fatness (42). This approach could also serve as an important education tool to highlight the importance of PA and to track progress through CRF levels.

PERSPECTIVES FOR PROGRESS

In order to develop and expand CRF surveillance initiatives there are several priority areas for future research:

1. This review provides an opinion that developing CRF guidelines could be an important strategy to help advance public health knowledge and research on CRF. Because this opinion may lack consensus throughout the public health and research communities, it would be important to develop a consensus statement on CRF guidelines.
2. This review demonstrates the importance of the 20mSRT as an indicator of current and future health among children and youth. Given the lack of temporal data among children and youth in low- and middle-income countries (parts of South America, Africa, and Asia) there is a need for coordinated efforts to measure the current state of CRF in these countries and worldwide. This could be carried out through an international survey on the 20mSRT performance of children and youth. These data would be vital for identifying populations in poor health and at-risk of potential elevated disease burden in the future, as well as for monitoring trends over time.

In conclusion, we recommend a paradigm shift that includes CRF surveillance among children and youth. CRF surveillance could supplement current PA surveillance, and in some cases CRF could replace the surveillance of PA, especially in countries with financial, human resource and/or economic constraints. Furthermore, the population-based measurement of CRF could provide a method to better understand the link between PA levels and health among children and youth at the population level. In order to expand the utility of CRF surveillance there is a need to establish clear CRF guidelines to provide a uniform interpretation of CRF among children and youth. The use of consistent and periodic CRF surveillance could help guide and evaluate health promoting policies and interventions aimed at reducing NCD burden and improving overall health in the population. Perhaps sufficient investment in CRF surveillance could be the first step

toward increasing healthy life expectancy where people can live longer without NCDs and disability — bridging the gap between increasing life expectancy and healthy life expectancy.

REFERENCES

1. GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*. 2016; 388:1459–544.
2. Salomon JA, Wang H, Freeman M, et al. Healthy life expectancy for 187 countries, 1990-2010: a systematic analysis for the Global Burden Disease Study 2010. *The Lancet*. 2012; 380:2144–62.
3. Bloom DE, Cafiero ET, Jané-Llopis E, et al. The global economic burden of non-communicable disease. Geneva: World Economic Forum. 2011.
4. Fletcher GF, Blair SN, Blumenthal J, et al. Statement on exercise: benefits and recommendations for physical activity programs for all Americans: a statement for health professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology, American Heart Association. *Circulation*. 1992; 86:340–4.
5. Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation*. 2016; 134(24):e653–e699
6. Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet*. 2012; 380(9838): 219–29
7. Berenson GS, Srinivasan SR, Bao W, Newman WP, Tracey RE, Wattigney WA. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults: The Bogalusa Heart Study. *N Engl J Med*. 1998; 338(23):1650–5.
8. Dollman J, Norton K, Norton L. Evidence for secular trends in children's physical activity behaviour. *Br J Sports Med*. 2005; 39:892–7.
9. Ekelund U, Tomkinson G, Armstrong N. What proportion of youth are physically active? Measurement issues, levels and recent time trends. *Br J Sports Med*. 2011; 45:859–65.
10. Booth VM, Rowlands AV, Dollman J. Physical activity temporal trends among children and adolescents. *J Sci Med Sport*. 2015; 18(4): 418–25.
11. Ortega FB, Konstabel K, Pasquali E, et al. Objectively measured physical activity and sedentary time during childhood, adolescence and young adulthood: a cohort study. *PLoS One*. 2013; 8(4):e60871.
12. Tomkinson GR, Olds TS. Secular changes in pediatric aerobic fitness test performance: The global picture. *Med Sport Sci*. 2007; 50:46–66.
13. Tremblay MS, Shields M, Laviolette M, Craig CL, Janssen I, Connor Gorber S. Fitness of Canadian children and youth: results from the 2007-2009 Canadian Health Measures Survey. *Health Reports*. 2010; 21(1):7–20.

14. Nelson MC, Neumark-Stzainer D, Hannan PJ, Sirard JR, Story M. Longitudinal and secular trends in physical activity and sedentary behavior during adolescence. *Pediatr*. 2006; 118(6):e1627.
15. Alwan A, MacLean DR, Riley LM, et al. Monitoring and surveillance of chronic non-communicable diseases: progress and capacity in high-burden countries. *The Lancet*. 2010; 376:1861–8.
16. Hallal PC, Andersen LB, Bull F, Guthold R, Haskell W, Ekelund U. Global physical activity levels: Surveillance progress, pitfalls, and prospects. *The Lancet*. 2012;380:247–257.
17. Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: A powerful marker of health. *Int J Obes*. 2008; 32:1–11.
18. Ruiz JR, Castro-Piñero, Artero EG, et al. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med*. 2009; 43:909–23.
19. Eklund U, Anderssen SA, Froberg K, Sardinha LB, Andersen LB, Brage S, et al. Independent associations of physical activity and cardiorespiratory fitness with metabolic risk factors in children: the European youth heart study. *Diabetologia*. 2007; 50: 1832–40.
20. IOM (Institute of Medicine). *Fitness measures and health outcomes in youth*. Washington, DC: The National Academies Press; 2012.
21. Pate RR, Welk GJ, McIver KL. Large-scale youth physical fitness testing in the United States: A 25-year retrospective review. *Pediatr Exerc Sci*. 2013; 25:515–23.
22. Ruiz JR, Castro-Piñero J, España-Romero V, et al. Field-based fitness assessment in young people: the ALPHA health-related fitness test battery for children and adolescents. *Br J Sports Med*. 2011; 45(6):518–24.
23. Mayorga-Vega M, Aguilar-Soto P, Viciania J. Criterion-related validity of the 20-M shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. *J Sports Sci Med*. 2015; 14:536–47.
24. Ruiz JR, Silva G, Oliveira N, Ribeiro JC, Oliveira JF, Mota J. Criterion-related validity of the 20-m shuttle run test in youths aged 13-19 years. *J Sports Sci*. 2009; 27(9):899–906.
25. Castro-Piñero J, Artero EG, España-Romero V, Ortega FB, Sjöström M, Suni J, Ruiz JR. Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med*. 2010; 44(13):934–43.
26. Tomkinson GR, Olds TS. Field tests of fitness. In: Armstrong N, van Mechlen W, eds. *Paediatric exercise science and medicine*. New York, NY: Oxford University Press. 2008:109–28.
27. Ortega FB, Artero EG, Ruiz JR, Vicente-Rodriguez G, Bergman P, Hagströmer M, et al. Reliability of health-related physical fitness tests in European adolescents. The HELENA Study. *Int J Obes (Lond)*. 2008; 32(Suppl 5):S49–57.

28. Artero EG, España-Romero V, Castro-Piñero J, Ortega FB, Suni J, Castillo-Garzon MJ, et al. Reliability of field-based fitness tests in youth. *Int J Sports Med*. 2011;32(3):159–69.
29. Bouchard C, Shephard RJ. Physical activity, fitness and health: the model and key concepts. In: Bouchard C, Shephard RJ, Stephens T (eds.). *Physical activity, fitness and health*. International Proceedings and Consensus Statement. Champaign III: Human Kinetics. 1994: 77–88.
30. Malina RM. Physical activity and fitness: pathways from childhood to adulthood. *Am J Hum Biol*. 2001;13:162–172.
31. Ruiz JR, Rizzo NS, Hurtig-Wennlöf A, Ortega FB, Wärnberg J, Sjöström M. Relations of total physical activity and intensity to fitness and fatness in children: the European Youth Heart Study. *Am J Clin Nutr*. 2006; 84(2):299–303.
32. Steele RM, Brage S, Corder K, Wareham NJ, Ekelund U. Physical activity, cardiorespiratory fitness, and metabolic syndrome in youth. *J Appl Physiol*. 2008; 105:342–51.
33. Armstrong N, Tomkinson GR, Ekelund U. Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during youth. *Br J Sports Med*. 2011; 45:849–58.
34. Armstrong N. Aerobic fitness and physical activity in children. *Pediatr Exerc Sci*. 2013; 25:548–60.
35. Kemper HCG, van Mechelen. Physical fitness testing of children: A European perspective. *Pediatr Exerc Sci*. 1996; 8(3):201–14.
36. Morrow Jr. JR, Zhu W, Franks DB, et al. 1958–2008 50 years of youth fitness tests in the United States. *Res Q Exerc Sport*. 2009; 80(1):1–11.
37. Ortega FB, Ruiz JR, España-Romero V, et al. The International Fitness Scale (IFIS): usefulness of self-reported fitness in youth. *Int J Epidemiol*. 2011; 40:701–11.
38. Sánchez-López M, Martínez-Vizcaíno V, García-Hermoso A, Jiménez-Pavón D, Ortega FB. Construct validity and test-retest reliability of the International Fitness Scale (IFIS) in Spanish children aged 9–12 years. *Scand J Med Sci Sports*. 2015; 25(4):543–51.
39. Rowland TW. Does peak VO₂ reflect VO₂max in children?: evidence from supramaximal testing. *Med Sci Sports Exerc*. 1993; 25(6):689–93.
40. Armstrong N, Barker AR. Oxygen uptake kinetics in children and adolescents: A review. *Pediatr Exerc Sci*. 2009; 21:130–47.
41. Buskirk E, Taylor HL. Maximal oxygen intake and its relation to body composition, with special reference to chronic physical activity and obesity. *J Appl Physiol*. 1957; 11(1):72–8.
42. Lloyd LK, Bishop PA, Walker JL, Sharp KR, Richardson MT. The influence of body size and composition on FITNESSGRAM® test performance and the adjustment of

- FITNESSGRAM® test scores for skinfold thickness in youth. *Meas Phys Edu Exerc Sci*. 2003; 7(4): 205–26.
43. Welsman JR, Armstrong N, Nevill AM, Winter EM, Kirby BJ. Scaling peak VO₂ for differences in body size. *Med Sci Sports Exerc*. 1996; 28(2):259–65.
 44. Léger L, Lambert J, Goulet A, Rowan C, Dinelle Y. Capacité aérobie des Québécois de 6 à 17 ans – test navette de 20 mètre avec paliers de 1 minute. *Can J Appl Sport Sci*. 1984; 9:64–9.
 45. Léger 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.
 46. Olds TS, Tomkinson GR, Léger 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.
 47. 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*. Published online first: 23 April 2016.
 48. Tomkinson GR, Léger 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.
 49. Welk GJ, Bai Y, Saint-Maurice PF, Allums-Featherston K, Candelaria N. Design and evaluation of the NFL PLAY 60 FITNESSGRAM® partnership project. *Res Q Exerc Sport*. 2016; 87(1):1–13.
 50. Melo X, Santa-Clara H, Almeida JP, Carnero EA, Sardinha LB, Bruno PM, et al. Comparing several equations that predict peak VO₂ using the 20-m multistage-shuttle run-test in 8-10-year-old children. *Eur J Appl Physiol*. 2011; 11(5):839–49.
 51. Moreira C, Santos R, Ruiz JR, et al. Comparison of different VO_{2max} equations in the ability to discriminate the metabolic risk in Portuguese adolescents. *J Sci Med Sport*. 2011; 14:79–84.
 52. Barnett A, Chan LYS, Bruce IC. A preliminary study of the 20-m multistage shuttle run as a predictor of peak VO₂ in Hong Kong Chinese students. *Pediatr Exerc Sci*. 1993; 5: 42–50.
 53. Matsuzaka A, Takahashi Y, Yamazoe M, et al. Validity of the multistage 20-m shuttle-run test for Japanese children, adolescents, and adults. *Pediatr Exerc Sci*. 2004; 16: 113–25.
 54. Mahar MT, Welk GJ, Rowe DA, et al. Development and validation of a regression model to estimate VO_{2peak} from PACER 20-m shuttle run performance. *J Phys Act Health*. 2006; 3(Suppl. 2):S34–S46.
 55. Ruiz JR, Ramirez-Lechuga J, Ortega FB, et al. Artificial neural network-based equation for estimating VO_{2max} from the 20 m shuttle run test in adolescents. *Artif Intell Med*. 2008; 44: 233–45.

56. Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9-17-year-old Australians since 1985. *Br J Sports Med.* 2013; 47:98–108.
57. Sandercock G, Voss C, Cohen D, Taylor M, Stasinopoulos DM. Centile curves and normative values for the twenty metre shuttle-run test in English schoolchildren. *J Sports Sci.* 2012; 30(7):679–87.
58. Santos R, Mota J, Santos DA, Silva AM, Baptista F, Sardinha LB. Physical fitness percentiles for Portuguese children and adolescents aged 10-18 years. *J Sports Sci.* 2014; 32(16):1510–8.
59. Sauka M, Priedite IS, Artjuhova L, et al. Physical fitness in northern European youth: reference values from the Latvian Physical Health in Youth Study. *Scand J Public Health.* 2011; 39(1):35–43.
60. Tambalis KD, Panagiotakos DB, Psarra G, et al. Physical fitness normative values for 6-18-year-old Greek boys and girls, using the empiracle distribution and the lambda, mu, and sigma statistical method. *Eur J Sport Sci.* 2016; 16(6):736–46.
61. De Miguel-Etayo P, Gracia-Marco L, Ortega FB, Intemann T, Foraita R, Lissner L, et al. Physical fitness reference standards in European children: the IDEFICS study. *Int J Obes (Lond).* 2014; 38: S57–S66
62. Ortega FB, Artero EG, Ruiz JR, Espana-Romero V, Jimenez-Pavon D, Vicente-Rodriguez G, et al. Physical fitness levels among European adolescents: the HELENA study. *Br J Sports Med.* 2011; 45(1): 20–9.
63. Ortega FB, Ruiz JR, Castillo MJ, Moreno LA, Gonzalez-Gross M, Warnberg J, Gutierrez A. [Low level of physical fitness in Spanish adolescents. Relevance for future cardiovascular health (AVENA study)]. *Rev Esp Cardiol.* 2005; 58(8):898–909.
64. Lang JJ, Tremblay MS, Ortega FB, Ruiz JR, Tomkinson GR. Review of criterion-referenced standards for cardiorespiratory fitness: what percentage of 1 142 026 international children and youth are apparently healthy? *Br J Sports Med.* Published Online First: 2 March 2017.
65. Lobelo F, Pate RR, Dowda M, Liese AD, Ruiz JR. Validity of cardiovascular fitness criterion-referenced standards for adolescents. *Med Sci Sports Exerc.* 2009; 41(6):1222–9.
66. Mesa JL, Ruiz JR, Ortega FB, et al. Aerobic physical fitness in relation to blood lipids and fasting glycaemia in adolescents: influence of weight status. *Nutr Metab cardiovasc Dis.* 2006; 16(4):285–93.
67. Ruiz JR, Caverio-Redendo 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.* Published online first: 26 September 2016.
68. Welk GJ, Laurson KR, Eisenmann JC, Cureton KJ. Development of youth aerobic-capacity using receiver operating characteristic curves. *Am J Prev Med.* 2011; 41(4 Suppl 2):S111–S116.

69. Ruiz JR, Ortega FB, Rizzo NS, et al. High cardiovascular fitness is associated with low metabolic risk scores in children: the European Youth Heart Study. *Pediatr Res*. 2007; 61(3):350–5.
70. Ruiz JR, Ortega FB, Meusel D, Harro M, Oja P, Sjöström M. Cardiorespiratory fitness is associated with features of metabolic risk factors in children. Should cardiorespiratory fitness be assessed in a European health monitoring system? The European Youth Heart Study. *J Public Health*. 2006; 14(2):94–102.
71. Ruiz JR, Ortega FB, Meusel D, Sjöström M. Traditional and novel cardiovascular risk factors in school-aged children: A call for the further development of public health strategies with emphasis on fitness. *J Public Health*. 2007; 15:171.
72. Ruiz JR, Huybrechts I, Cuenca-García M, et al. Cardiorespiratory fitness and ideal cardiovascular health in European adolescents. *Heart*. 2015; 101(10):766–73.
73. Hurtig-Wennlöf A, Ruiz JR, Harro M, Sjöström 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*. 2007; 14(4):575–81.
74. Rizzo NS, Ruiz JR, Hurtig-Wennlöf A, Ortega FB, Sjöström M. Relationship of physical activity, fitness, and fatness with clustered metabolic risk in children and adolescents: the European youth heart study. *J Pediatr*. 2007; 150(4):388–94.
75. Erikssen G, Liestøl K, Børnholt J, Thaulow E, Sandvik L, Erikssen J. Changes in physical fitness and changes in mortality. *The Lancet*. 1998; 352:759–62.
76. Slaterry M, Jacobs DR Jr. Physical fitness and cardiovascular disease mortality: the US railroad study. *Am J Epidemiol*. 1988; 127:571–80.
77. Lang JJ, Belanger K, Poitras V, et al. Systematic review of the relationship between 20 m shuttle run test performance and health markers in children and youth. *J Sci Med Sport*. 2017. In press. doi:10.1016/j.jsams.2017.08.002
78. Höglström G, Nordström A, Nordström P. High aerobic fitness in late adolescence is associated with a reduced risk of myocardial infarction later in life: a nationwide cohort study in men. *Eur Heart J*. 2014; 35:3133–40.
79. Höglström G, Nordström A, Nordström P. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. *Int J Epi*. Published Online First: 20 December 2015. doi:10.1093/ije/dyv321.
80. Lang JJ, Chaput JP, Longmuir PE, Barnes JD, Belanger K, Tomkinson GR, et al. Cardiorespiratory fitness is associated with physical literacy among Canadian children aged 8 to 12 years. Submitted to *BMC Public Health*.
81. Lang JJ, Tremblay MS, Léger L, et al. International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results. *Br J Sports Med*. Published Online First: 20 Sept 2016.

82. Tremblay MS, Barnes JD, González SA, Katzmarzyk PT, Onywera VO, Reilly JJ, et al. Globla matrix 2.0: Report card grades on the physical activity of children and youth comparing 38 countries. *J Phys Act Health*. 2016;13(Suppl 2):S343–66.
83. Sherar LB, Griew P, Esliger DW, et al. International children's accelerometry database (ICAD): Design and methods. *BMC Public Health*. 2011; 11:485.
84. Cooper AR, Goodman A, Page AS, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). *Int J Behav Nutr Phys Act*. 2015; 12:113.
85. Ortega FB, Ruiz JR, Labayen I, et al. Health inequality in urban adolescents: role of physical activity, diet, and genetics. *Pediatr*. 2014; 133(4):e884–95.
86. Ruiz JR, Ortega FB, Martínez-Gómez D, Labayen I, Moreno LA, De Bourdeaudhuij I, et al. Objectively measured physical activity and sedentary time in European adolescents: the HELENA study. *Am J Epidemiol*. 2011; 174(2):173–84.
87. Katzmarzyk PT, Mason C. The physical activity transition. *J Phys Act Health*. 2009; 6(3):269–80.
88. Inoue S, Ohya Y, Tudor-Locke C, Tanaka S, Yoshiike N, Shimomitsu T. Time trends for step-determined physical activity among Japanese adults. *Med Sci Sports Exerc*. 2011; 43:1913–9.
89. Sigmundova D, El Ansari W, Sigmund E, Fromel K. Secular trends: a ten-year comparison of the amount and type of physical activity and inactivity of random samples of adolescents in the Czech Republic. *BMC Public Health*. 2011; 11:731.
90. Thompson AM, McHugh TL, Blanchard CM, et al. Physical activity of children and youth in Nova Scotia from 2001/02 and 2005/06. *Prev Med*. 2009; 49:407–9.
91. Raustorp A, Ludvigsson J. Secular trends of pedometer-determined physical activity in Swedish school children. *Acta Paediatr*. 2007; 96:1824–8.
92. Tomkinson GT, Lang JJ, Tremblay MS. Temporal trends in the cardiorespiratory fitness of 965,264 children and youth representing 19 countries since 1981. *Br J Sports Med*. Under review
93. Pate RR. The case for large-scale physical fitness testing in American youth. *Pediatr Exerc Sci*. 1989; 1:290–4.
94. Csányi T, Finn KJ, Welk GJ, Zhu W, Karsai I, Ihász F, et al. Overview of the Hungarian National Youth Fitness Study. *Res Q Exerc Sport*. 2015; 86(Suppl 1):S3–S12.
95. Tomkinson GR, Olds TS, Kang SJ, Kim DY. Secular trends in the aerobic fitness test performance and body mass index of Korean children and adolescents (1968 – 2000). *Int J Sports Med*. 2007; 28:314–20.
96. Macfarlane DJ, Tomkinson G. Evolution and variability in fitness test performance of Asian children and adolescents. *Med Sport Sci*. 2007;50:143–67.

97. Strel J. Analiza programa »Zdrav življenjski slog« za leti 2010/11 in 2011/12. Ljubljana: Institute for Sport Planica; 2013. [Analysis of the program Healthy Lifestyle for the years 2010/11 and 2011/12].
98. YMCA. YMCA fitness testing and assessment manual (4th edition). Champaign III: Human Kinetics. 2002.
99. Campbell PT, Katzmarzyk P, Malina RM, Rao DC, Pérusse L, Bouchard C. Prediction of physical activity and physical work capacity (PWC150) in young adulthood from childhood and adolescence with consideration of parental measures. *Am J Hum Biol.* 2001; 13(2):190–6.
100. Buono MJ, Roby JJ, Micale FG, Sallis JF. Predicting maximal oxygen uptake in children: modification of the Astrand-Ryhming nomogram for predicting maximal oxygen intake. *Br J Sports Med.* 1981; 15(3): 182–5.
101. Weller IMR, Thomas SG, Corey PN, Cox MH. Prediction of maximal oxygen uptake from a modified Canadian aerobic fitness test. *Can J Appl Physiol.* 1993; 18(2):175–88.
102. Breithaupt P, Adamo KB, Colley RC. The HALO submaximal treadmill protocol to measure cardiorespiratory fitness in obese children and youth: a proof of principle study. *Appl Physiol Nutr Metab.* 2012; 37:308–14.
103. Lobelo F, Stoutenberg M, Hutber A. The exercise is medicine global health initiative: a 2014 update. *Br J Sports Med.* Published online first: 28 April 2014.
104. Baquet G, Van Praagh E, Berthoin S. Endurance training and aerobic fitness in young people. *Sports Med.* 2003; 33(15):1127:1143.

Table 1 20mSRT prediction equations to estimate $\dot{V}O_{2\text{peak}}$ in children and youth.

Citation (publication year)	Sample	Country	Equation to estimate $\dot{V}O_{2\text{peak}}$
Léger (43; 1988)	188 boys and girls, aged 8–19 years	Canada	$\dot{V}O_{2\text{peak}} = 31.025 + 3.238 \cdot S - 3.248 \cdot A + 0.1536 \cdot S \cdot A$
Barnett (50; 1993)	27 boys, 28 girls, aged 12–17 years	Hong Kong	$\dot{V}O_{2\text{peak}} = 25.8 - 6.6 \cdot G_1 - 0.2 \cdot W + 3.2 \cdot S$ $\dot{V}O_{2\text{peak}} = 24.4 - 5.0 \cdot G_1 - 0.8 \cdot A + 3.4 \cdot S$
Matsuzaka (51; 2004)	62 boys, 70 girls, aged 8–17 years	Japan	$\dot{V}O_{2\text{peak}} = 25.9 - 2.21 \cdot G_1 - 0.449 \cdot A - 0.831 \cdot BMI + 4.12 \cdot S$
Mahar (52; 2006)	61 boys, 74 girls, aged 12–14 years	USA	$\dot{V}O_{2\text{peak}} = 47.438 + L \cdot 0.242 + G_2 \cdot 5.134 - W \cdot 0.197$
Ruiz (53; 2008)	122 boys, 71 girls, aged 13–19 years	Spain	Artificial neural network equation available: http://www.helenastudy.com/scientific.php

Note: S = running speed ($\text{km} \cdot \text{h}^{-1}$) at the last completed stage; A = participant's age at last birthday; G_1 = gender (male = 0, female = 1); W = body weight (kg); BMI = body mass index (kg/m^2); L = number of laps completed; G_2 = gender (female = 0, male = 1)

Table 2 20mSRT minimum recommended standards of 42 and 35 ml·kg⁻¹·min⁻¹ for boys and girls, respectively. Data available from Ruiz et al. (65).

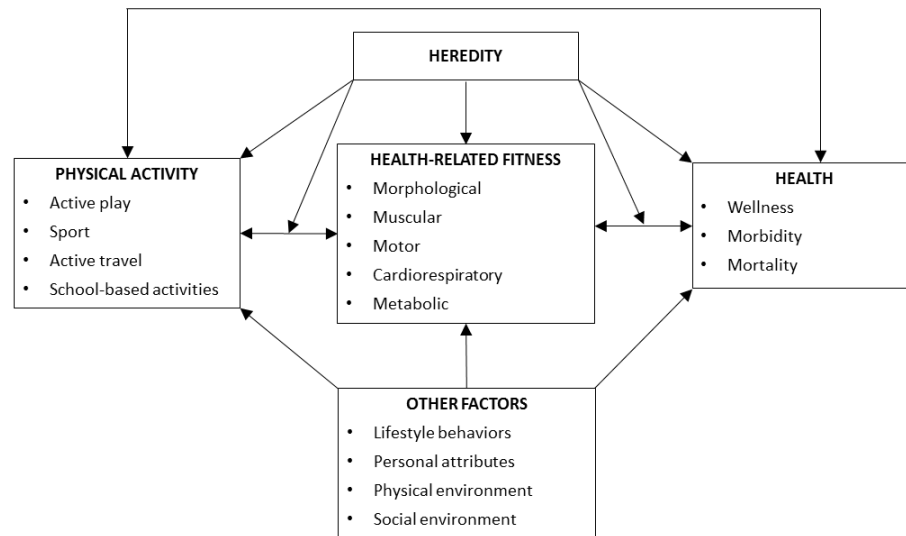
Age (years)	Boys			Girls		
	No. of completed 20mSRT laps	Running speed (km·h ⁻¹)	International percentage (%) meeting standards ^a	No. of completed 20mSRT laps	Running speed (km·h ⁻¹)	International percentage (%) meeting standards ^a
8	7	8.5		7	8.5	
9	15	9.0	93.0	7	8.5	99.9
10	15	9.0	87.2	7	8.5	99.6
11	23	9.5	78.5	7	8.5	98.6
12	31	10.0	72.3	7	8.5	93.9
13	40	10.5	70.3	15	9.0	88.3
14	40	10.5	62.8	15	9.0	78.6
15	49	11.0	62.3	23	9.5	67.2
16	49	11.0	56.2	31	10.0	54.1
17	59	11.5	51.7	31	10.0	43.6
18	59	11.5		40	10.5	

^a Percentage data for 9–17 year olds available from Lang et al. (62).

FIGURE CAPTION

Figure 1 Modified version of Bouchard and Shephard's health-related fitness, physical activity and health model (29). Reproduced with permission.

Figure 1



CHAPTER 8: GENERAL DISCUSSION

The overarching purpose of this dissertation was to investigate the utility of cardiorespiratory fitness (CRF), measured using the 20 m shuttle run test (20mSRT), as a holistic health indicator for population surveillance among school-aged children and youth. To achieve this purpose I conducted seven studies that fit into one of two categories: 1) associations between 20mSRT performance and health among children and youth; and 2) how the 20mSRT performance of children and youth varies across countries to provide up-to-date international surveillance data. Thus, the remainder of this chapter will complement the discussion found in Chapter 7 by providing an overview of how each Chapter informs these two categories. I will follow by providing an overview of the dissertation strengths and limitations, important future research direction, suitability for a dissertation in Population Health, and the final conclusions.

8.1 Cardiorespiratory Fitness and Health among Children and Youth

As described in Chapter 1, this dissertation was guided by a modified version of the Bouchard and Shephard's health-related fitness model, which describes the interconnected relationships between physical activity, health-related fitness, heredity, other factors, and health (see Figure 1.1).¹ In the literature review provided in Chapter 2, I identified a literature gap informing the association between 20mSRT performance and health among school-aged children and youth. Thus, Chapter 3 was included to help synthesize the available evidence using a systematic review method. Through this study

we identified 142 papers that explored associations between 20mSRT performance and health, combining to provide data for over 300,000 children and youth aged 5–17 years. We concluded that 20mSRT performance was favourably associated with physical, physiological, cognitive, and psychosocial health in these age groups. Thus, children who performed better during 20mSRT had better holistic health in comparison with children and youth who performed poorly on the test.

Next, we identified two methods to interpret 20mSRT performance data: normative-referenced and criterion-referenced standards. In Supplemental Manuscript 1 we applied a systematic review method to identify papers that reported 20mSRT descriptive statistics for children and youth aged 9–17 years. We identified and extracted data from 177 studies, representing 1 142 026 children and youth from 50 countries. All data were standardized to reflect a common protocol (Léger's 20mSRT²) and metric (running speed (km/h) at the last completed stage). These data were used to develop international sex- and age-specific 20mSRT normative-referenced standards using the Lambda Mu and Sigma method.³ Our normative-referenced standards can help children and youth identify their centile rank in comparison with their peers, providing an interpretation of their overall 20mSRT performance.

Chapter 4 helped clarify and provide direction in the use of criterion-referenced standards for CRF among children and youth. We conducted a review and identified ten sets of criterion-referenced standards, published between 2006 and 2016, and developed using samples of children and youth from the Americas or Europe. We used data from Supplement 1 to help describe the differences in children and youth with apparently healthy CRF while using each of the ten standards. As a result of heterogeneity across

studies we highlighted the need to conform to a single set of criterion-referenced standards to improve comparability in this area of research. Thus, we recommended the Ruiz standards of 35 and 42 mL/kg/min for girls and boys, respectively, developed using a meta-analysis method combining results from seven published studies.⁴

In Chapter 5 we provided an original research article that aimed to assess the associations between CRF, as measured using the 20mSRT, and components of physical literacy among Canadian children aged 8–12 years. The secondary aim was to determine the percentage of Canadian children with healthy CRF using the Ruiz criterion-referenced standards.⁴ Results from this study showed clear benefits across all components of physical literacy in those who had high CRF, in comparison with those who had low CRF. Furthermore, almost all Canadian girls were identified as having apparently healthy CRF, whereas the percentage of boys with apparently healthy CRF declined with age. These results suggest a potential limitation of the Ruiz standards in correctly identifying younger children with healthy CRF, which is likely a result of using a single standard for all age groups. Consequently, there is more work needed to refine the existing criterion-referenced standards for CRF before they are used to compare the health of children and youth within and between countries.

8.2 Cardiorespiratory Fitness as a Population Health Surveillance Indicator

In Chapter 1 I described the importance of population health surveillance as a means to *guide* healthy public policy efforts by identifying populations at risk, and to *monitor* and *evaluate* the effectiveness of implemented policies. To help describe the policy process I incorporated the Bauman *et al.* ecological model which describes the

nested environments in which we live (i.e., individual, interpersonal, environment, regional or national policy, and global; Figure 1.2).⁵ Through a top-down mechanism, healthy public policy has the potential to influence our environment, interpersonal relationships, and our individual health. Likewise, the surveillance of population health indicators at the individual level provides valuable data that could help determine whether implemented healthy public policies are effective. Thus, I have provided three studies in this dissertation that could help inform future national or international CRF surveillance systems.

Both Chapters 6 and Supplemental Manuscript 2 used data gathered through Supplemental Manuscript 1. In Chapter 6 we pooled 20mSRT results from 50 countries to assess the international variability in 20mSRT performance among children and youth. Our results indicated that 20mSRT performance varies substantially worldwide. In general, countries from North-Central Europe and Africa were the best 20mSRT performers, whereas those from South America were the worst performers. These results highlight a need for better policy approaches in some countries to help improve and/or protect the health of their young citizens. Indeed, these data could help highlight countries with effective policy-based strategies aimed at improving the CRF of children and youth.

In Supplemental Manuscript 2 we report analyses of a subset of the data from Supplemental Manuscript 1 to calculate temporal trends in 20mSRT performance from 19 countries (high- and middle-income) since 1981, representing 965 264 children and youth. The results suggest that all but two countries (Japan and Brazil) have experienced declining temporal trends in 20mSRT performance. Across countries, the rate of decline

in performance has diminished with each decade and recently stabilized near zero. These results are encouraging and may suggest that high- and middle-income countries have implemented successful healthy public policy to improve the health of children and youth in their respective country. However, more work is needed to help children and youth improve their overall CRF to levels seen in the 80's – insight that Brazil and Japan may be able to provide.

Chapter 7 provided a narrative literature review that discussed the state of the international health surveillance among children and youth. Using recent literature and evidence from this dissertation, we helped make a case for CRF surveillance among children and youth. This Chapter provides a detailed dissertation discussion by describing how each of the six dissertation manuscripts fit into the broad literature, including the implications of our research and future directions. We concluded by highlighting the need for consensus on the importance of CRF among children and youth, and we further describe how consensus could take place through CRF guidelines.

8.3 Summary of Strengths and Limitations

There are several strengths and limitations of this dissertation described in detail in each of the seven dissertation manuscripts, several of which are summarized below.

8.3.1 Strengths

The primary strength of this dissertation was the effort to pool 20mSRT data from 287 unique studies, through two systematic reviews of literature (Chapter 3 and Supplemental Manuscript 1). As a result, our work represents well over 1 million

children and youth from countries that spans the Human Development Index, and represents six of seven continents. To obtain these data we used rigorous systematic review methods following standardized procedures. We used the GRADE method to analyze quality of the evidence in Chapter 3. To compare data across studies we used standardized and rigorous data treatment methods in Chapters 4, 6, and Supplemental Manuscript 1 and 2. In these Chapters we also implemented a post-stratification population-weighted technique to correct for over- or under-representation across countries. In Chapter 5 we analyzed the RBC-CAPL dataset which provided several strengths, including harmonized data collection methods, large sample size, and a well-balanced cross-country Canadian representation.

8.3.2 Limitations

The manuscripts herein are not without limitations. The majority of these limitations are inherent when pooling data from numerous studies that used different data collection methods, sampling strategies, and various sample sizes and characteristics. For instance, there may have been unaccounted differences in testing conditions (i.e., indoor vs outdoor, elevation, and temperature) across included studies, and study specific measurement error may have been present which could have biased our estimates in some studies. Furthermore, uncontrolled temporal time trends in 20mSRT performance may have biased our estimates in Chapters 4, 6, and Supplemental Manuscript 1, given that our mean measurement year was 2000. In Chapter 5 we were unable to identify causal relationships because we analyzed cross-sectional data (RBC-CAPL dataset). In all studies we used data from convenience samples, with a few minor exceptions (e.g., Japan,

Hungary), making our results not generalizable to the broad population. As a result, our normative-referenced data may not be a true estimate of the international context since many countries were not represented and others used small convenience samples. In most studies we were unable to control for unmeasured confounders (i.e., socioeconomic status, maturation, heredity) which may have influenced our results. In all cases we interpret our results using chronological age, however given our age range (9–17 years old) using maturational age may have provided different results by accounting for youth who developed early or late compared with their peers. In addition, the 20mSRT is a maximal test to assess CRF which might be stigmatized in certain populations resulting in elevated non-response bias. This may have resulted in the ‘best’ performers participating in the test, whereas those who are not good at the 20mSRT choosing not to participate. This potential non-response bias may have had a large impact on our cross-country comparison of 20mSRT performance in Chapter 6. Finally, there may have been systematic error in underestimating $\text{VO}_{2\text{peak}}$ when using the Léger equation across studies.

In addition to these study specific limitations, there are several limitations inherent in using normative- and criterion-referenced standards. First, temporal trends in 20mSRT performance and health indicators among children and youth would result in the need to consistently update these standards. Second, the standards are only as good as the quality of the underlying data, suggesting that there is a need for large representative samples when calculating standards. Third, the use of normative-referenced standards may discourage those with poor performance while encouraging those who already have high

CRF levels. Fourth, the normative-referenced standards do not provide an indication of health. Last, criterion-referenced standards provide a dichotomous or categorical interpretation of results which eliminates or reduces our understanding of the distribution of performance across the sample.

8.4 Summary of Important Future Directions

Throughout the manuscripts included in this dissertation we provide future research directions. However, there are three specific areas for future research identified as important, and described in detail below:

1. There is a need for consensus on the importance, and how to best define healthy levels of CRF in childhood and adolescence. One important step needed to attain consensus is to develop CRF guidelines for children and youth, as described in Chapter 7. This could be achieved through a consensus statement that brings together an international expert panel of researchers and public health officials, which could serve as the basis for future CRF guidelines in these populations.
2. As described in Chapter 6 and 7, there is a need for consistent 20mSRT reporting standards to facilitate international pooling of data for surveillance purposes. To attain homogeneity in this field of research there is a need to develop an international 20mSRT data repository where researchers worldwide can share their data through a free and easy to access website. This type of a website could serve as an international CRF surveillance system, and it could be used to further advance the importance of CRF for health in children and youth.

3. There is a need to collect recent 20mSRT data in low- and middle-income countries (e.g., parts of Africa and Asia). These data would help update temporal trend data for these countries, providing more insight into the potential epidemiological transition seen in some countries.

8.5 Suitability for a Doctoral Dissertation in Population Health

Through this dissertation I demonstrated many skills and abilities that are integral criteria for a doctoral dissertation in population health. I critically appraised the CRF literature and identified several important research gaps, and followed by developing a combination of studies that aimed to fill these gaps through unique contributions to the literature. A skill that I demonstrated was the capability of developing and managing large datasets, and using these datasets to answer specific research questions. I also demonstrated the ability to synthesize a large body of literature, using systematic methods. Through a variety of secondary data analysis techniques, including Monte Carlo simulation and post-stratification population weighted methods I demonstrated my ability to carry out population-based research. Several analyses were also carried out using primary data with large population-based sample sizes. I was able to interpret the findings from my dissertation and link them to future directions aimed at improving population health, worldwide. Finally, through peer-reviewed publications in high impact journals and several academic conference presentations, I was able to display my abilities to disseminate my doctoral research.

8.6 Conclusion

This dissertation provides the largest effort to date to synthesize research that describes the international 20mSRT performance of children and youth. Specifically, it aimed to assess the utility of CRF as a holistic health indicator for population surveillance among school-aged children and youth. Through four studies we described the favourable associations between 20mSRT performance and holistic health among children and youth, and through the remaining three studies we described the broad utility of the 20mSRT as a population health surveillance indicator. The implications of this dissertation lie in its ability to potentially inform future international consensus on the importance of CRF in childhood and adolescence. More importantly, I hope this dissertation can provide a base to advocate for change in the way we monitor the population health of children and youth within and between countries. Using established CRF surveillance systems, future research could assess the effectiveness of the 20mSRT in guiding and evaluating healthy public policy aimed at improving CRF and reducing the population burden of non-communicable disease.

8.7 References

1. Bouchard C, Shephard RJ. Physical activity, fitness and health: the model and key concepts. In: Bouchard C, Shephard RJ, Stephens T, eds. *Physical Activity, Fitness and Health. International Proceedings and Consensus Statement*. Champaign III: Human Kinetics; 1994:77–88.
2. Léger L, Lambert J, Goulet A, Rowan C, Dinelle Y. Capacité aérobie des Québécois de 6 à 17 ans – test navette de 20 mètres avec paliers de 1 minute. *Canadian Journal of Applied Sport Science*. 1984; 9:64–9.
3. Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Statistics in Medicine*. 1992; 11:1305–19.
4. Ruiz JR, Cavero-Redondo I, Ortega FB, Welk GJ, Andersen LB, Martinez-Vizcaino V. 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*. 2015; 50:1451–8.
5. Bauman A, Reis RS, Sallis JF, Wells JC, Loos RJF, Martin BW, et al. Correlates of physical activity: why are some people physically active and others not? *Lancet*. 2012; 380:258–71.

SUPPLEMENTAL MANUSCRIPT 1

This manuscript has been written and formatted following the publication standards of the peer-reviewed journal, *British Journal of Sport Medicine*. It appears in its published format.

Title:

International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries

Citation:

Tomkinson GR, **Lang JJ**, Dale M, LeBlanc AG, Belanger K, Ortega FB, Léger L. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. *British Journal of Sports Medicine*. Published Online First: 20 May 2016. doi:10.1136/bjsports-2016-095987

Online link:

<http://bjsm.bmj.com/content/early/2016/05/20/bjsports-2016-095987.abstract>

International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries

Grant R Tomkinson,^{1,2} Justin J Lang,³ Mark S Tremblay,³ Michael Dale,⁴ Allana G LeBlanc,⁵ Kevin Belanger,³ Francisco B Ortega,^{6,7} Luc Léger⁸

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bjsports-2016-095987>).

For numbered affiliations see end of article.

Correspondence to Dr Grant R Tomkinson, Department of Kinesiology and Public Health Education, University of North Dakota, 2751 2nd Avenue North, Stop 8235, Grand Forks, ND 58202, USA; grant.tomkinson@und.edu

Accepted 23 April 2016

ABSTRACT

Objective To develop sex-specific and age-specific international norms for the 20 m shuttle run test (20mSRT) in children and youth (aged 9–17 years), and to estimate the prevalence meeting the FITNESSGRAM criterion-referenced standards for healthy cardiorespiratory endurance (CRE).

Methods A systematic review was undertaken to identify papers explicitly reporting descriptive 20mSRT (with 1 min stages) data on children and youth since 1981. Data were included on apparently healthy (free from known disease/injury) 9–17 years old. Following standardisation to a common metric and for protocol differences, pseudo data were generated using Monte Carlo simulation, with population-weighted sex-specific and age-specific normative centiles generated using the Lambda Mu and Sigma (LMS) method. Sex-related and age-related differences were expressed as per cent and standardised differences in means. The prevalence with healthy CRE was estimated using the sex-specific and age-specific FITNESSGRAM criterion-referenced standards for $\dot{V}O_{2peak}$.

Results Norms were displayed as tabulated centiles and as smoothed centile curves for the 20mSRT using 4 common metrics (speed at the last completed stage, completed stages/minutes, laps and relative $\dot{V}O_{2peak}$). The final data set included 1 142 026 children and youth from 50 countries, extracted from 177 studies. Boys consistently outperformed girls at each age group (mean difference $\pm$ 95% CI: 0.86 ± 0.28 km/h or 0.79 ± 0.20 standardised units), with the magnitude of age-related increase larger for boys than for girls. A higher proportion of boys (mean $\pm$ 95% CI: $67 \pm 14\%$) had healthy CRE than girls (mean $\pm$ 95% CI: $54 \pm 17\%$), with the prevalence of healthy CRE decreasing systematically with age.

Conclusions This study provides the most comprehensive and up-to-date set of international sex-specific and age-specific 20mSRT norms for children and youth, which have utility for health and fitness screening, profiling, monitoring and surveillance.

BACKGROUND

Cardiorespiratory endurance (CRE) is the ability to deliver oxygen to the muscles and to utilise it to generate energy to support muscle activity during exercise.^{1–2} In adults, low CRE is strongly associated with cardiovascular disease and all-cause mortality and morbidity (independent of adiposity),^{3–4} stroke,⁵ diabetes,⁶ mental health,⁷ health-related quality of life⁸ and cardiometabolic

disease risk.^{6–9} In children and youth, CRE is a weak-to-strong predictor of cardiovascular disease risk, cancer and mental health.^{10–11} There is also evidence of an interaction between adiposity and CRE, suggesting that high levels of CRE may attenuate the deleterious effects of being overweight or obese in children and youth, the so-called ‘fat but fit’ phenotype.¹² Thus, CRE provides insight into the synergistic capabilities of several bodily systems and organs that are involved in the performance of physical activity and exercise, providing a strong and summative measure of health in children and youth.¹⁰ CRE also tracks moderately well from childhood to adulthood,^{13–14} indicating that the measurement and surveillance of CRE in children provides insight into current and future population health statuses. Low fitness in childhood and adolescence is substantially linked with increased cardiometabolic disease risk,^{15–16} obesity,¹⁷ reduced quality of life,¹⁸ poorer skeletal¹⁹ and mental health¹¹ in adulthood. In addition to health implications, CRE is an important determinant of sporting success for many popular youth sports (eg, hockey, basketball, football (soccer), distance running, swimming, rugby).¹

The 20 m shuttle run test (20mSRT) is arguably the most popular field-based assessment and estimate of CRE in children and youth worldwide.^{20–23} It is an excellent tool for population-based surveillance and monitoring because it demonstrates strong-to-very strong test–retest reliability, and moderate-to-strong validity.²⁰ The 20mSRT also has excellent field-based utility due to its low cost, flexibility with testing locations (indoors, outdoors, smaller spaces) and its ability to test multiple individuals simultaneously with minimal equipment and personnel.^{20–24–25} In order to extend the utility of the 20mSRT as a surveillance instrument, there is a need for international norms, which, to date, have only been published for a single 20mSRT metric or for selected geographical regions,²⁰ including Europe,²⁶ North America²⁷ and Oceania.²⁸ Olds *et al*²¹ cumulated 20mSRT data on children and youth from 37 countries, representing six continents, suggesting that there is an international appetite for assessing CRE in children and youth. Harmonising reference values by creating international normative centiles for the 20mSRT in children and youth would provide opportunity for international surveillance and a means to compare CRE across geographic areas and time.

This study provides a 10-year update of the comprehensive 20mSRT reviews by Tomkinson *et al*²²

To cite: Tomkinson GR, Lang JJ, Tremblay MS, *et al*. *Br J Sports Med* Published Online First: [please include Day Month Year] doi:10.1136/bjsports-2016-095987

and Olds *et al.*²¹ which were the first studies to describe a method to harmonise 20mSRT data by standardising for differences in test protocols and performance metrics. The primary aim was to develop the most comprehensive set of sex-specific and age-specific international normative centiles for CRE. The secondary aim was to estimate the proportion of children and youth meeting the FITNESSGRAM criterion-referenced standards for healthy CRE. These data will facilitate the identification of children and youth with very low CRE in order to set appropriate goals and promote positive health-related fitness behaviours, and conversely those with very high CRE, which may be important for sporting or athletic success.

METHODS

Data sources

A systematic review of the scientific literature was registered (PROSPERO 2013:CRD42013003622) and completed to locate studies that reported descriptive 20mSRT data on 9–17 years old (see online supplement 1). Studies were identified up until October 2015 using the following bibliographic databases: MEDLINE (1946–2015), PsycINFO (1806–2014), EMBASE (1947–2014), SPORTDiscus (1949–2014) and Cochrane Central Register of Controlled Trials (2005–2014). The search strategy included the following terms: shuttle run*, OR beep test, OR multi-stage, OR aerobic, OR cardio*, OR endurance; with child*, adolescen*, pubescen*, boy, girl, young and youth as search term modifiers. All studies were extracted as text files, imported into Reference Manager (Thompson Reuters, San Francisco, California, USA), and assigned a unique reference identification number. Duplicate studies were first removed using Reference Manager with the remaining duplicates removed manually. Two independent reviewers screened all titles and abstracts for eligibility, with full-text copies obtained for all studies meeting initial screening criteria according to at least one reviewer. These two independent reviewers then examined all full-text articles and discrepancies were resolved by discussion and consensus. A third reviewer examined an article when the two reviewers were unable to reach consensus, with consensus reached for all included articles. Email contact with the corresponding authors of studies occurred when necessary, in order to provide clarification, to avoid 'double counting' previously reported data, and/or to request additional descriptive or raw data. The reference lists of all included studies were manually reviewed to identify new studies. Reviewers contacted content experts to obtain grey literature. In addition, the extensive personal libraries of the study authors were examined for relevant studies.

Inclusion/exclusion criteria

Studies were included if they explicitly reported descriptive 20mSRT data (using the 1 min stage protocol; see Tomkinson *et al.*²² for protocol variants) at the sex by age by country level. Study participants must have been apparently healthy (free from known disease or injury) 9–17 years old who were tested from 1981 onwards—the inception year of Léger's 20mSRT with 1 min stages. Studies were excluded if they reported descriptive data on: (1) other versions of the shuttle run (eg, the 15 m test or estimates of 20mSRT performance based on the 15 m test); (2) duplicate data published in another included study; or (3) on only special interest groups of children who were atypical of their source population (eg, elite athletes, physically or mentally retarded children). Figure 1 shows a flow chart of the included studies.

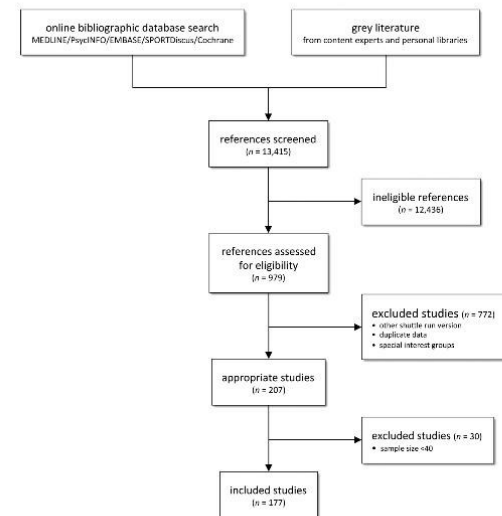


Figure 1 PRISMA flow chart outlining the identification of the included studies.

Data treatment and statistical analysis

All descriptive data were extracted into Excel (Microsoft Office 2010, USA) using a standardised data extraction table. The following descriptive data were extracted by one author and checked by another for accuracy: authors, country of testing, year of testing, sex, age, 20 m shuttle run protocol, 20 m shuttle run metric, sample size, mean, SD, median, sampling method and the sampling base. All data were examined for anomalies by running range checks and examining sex-specific and age-specific scattergrams. While only data on children and youth aged 9–17 years inclusive were retained for further analysis, individual study by sex by age by country by year groups were excluded when the sample size fell below 40 as the means and SDs for smaller samples were too labile. The final data set included 1896 study by sex by age by country by year groups.

The general procedure used to generate the sex-specific and age-specific normative centiles from extracted data is shown in figure 2. In most studies (76% or 135/177 studies), age was reported as age at last birthday; however, age was also reported in 7% (13/177) of studies as a span of years (maximum range: 3 years) and in 17% (29/177) as mean and SD years. In studies reporting age as a span of years, a new sample size was assigned to each study by sex by age by country group by dividing the reported sample size by the number of age groups (eg, a sample size of 162 was assigned to boys aged 10, 11 and 12 years in the study by Vandongen *et al.*²⁹ which reported testing 486 boys aged 10–12 years). In studies that reported age as mean $\pm$ SD years, Monte Carlo simulation was used to produce pseudo age data (using a random normal generator) based on reported means and SDs to estimate the sample size in each study by sex by age by country group.

Testing year was recorded as the midpoint year of testing (eg, 2009.5 was recorded as the measurement year for a study that reported testing children in 2009) in 55% (98/177) of studies, with 34% (60/177) reporting a span of testing years and 11% (19/177) not reporting it at all. The midpoint year was recorded for studies reporting a span of testing years (eg, 2010.0 was recorded for a study reporting testing over the period 2009–

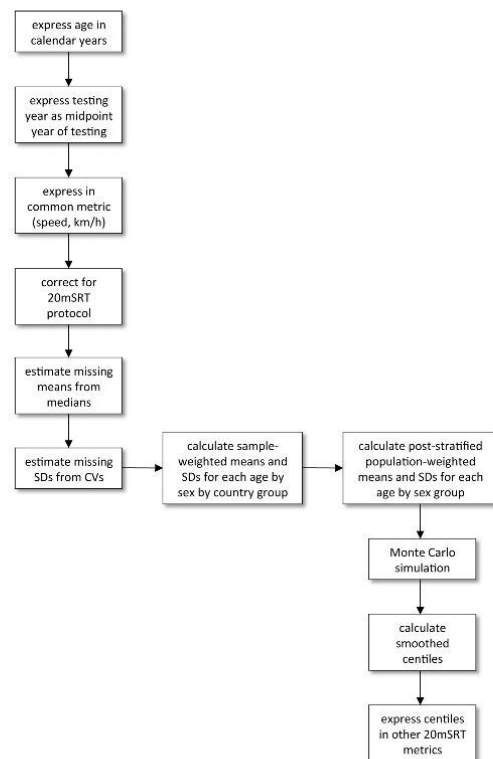


Figure 2 Flow chart showing the methodological procedure used in this study. Results from studies were first expressed in the common metric of running speed (km/h) at the final completed stage and then corrected to Léger's original 1 min protocol according to which protocol they used (Léger, Eurofit or Queen's University of Belfast). Following the estimation of missing means (from reported medians) and SDs (from calculated CVs), poststratified population-weighted means and SDs were estimated for each sex by age group, with pseudo data then generated using Monte Carlo simulation. Smoothed centiles were then generated using the LMS method, with international normative 20mSRT values expressed in several different metrics (speed at the last completed stage, the number of completed full stages/minutes and relative $\dot{V}O_{2peak}$ (ml/kg/min)). 20mSRT, 20 m shuttle run test; CV, coefficient of variation; LMS, Lambda Mu and Sigma; and $\dot{V}O_{2max}$ peak oxygen uptake.

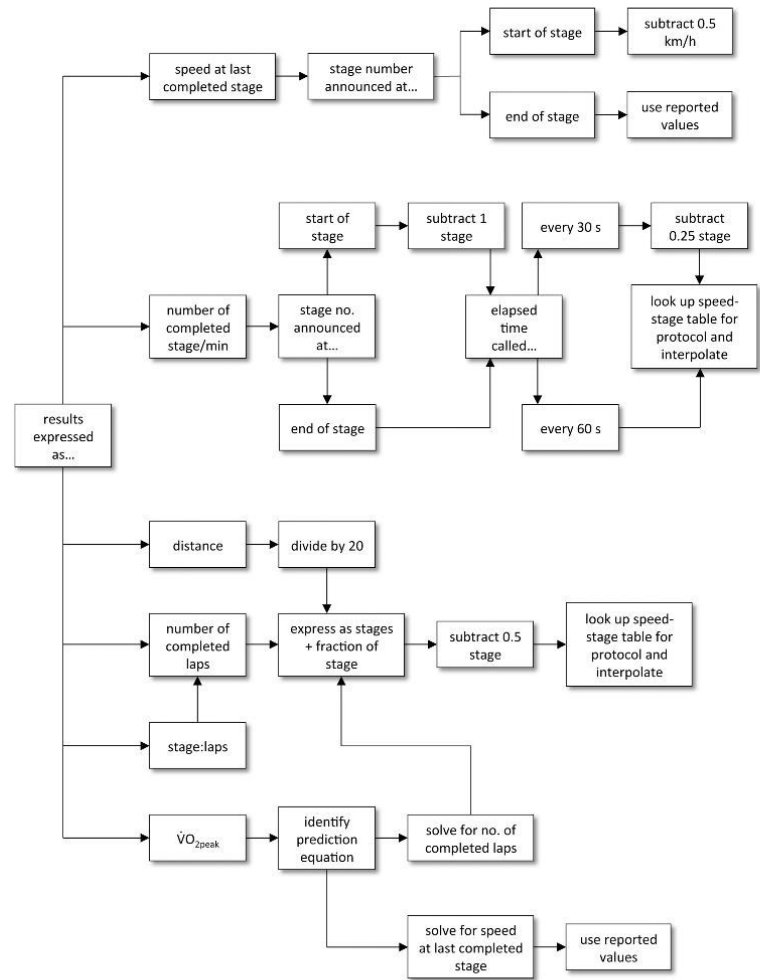
2010), with 2 years prior to the publication year assumed for studies when it was not reported, which was the median difference for those studies in which the testing year was known.

To combine data from different studies, all 20mSRT data were standardised to a common metric and for protocol differences. To do this, we used an updated version of the standardisation procedure described in detail by Tomkinson *et al.*²² Figure 3 shows this updated standardisation procedure and summarises the steps used to express 20mSRT performances in the common metric of speed (km/h) at the last completed stage. All 20mSRT data were then standardised for protocol differences to Léger's original 1 min protocol,³⁰ which starts at a speed of 8.5 km/h and increases in speed by 0.5 km/h each minute.²²

As part of the modelling procedure used to generate sex-specific and age-specific normative centiles, means and SDs were required at the study by sex by age by country by year level. If no mean was available (1% or 2/177 studies), then mean values were estimated from the reported median values. This was done by first locating all studies reporting both median and mean values at the study by sex by age by country level, and second, by determining the best-fitting and most parsimonious linear and curvilinear (second and third order polynomials) regression models between median (predictor variable) and mean (response variable) speed values. Median and mean speed values were available for 418 study by sex by age by country groups, with the relationship nearly perfectly described by the following linear regression model: $\text{mean} = 0.9408 \times \text{median} + 0.6566$ (where r (95% CI) = 0.988 (0.985 to 0.990) and $SE = 0.128$). Furthermore, 5% (9/177) of studies did not report SD values. Missing SD values were estimated by first locating all studies reporting both means and SDs at the study by sex by age by country by year level; second, by calculating the corresponding coefficient of variation (CV) values; and third, by calculating the sample-weighted mean CVs for boys and girls separately. Mean and SD speed values were available for 1585 study by sex by age by country groups, with sample-weighted mean CVs ($\pm 95\%$ CI) of $10.8 \pm 0.1\%$ and $9.2 \pm 0.1\%$ for boys and girls, respectively.

Sample-weighted means and SDs (the latter calculated from sample-weighted mean CVs) were then calculated at the sex by age by country level. While these data represent the best available 20mSRT data, in order to best generate internationally representative sex-specific and age-specific normative centiles and to correct for systematic bias associated with oversampling and undersampling, means and SDs were corrected using the post-stratification population-weighting procedure described by Levy and Lemeshow.³¹ This procedure ensures that our population estimates reflect the underlying international age-specific and sex-specific country demographics. Thus, population estimates standardised to the mean measurement year of 2000 were extracted from the United Nations World Population Prospects report.³² Monte Carlo simulation was then used to create pseudo data using a random normal generator based on population-weighted means and SDs at the sex by age level. Monte Carlo simulation assumes that the distributions are approximately normal, which is true of the available raw 20mSRT speed data. Pseudo data sets were repeatedly generated until the calculated mean differed from the reported mean by $<0.5\%$, and the calculated SD differed from the reported SD by $<2.5\%$. These pseudo data sets were then used to generate sex-specific and age-specific normative centiles in LMSchartmaker Pro (V2.43, The Institute of Child Health, London), which analyses data using the Lambda Mu and Sigma (LMS).³³ The LMS method fits smooth centile curves to reference data by summarising the changing distribution of three sex-specific and age-specific curves representing the skewness (L ; expressed as a Box-Cox power), the median (M) and the CV (S). Using penalised likelihood, the curves can be fitted as cubic splines using non-linear regression, and the extent of smoothing required can be expressed in terms of smoothing parameters or equivalent degrees of freedom.³⁴ The effective degrees of freedom for 20mSRT speed were 1 (L curve), 4 (M curve) and 3 (S curve) for boys and 1 (L curve), 3 (M curve) and 3 (S curve) for girls. Normative centiles were also expressed in other common 20mSRT metrics, including the number of completed stages/minutes, the number of completed laps and relative peak

Figure 3 Performance metric conversion flow chart adapted from Tomkinson *et al.*²² $\dot{V}O_{2peak}$, peak oxygen uptake.



oxygen uptake ($\dot{V}O_{2peak}$, mL/kg/min) values using the Léger *et al.*³⁵ prediction equation:

$$\dot{V}O_{2peak}(\text{mL/kg/min}) = 31.025 + 3.238 \text{ speed} - 3.248 \text{ age} + 0.1536 \text{ speed} \times \text{age}$$

where speed is the running speed of the last completed stage (km/h) and age is age at last birthday. In a sample of Québec children and youth, this equation had a SE of estimate of 5.9 mL/kg/min or 12.1%.³⁵

The prevalence of children and youth (10–17 years old) with 'healthy' CRE was estimated using the new sex-specific and age-specific FITNESSGRAM criterion-referenced standards for $\dot{V}O_{2peak}$.^{36,37} Differences in mean 20mSRT performance (km/h) between (1) age-matched boys and girls (eg, boys aged 10 years vs girls aged 10 years), and (2) sex-matched children of different ages (eg, boys aged 10 vs 11 years), were expressed as absolute and standardised differences. Positive differences indicated that

20mSRT performances for boys or older children were higher than those for girls or younger children. Standardised differences of 0.2, 0.5 and 0.8 were used as thresholds for small, moderate and large, respectively.³⁸

RESULTS

The final data set included 1 142 026 children and youth aged 9–17 years (1896 study by sex by age by country by year groups extracted from 177 studies) from 50 countries tested over the period 1981–2014 (figure 4, see online supplement 2^{w39–w190}). These 50 countries represented six major geographical regions (Africa, Asia, Europe, Latin America and the Caribbean, Northern America and Oceania),³² including 33 high-income, 9 upper middle-income, 5 lower middle-income and 3 low-income economies.¹⁹¹ Norms are presented as tabulated centiles from 5% to 95% for four common 20mSRT metrics in tables 1–4, with the smoothed centile curves for 20mSRT speed presented in figure 5 and the sex-specific and age-specific LMS values shown in online supplement 3.



Figure 4 World map representing the 50 countries (shown in black) for which 20mSRT data on 9–17 years old were available. 20mSRT, 20 m shuttle run test.

On average, approximately two-thirds of boys (mean $\pm$ 95% CI: 67 $\pm$ 14%) and half of the girls (mean $\pm$ 95% CI: 54 $\pm$ 17%) had healthy CRE, with the prevalence of healthy CRE decreasing by about 8% (boys) and 10% (girls) with every increasing year from age 10 to 17 years (figure 6).

Boys consistently outperformed girls at each age group (mean difference $\pm$ 95% CI: 0.86 $\pm$ 0.28 km/h or 0.79 $\pm$ 0.20 standardised units), with the sex-related differences increasing with age from a small difference at age 9 years (difference in means $\pm$ 95% CI: 0.32 $\pm$ 0.01 km/h or 0.40 $\pm$ 0.01 standardised units)

Table 1 Twenty-metre shuttle run (speed (km/h) at the last complete stage) centiles by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981

Age (years)	P ₅	P ₁₀	P ₂₀	P ₃₀	P ₄₀	P ₅₀	P ₆₀	P ₇₀	P ₈₀	P ₉₀	P ₉₅
Boys											
9	8.63	8.94	9.31	9.58	9.82	10.03	10.25	10.48	10.75	11.13	11.44
10	8.61	8.95	9.35	9.65	9.90	10.13	10.36	10.61	10.90	11.31	11.64
11	8.60	8.97	9.41	9.72	10.00	10.25	10.50	10.77	11.09	11.53	11.89
12	8.64	9.05	9.53	9.89	10.19	10.47	10.75	11.05	11.40	11.89	12.29
13	8.74	9.18	9.72	10.10	10.43	10.73	11.04	11.37	11.75	12.29	12.72
14	8.85	9.32	9.88	10.29	10.64	10.96	11.29	11.64	12.04	12.61	13.08
15	8.94	9.42	10.01	10.43	10.79	11.13	11.47	11.83	12.25	12.84	13.32
16	9.01	9.51	10.11	10.55	10.92	11.27	11.62	11.99	12.42	13.03	13.53
17	9.08	9.60	10.22	10.67	11.05	11.41	11.77	12.16	12.60	13.23	13.74
Girls											
9	8.56	8.82	9.13	9.35	9.54	9.72	9.90	10.08	10.31	10.61	10.87
10	8.48	8.76	9.10	9.35	9.56	9.75	9.95	10.15	10.40	10.74	11.01
11	8.42	8.72	9.09	9.35	9.57	9.78	10.00	10.22	10.48	10.85	11.15
12	8.37	8.69	9.08	9.36	9.60	9.83	10.05	10.29	10.57	10.95	11.27
13	8.36	8.69	9.09	9.38	9.63	9.86	10.09	10.34	10.63	11.03	11.36
14	8.36	8.70	9.11	9.40	9.65	9.89	10.12	10.37	10.67	11.07	11.41
15	8.36	8.70	9.12	9.42	9.67	9.91	10.15	10.40	10.70	11.11	11.44
16	8.36	8.71	9.13	9.43	9.69	9.93	10.17	10.43	10.73	11.14	11.49
17	8.36	8.72	9.14	9.45	9.71	9.96	10.20	10.46	10.77	11.19	11.54

The ages shown represent age at last birthday (eg, 9=9.00–9.99).

Table 2 Twenty-metre shuttle run (number of completed stages/minutes) centiles by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981

Age (years)	P ₅	P ₁₀	P ₂₀	P ₃₀	P ₄₀	P ₅₀	P ₆₀	P ₇₀	P ₈₀	P ₉₀	P ₉₅
Boys											
9	1.25	1.87	2.63	3.17	3.63	4.07	4.50	4.96	5.50	6.26	6.88
10	1.23	1.90	2.71	3.29	3.79	4.26	4.72	5.22	5.80	6.61	7.28
11	1.21	1.94	2.82	3.45	3.99	4.50	5.00	5.55	6.18	7.06	7.78
12	1.29	2.09	3.07	3.77	4.37	4.93	5.50	6.10	6.80	7.77	8.58
13	1.48	2.36	3.43	4.20	4.85	5.47	6.08	6.74	7.51	8.57	9.45
14	1.70	2.64	3.77	4.58	5.28	5.93	6.58	7.28	8.09	9.22	10.15
15	1.87	2.84	4.01	4.86	5.58	6.26	6.93	7.66	8.50	9.67	10.64
16	2.02	3.01	4.22	5.10	5.84	6.54	7.23	7.98	8.85	10.06	11.05
17	2.17	3.20	4.44	5.34	6.11	6.83	7.54	8.31	9.21	10.45	11.48
Girls											
9	1.12	1.63	2.25	2.70	3.08	3.43	3.79	4.17	4.61	5.23	5.74
10	0.97	1.53	2.21	2.70	3.11	3.50	3.89	4.31	4.80	5.47	6.03
11	0.83	1.44	2.17	2.70	3.15	3.57	3.99	4.44	4.96	5.69	6.29
12	0.74	1.39	2.17	2.73	3.21	3.65	4.10	4.57	5.13	5.90	6.54
13	0.72	1.38	2.19	2.77	3.26	3.72	4.19	4.68	5.25	6.05	6.71
14	0.72	1.40	2.21	2.80	3.31	3.78	4.24	4.75	5.33	6.14	6.81
15	0.73	1.41	2.24	2.83	3.34	3.82	4.29	4.80	5.39	6.21	6.89
16	0.73	1.42	2.26	2.87	3.38	3.86	4.34	4.86	5.46	6.29	6.97
17	0.73	1.44	2.29	2.90	3.43	3.92	4.40	4.93	5.54	6.38	7.08

to a large difference at age 17 years (difference in means $\pm$ 95% CI: 1.46 ± 0.02 km/h or 1.20 ± 0.02 standardised units; figure 7A). From age 9 to 17 years, boys' performance improved at the rate of 0.17 km/h (or 0.15 standardised units) per year, with the largest rate of increase occurring at age 12 years (0.27 km/h or 0.23 standardised units). Girls' performance steadily improved at the rate of 0.03 km/h (or 0.03 standardised units) per year (figure 7B). These age-related changes were cumulatively large for boys and small for girls.

DISCUSSION

This study systematically analysed 20mSRT data on 1 142 026 children and youth aged 9–17 years to generate the most comprehensive and up-to-date set of international sex-specific and age-specific norms for CRE. These international norms have utility for health and fitness screening, athlete profiling, and monitoring and surveillance in health, clinical, educational or sporting settings. They complement a growing body of literature reporting national, regional and international growth centiles

Table 3 Twenty-metre shuttle run (number of laps) centiles by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981

Age (years)	P ₅	P ₁₀	P ₂₀	P ₃₀	P ₄₀	P ₅₀	P ₆₀	P ₇₀	P ₈₀	P ₉₀	P ₉₅
Boys											
9	9	14	20	24	28	32	36	40	45	52	58
10	9	14	21	25	29	33	37	42	47	55	62
11	9	15	22	27	31	36	40	45	51	60	67
12	9	16	24	29	34	39	45	50	57	67	75
13	11	18	26	33	39	44	50	56	64	75	84
14	13	20	29	36	43	48	55	62	70	81	92
15	14	22	31	39	45	52	58	66	74	86	97
16	15	23	33	41	48	54	61	69	78	91	102
17	16	25	35	43	50	57	64	72	81	95	107
Girls											
9	8	12	17	21	24	26	29	33	36	42	47
10	7	11	17	21	24	27	30	34	38	44	49
11	6	11	16	21	24	28	31	35	40	46	52
12	5	10	16	21	25	28	32	36	41	48	54
13	5	10	17	21	25	29	33	37	42	50	56
14	5	10	17	21	25	29	33	38	43	50	57
15	5	10	17	22	26	30	34	38	44	51	58
16	5	10	17	22	26	30	34	39	44	52	59
17	5	11	17	22	26	30	35	39	45	53	60

Table 4 Relative peak oxygen uptake ($\dot{V}O_{2peak}$, mL/kg/min) centiles by age and sex in 1 142 026 children and youth aged 9–17 years from 50 countries since 1981

Age (years)	P ₅	P ₁₀	P ₂₀	P ₃₀	P ₄₀	P ₅₀	P ₆₀	P ₇₀	P ₈₀	P ₉₀	P ₉₅
Boys											
9	41.7	43.1	44.8	46.1	47.1	48.1	49.1	50.2	51.5	53.2	54.6
10	39.7	41.3	43.2	44.6	45.8	46.9	48.0	49.2	50.6	52.5	54.1
11	37.7	39.5	41.7	43.2	44.6	45.8	47.0	48.4	49.9	52.1	53.9
12	36.0	38.0	40.5	42.3	43.8	45.2	46.7	48.2	50.0	52.4	54.5
13	34.6	36.9	39.7	41.7	43.4	45.0	46.6	48.3	50.3	53.1	55.4
14	33.3	35.8	38.8	41.0	42.9	44.6	46.4	48.3	50.5	53.5	56.0
15	31.8	34.5	37.8	40.1	42.1	44.0	45.9	47.9	50.2	53.4	56.1
16	30.4	33.2	36.6	39.1	41.3	43.2	45.2	47.3	49.8	53.3	56.1
17	28.9	32.0	35.6	38.2	40.5	42.6	44.7	46.9	49.5	53.2	56.2
Girls											
9	41.3	42.5	44.0	45.0	45.9	46.7	47.5	48.4	49.4	50.8	52.0
10	39.0	40.4	42.0	43.2	44.2	45.1	46.0	47.0	48.2	49.8	51.1
11	36.8	38.3	40.1	41.4	42.5	43.5	44.5	45.7	47.0	48.7	50.2
12	34.6	36.2	38.2	39.6	40.8	42.0	43.1	44.3	45.7	47.7	49.3
13	32.6	34.3	36.4	37.9	39.2	40.4	41.6	42.9	44.4	46.5	48.2
14	30.6	32.4	34.6	36.2	37.6	38.8	40.1	41.4	43.0	45.2	47.0
15	28.7	30.5	32.8	34.5	35.9	37.2	38.5	39.9	41.6	43.9	45.7
16	26.7	28.7	31.1	32.8	34.2	35.6	37.0	38.4	40.2	42.5	44.5
17	24.7	26.8	29.3	31.1	32.6	34.1	35.5	37.0	38.8	41.3	43.3

across a range of cardiometabolic disease risk factors, including adiposity (eg, body mass index^{192–193} and waist circumference^{194–198}), blood pressure,^{199–200} cholesterol,²⁰⁰ triglycerides,²⁰⁰ glucose²⁰⁰ and health-related fitness.^{28–50–126–201–203}

Using a quintile framework, children in the bottom 20% can be classified as having ‘very low’ CRE; between the 20th and 40th centiles as having ‘low’ CRE; between the 40th and 60th centiles as having ‘moderate’ CRE; between the 60th and 80th centiles as having ‘high’ CRE; and above the 80th centile as having ‘very high’ CRE. Single measures of 20mSRT performance taken in health, clinical, educational or sporting settings can then be qualitatively interpreted using these quintile-based thresholds, with longitudinal changes (eg, due to growth or exercise training interventions) monitored by tracking changes against centile bands. For example, Armstrong *et al*¹ estimated that in children an appropriate 12-week CRE training programme will induce, on average, an 8–9% increase in $\dot{V}O_{2peak}$ independent of sex, age and maturation, equivalent to an increase of ~0.5 standardised units or ~20 centile points, which should be enough for a child to shift upwards from one quintile band to the next or above the relevant criterion-referenced standard for low cardiometabolic risk.²⁰⁴

While these norms are not criterion-referenced in that they do not indicate whether children have healthy CRE or low cardiometabolic risk, this study does provide an estimate of the prevalence with healthy CRE according to the new FITNESSGRAM (V10) standards, which have been shown to discriminate with moderate accuracy between youth with and without metabolic syndrome.³⁷ There are currently no agreed on international criterion-referenced standards for CRE, and while we estimated prevalence using the US-based FITNESSGRAM standards, other national and regional standards have been published elsewhere.^{26–205–211} It is important to note that the differences between published standards are (in some cases) substantial, with the new FITNESSGRAM (V10) standards generally higher for girls and lower for boys relative to other standards,^{26–205–211} meaning that our prevalence

estimates would be substantially different if calculated using other standards. In contrast, $\dot{V}O_{2peak}$ can be estimated using different test protocols and prediction equations, and special care must be taken when comparing $\dot{V}O_{2peak}$ values with standards that were estimated using different test protocols and prediction equations.²¹² For example, we predicted $\dot{V}O_{2peak}$ from 20mSRT performance using the Léger *et al*³⁵ equation, whereas the new FITNESSGRAM standards were developed using predicted $\dot{V}O_{2peak}$ from a laboratory-based treadmill test.³⁷ To our knowledge, at least 17 prediction equations (from 10 studies of apparently healthy children and youth^{37–213–221}) are currently available to estimate $\dot{V}O_{2peak}$ from 20mSRT performance, equations that differ in validity and can result in substantially different estimates of $\dot{V}O_{2peak}$ and hence the prevalence of healthy CRE. Future studies need to build on multinational efforts (eg, the HELENA study in Europe) in order to develop the most valid international normative-referenced and criterion-referenced standards for CRE.²²²

Using a cross-sectional approach, this study quantified the sex-related and age-related differences in 20mSRT speed, showing that boys consistently outperformed girls and experienced larger age-related changes. Given that longitudinal data are required to determine the true developmental patterns of CRE (because they control for within-participant changes in timing and tempo), and that larger age-related increases have been observed in children followed longitudinally than in those examined cross-sectionally,²²³ it is possible that our data underestimate the true developmental patterns. While the developmental patterns of children's $\dot{V}O_{2peak}$ has been well studied in non-representative longitudinal samples,^{224–226} other aspects of CRE (eg, mechanical efficiency, fractional utilisation, $\dot{V}O_2$ kinetics, etc) are less well understood, making it difficult to describe the mechanistic causes underlying the developmental patterns in 20mSRT performance. Given that relative $\dot{V}O_2$ and $\dot{V}O_{2peak}$ vary linearly with speed and peak speed, changes in 20mSRT speed should therefore reflect changes in underlying $\dot{V}O_2$ (ie, the oxygen cost of the activity). Interestingly, however, this

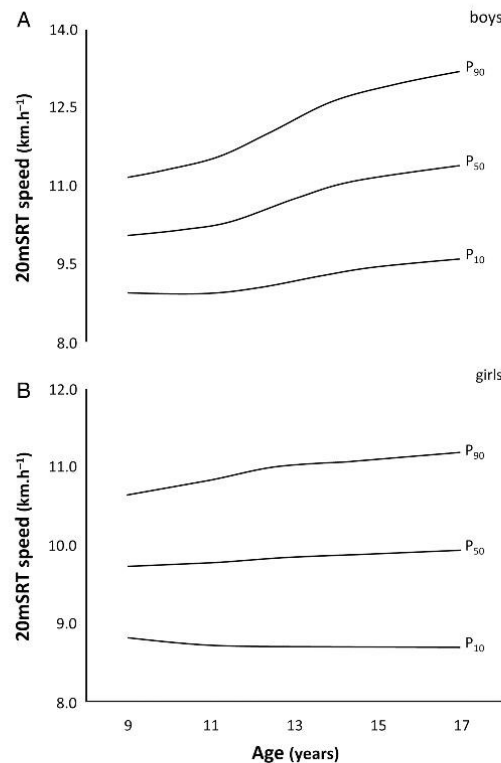


Figure 5 Smoothed centile curves (P_{10} , P_{50} and P_{90}) for the 20mSRT (speed (km/h) at the last completed stage) performance of (A) boys and (B) girls. 20mSRT, 20 m shuttle run test.

study observed decreases in boys' relative $\dot{V}O_{2peak}$ throughout childhood and adolescence, rather than the expected plateau.²²⁷ This unexpected developmental pattern may be due in part to $\dot{V}O_{2peak}$ prediction error or the fact that the developmental patterns were calculated using pooled cross-sectional data. In

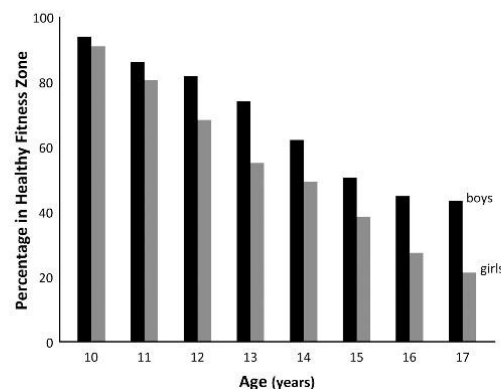


Figure 6 Prevalence of 10–17 years old from 50 countries meeting the FITNESSGRAM Healthy Fitness Zone (V10) thresholds for $\dot{V}O_{2peak}$ (mL/kg/min). $\dot{V}O_{2peak}$, peak oxygen uptake.

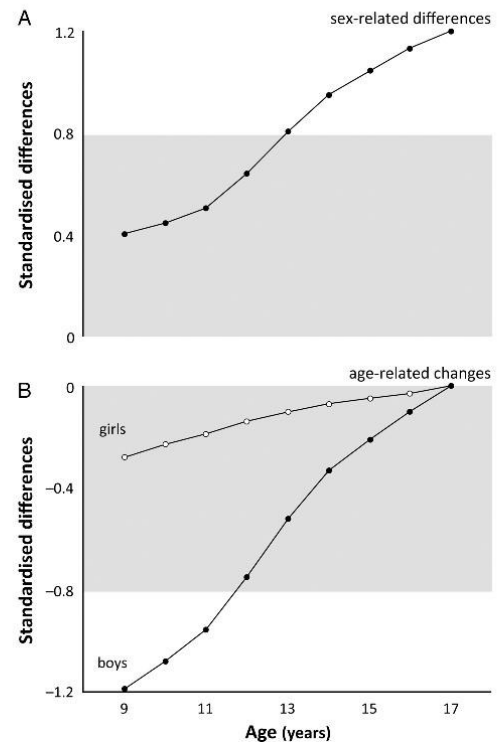


Figure 7 Standardised differences in mean 20mSRT performance (speed (km/h) at the last completed stage) between (A) age-matched boys and girls and (B) sex-matched children of different ages (anchored to age 17 years=0). Positive differences indicate that 20mSRT performances were higher for boys than for girls (top panel) or for older children than for younger children (bottom panel). The limits of the grey zone represent the thresholds for a large standardised difference (0.8 or -0.8). 20mSRT, 20 m shuttle run test.

contrast, the premise of a plateau in boys' relative $\dot{V}O_{2peak}$ throughout childhood and adolescence is largely based on progressive treadmill or cycle testing of volunteer recruits who were athletically inclined, non-obese and motivated to exercise.^{226 228} This unexpected finding in the developmental pattern in boys' $\dot{V}O_{2peak}$ might reflect current maturational development, or our use of a large, internationally representative sample of children and youth.

Strengths and limitations

This study brings together 20mSRT data from 177 studies in what is to date the largest international CRE database in the world. We conducted a systematic review, using a strict set of inclusion and exclusion criteria, applied rigorous data treatment procedures to systematically control for bias (eg, differences in test protocols and performance metrics), a poststratification population weighting procedure, a novel pseudo data method and the LMS method, to generate international sex-specific and age-specific smoothed centiles (across four performance metrics) for CRE. While it is not the first comprehensive study of children's 20mSRT performance, it updates the comprehensive reviews of Olds *et al*²¹ and Tomkinson *et al*²² by: (1) extending the data coverage from 2003 to 2014 through a rigorous

systematic review process, (2) producing sex-specific and age-specific international normative centiles and (3) estimating the international prevalence of children and youth with healthy CRE.

Unfortunately, there are several sampling-related and methodology-related limitations to this study. First, it pooled data from studies that used different sampling methods which raises the issue of representativeness, with some studies using probability sampling (eg, stratified proportional, stratified random or cluster sampling) and others non-probability sampling (eg, convenience sampling; see online supplement 2). Second, differences in sampling base also exist across national, state/provincial and local/city/school levels. Third, differences in testing conditions (eg, environmental conditions such as climate or altitude, practice and running surfaces) and measurement errors (eg, calibration and type of equipment, methodological drift and diurnal variation) are inherent to any large data synthesis, but the very large number of data points captured should minimise these issues. Fourth, the vigorous nature of the 20mSRT may result in difficulties in test administration in, or exclusion of individuals with physical disabilities and diseases. The absence of data from these populations is likely to have inflated our normative values within the lower centile range.

The 20mSRT data were also collected at different times over the period 1981–2014, and given convincing evidence of international declines in children's CRE in recent decades,^{1 23 229} our norms may be biased. With a mean measurement year of 2000, and assuming an international decline of ~5% per decade since 1975,^{1 23} our normative data may overestimate values in 2014 by ~7.5% (equivalent to ~0.65 standardised units or ~24 centile points) and represent a better health-related picture than what would be true today. However, without time trend data for all included countries and evidence of recent improvements in children's CRE in some included countries (eg, Australia,²⁸ Japan²²⁹ and Spain²³⁰), time-based corrections of our normative data are not recommended. Despite relying predominantly on data from high-income economies, there is no good evidence that 20mSRT performance is meaningfully related to a country's affluence or distribution of wealth,^{21 231} so our norms are unlikely to be biased. Future studies need to examine CRE in low-income and middle-income economies and at multiple stages of the CRE transition. It must also be remembered that the 20mSRT is affected by factors other than underlying construct CRE.^{20 232} Validity data show that a moderate-to-large (35–70%) amount of the variance in 20mSRT performance is explained by the variance in underlying $\text{VO}_{2\text{peak}}$,^{20 232 233} indicating that other physiological (eg, mechanical efficiency),²³⁴ fractional utilisation,²³⁵ VO_2 kinetics,²³⁶ lactate threshold,²³⁷ anaerobic capacity,²³⁸ physical (eg, fat mass)²³⁹ and psychosocial factors (eg, motivation, effort and self-efficacy²⁴⁰) also contribute.

Recommendations

Over the past few decades, the 20mSRT has been widely used to assess the CRE of children and youth, and yet data pooling is nearly impossible due to the difficulty with standardising performances (eg, because of differences in protocols, performance metrics, the way in which age is expressed, etc). To facilitate data pooling in the future, and to assist with the eventual update of 20mSRT norms, we make the following recommendations:

1. The test protocol used should be thoroughly and accurately reported;
2. Care should be taken to minimise and report factors that affect 20mSRT performance (eg, testing conditions and measurement errors);

3. Best practice should include 20mSRT results that are reported as the running speed (km/h) at the last completed stage;
4. Descriptive statistics (sample sizes, means and SDs) should be reported at the sex by age (at last birthday) level; and
5. The year(s) of testing should be reported;

Furthermore, because the 20mSRT is a maximal effort test, in order to ensure that a child has performed with 'good effort' (ie, they have tried very hard), perceptual (eg, ratings of perceived exertion) and/or physiological (eg, heart rate) effort should be measured in addition to performance effort.²⁴⁰ Any adverse events (or lack thereof) associated with maximal effort tests such as the 20mSRT should also be reported.²⁴¹

CONCLUSION

CRE is considered to be an excellent marker of current and future health. The 20mSRT is arguably the most popular measure of CRE because it is suitable for mass testing, is simple, cheap, easy, reliable, reasonably valid and is part of widely used health-related fitness test batteries (eg, Assessing Levels of Physical Activity and fitness (ALPHA),²⁴² Canadian Assessment of Physical Literacy (CAPL),²⁴³ Eurofit,²⁴⁴ FITNESSGRAM²⁴⁵ and even the PREFIT battery (Assessing FITness in PRESchoolers)).^{246 247} Using a systematic review and analytical approach, this study used the best available 20mSRT data to: (1) provide the most comprehensive and up-to-date set of international sex-specific and age-specific norms for children and youth; and (2) estimate the prevalence with healthy CRE according to the FITNESSGRAM standards. These data have utility for health and sport promotion given that they help to identify children and youth with: (1) very low CRE in order to set appropriate fitness goals, monitor longitudinal changes and promote positive health-related fitness behaviours; and (2) very high CRE in the hope of recruiting them into elite sporting or athletic development programmes.

What are the findings?

- This study provides the most comprehensive and up-to-date set of sex-specific and age-specific international norms for the 20 m shuttle run test in children and youth.
- This study estimates the international prevalence of children and youth with healthy cardiorespiratory endurance according to the FITNESSGRAM standards.
- These international norms have utility for health and fitness screening, profiling, monitoring and surveillance.

How might it impact on clinical practice in the future?

International sex-specific and age-specific 20mSRT norms for children and youth are important for health and fitness screening, profiling, monitoring and surveillance.

Author affiliations

¹Department of Kinesiology and Public Health Education, University of North Dakota, Grand Forks, North Dakota, USA

²Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health Sciences & Sansom Institute for Health Research, University of South Australia, Adelaide, South Australia, Australia

³Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada

⁴School of Health Sciences, University of South Australia, Adelaide, South Australia, Australia

⁵ParticipACTION, Toronto, Ontario, Canada

⁶University of Granada, Granada, Spain

⁷Karolinska Institute, Huddinge, Sweden

⁸Département de kinésiologie, Université de Montréal, Montréal, Québec, Canada

Acknowledgements The authors would like to thank the authors of the included studies for generously clarifying details of their studies and/or for providing raw data.

Contributors GRT, JLL and MST developed the systematic review research question and objectives. GRT, JLL, AGL and KB created the search strategy and provided guidance on review methodology. GRT, JLL, MD, AGL and KB screened and extracted the data. GRT and JLL led data analysis, synthesis and writing of the manuscript. All authors contributed to interpretation of the results, edited, reviewed and approved the final manuscript.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- Armstrong N, Tomkinson GR, Ekelund U. Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during youth. *Br J Sports Med* 2011;45:849–58.
- Institute of Medicine. Health-related fitness measures for youth: cardiorespiratory endurance. In: Institute of Medicine, ed. *Fitness measures and health outcomes in youth*. Washington DC: The National Academy Press, 2012:111–51.
- Blair SN, Kohl HW III, Paffenbarger RS Jr, et al. Physical fitness and all-cause mortality. A prospective study of healthy men and women. *JAMA* 1989;262:2395–401.
- 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;301:2024–35.
- Lee CD, Blair SN. Cardiorespiratory fitness and stroke mortality in men. *Med Sci Sports Exerc* 2002;34:592–5.
- Katzmarzyk P, Church TS, Janssen I, et al. Metabolic syndrome, obesity, and mortality: impact of cardiorespiratory fitness. *Diabetes Care* 2005;28:391–7.
- Dishman RK, Washburn RA, Heath GW. *Physical activity epidemiology*. Champaign, IL: Human Kinetics, 2004:358–9.
- Gillespie LD, Robertson MC, Gillespie WJ, et al. Interventions for preventing falls in older people living in the community. *Cochrane Database Syst Rev* 2009;(2): CD007146.
- Katzmarzyk PT, Church TS, Blair SN. Cardiorespiratory fitness attenuates the effects of the metabolic syndrome on all-cause and cardiovascular disease mortality in men. *Arch Intern Med* 2004;164:1092–7.
- Ortega FB, Ruiz JR, Castillo MJ, et al. Physical fitness in children and adolescence: a powerful marker of health. *Int J Obes* 2008;32:1–11.
- Ruiz JR, Castro-Piñero J, Antero EG, et al. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009;43:909–23.
- Eisenmann JC, Katzmarzyk PT, Perusse L. Aerobic fitness, body mass index, and CVD risk factors among adolescents: the Québec family study. *Int J Obes* 2005;29:1077–83.
- Malina RM. Physical activity and fitness: pathways from childhood to adulthood. *Am J Hum Biol* 2001;13:162–72.
- Ortega FB, Ruiz JR, Labayen I, et al. Role of socio-cultural factors on changes in fitness and adiposity in youth: a 6-year follow-up study. *Nutr Metab Cardiovasc Dis* 2013;23:883–90.
- Dwyer T, Magnusson CG, Schmidt MD, et al. Decline in physical fitness from childhood to adulthood associated with increased obesity and insulin resistance in adults. *Diabetes Care* 2009;32:683–7.
- Ortega FB, Silventoinen K, Tynelius P, et al. Muscular strength in male adolescents and premature death: cohort study of one million participants. *BMJ* 2012;345:e7279.
- Ortega FB, Labayen I, Ruiz JR, et al. Improvements in fitness reduce the risk of becoming overweight across puberty. *Med Sci Sports Exerc* 2011;43:1891–7.
- Morales PF, Sánchez-López M, Moya-Martínez M, et al. Health-related quality of life, obesity, and fitness in schoolchildren: the Cuenca study. *Qual Life Res* 2013;22:1515–23.
- Moliner-Urdiales D, Ruiz JR, Ortega FB, et al. Association of physical activity with muscular strength and fat-free mass in adolescents: the HELENA study. *Eur J Appl Physiol* 2010;109:1119–27.
- Tomkinson GR, Olds TS. Field tests of fitness. In: Armstrong N, van Mechelen W, eds. *Paediatric exercise science and medicine*. New York, NY: Oxford University Press, 2008:109–28.
- Olds TS, Tomkinson GR, Léger L, et al. 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;24:1025–38.
- Tomkinson GR, Léger LA, Olds TS, et al. 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;33:285–300.
- Tomkinson G, Olds T. Secular changes in pediatric aerobic fitness test performance: the global picture. *Med Sport Sci* 2007;50:46–66.
- Ruiz JR, Silva G, Oliveira N, et al. Criterion related validity of the 20m shuttle run test in adolescents aged 13–19 years. *J Sports Sci* 2009;27:899–906.
- Melo X, Santa-Clara H, Almeida JP, et al. Comparing several equations that predict peak VO₂ using the 20-m multistage-shuttle run-test in 8-10-year-old children. *Eur J Appl Physiol* 2011;111:839–49.
- Silva G, Aires L, Mota J, et al. Normative and criterion-related standards for shuttle run performance in youth. *Pediatr Exerc Sci* 2012;24:157–69.
- Carrel AL, Bowser J, White D, et al. Standardized childhood fitness percentiles derived from school-based testing. *J Pediatr* 2012;161:120–4.
- Catley MJ, Tomkinson GR. Normative health-related fitness values for children: analysis of 85347 test results on 9-17-year-old Australians since 1985. *Br J Sports Med* 2013;47:98–108.
- Vandongen R, Jenner DA, Thompson C, et al. A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10- to 12-year-old children. *Prev Med* 1995;24:9–22.
- Léger L, Lambert J, Goulet A, et al. Capacité aérobie des Québécois de 6 à 17 ans —Test navette de 20 mètres avec paliers de 1 minute. *Can J Appl Sport Sci* 1984;9:64–9.
- Levy PS, Lemeshow S. Stratification random sampling: further issues. In: Levy PS, Lemeshow S, eds. *Sampling of populations: methods and application*. Hoboken, NJ: John Wiley & Sons, Inc, 2008:143–88.
- United Nations, Department of Economic and Social Affairs, Population Division (2015). *World Population Prospects: The 2015 Revision, Key Findings and Advance Tables Working Paper No. ESA/P/WP.241*.
- Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stats Med* 1992;11:1305–19.
- Pan H, Cole T. *User's guide to LMSchartmaker*. UK: Medical Research Council, 2010:1–42.
- 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.
- Plowman SA, Meredith MD. *Fitnessgram/activitygram reference guide*. 4th edn. Dallas, TX: The Cooper Institute, 2013.
- Welk GJ, Laursen KR, Eisenmann JC, et al. Development of youth aerobic-capacity standards using receiver operating characteristic curves. *Am J Prev Med* 2011;41: S111–16.
- Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd edn. New Jersey: Lawrence Erlbaum, 1988.
- Secchi JD, García GC, Espana-Romero V, et al. Physical fitness and future cardiovascular risk in Argentine children and adolescents: an introduction to the ALPHA test battery. *Arch Argent Pediatr* 2014;112:132–40.
- Hardy LL, King L, Espinel P, et al. *NSW Schools Physical Activity and Nutrition Survey (SPANS) 2010: full report*. Sydney: NSW Ministry of Health, 2010.
- McIntyre S. *Trends in aerobic fitness from 1994 to 2009 in 10- and 11-year-old Australian children [thesis]*. Notre Dame, IN: University of Notre Dame, 2009.
- Booth M, Okely AD, Denney-Wilson E, et al. *NSW Schools Physical Activity and Nutrition Survey (SPANS) 2004: full report*. Sydney: NSW Department of Health, 2006.
- Lloyd KC, Antonas KN. Nutritional habits and fitness levels of schoolchildren. In: *Proceedings of the Nutrition Society of Australia—twenty-fourth annual scientific meeting, Fremantle, WA, 3-6 December*. Adelaide, SA: Nutrition Society of Australia, 2000:138.
- Cooley D, McNaughton L. Aerobic fitness of Tasmanian secondary school children using the 20-m shuttle run test. *Percept Mot Skills* 1999;88:188–98.
- Booth M, Macaskill P, McLellan L, et al. *NSW schools fitness and physical activity survey 1997*. Sydney: NSW Department of Education and Training, 1997.
- Okely AD, Gray T, Cotton WG. Effect of an extended stay outdoor education program on aerobic fitness. In: Gray T, Hayllar B, eds. *Catalysts for change: proceedings from the 10th National Outdoor Education Conference*. Sydney, Australia, 1997:206–10.
- Australian Council for Health, Physical Education and Recreation. *Australian fitness education award: user's manual and curriculum ideas*. Adelaide, SA: Australian Council for Health, Physical Education and Recreation, 1996.
- Australian Sports Commission. *Norms for sport related fitness tests in Australian students aged 12-17 years*. Belconnen, ACT: Australian Sports Commission, 1994.
- Jenner DA, Vandongen R, Bellin L. Relationships between blood pressure and measures of dietary energy intake, physical fitness, and physical activity in Australian children aged 11-12 years. *J Epidemiol Community Health* 1992;46:108–13.

- 50 Ortega FB, Artero EG, Ruiz JR, et al. Physical fitness levels among European adolescents: the HELENA study. *Br J Sports Med* 2011;45:20–9.
- 51 Vandendriessche JB, Vandorpe BFR, Vaeyens R, et al. Variation in sport participation, fitness and motor coordination with socioeconomic status among Flemish children. *Pediatr Exerc Sci* 2012;24:113–28.
- 52 Seghers J, Rutten C. Clustering of multiple lifestyle behaviours and its relationship with weight status and cardiorespiratory fitness in a sample of Flemish 11- to 12-year olds. *Public Health Nutr* 2010;13:1838–46.
- 53 Verstraete SJ, Cardon GM, De Clercq DL, et al. A comprehensive physical activity promotion programme at elementary school: the effects on physical activity, physical fitness and psychosocial correlates of physical activity. *Public Health Nutr* 2007;10:477–84.
- 54 Cardon G, De Bourdeaudhuij I, De Clercq D, et al. Physical fitness, physical activity, and self-reported back and neck pain in elementary schoolchildren. *Pediatr Exerc Sci* 2004;16:147–57.
- 55 Heyters C, Marique T. *Le baromètre de la condition physique*. Bruxelles: Ministère de la Communauté française, Direction générale du sport, 2004.
- 56 Telama R, Naul R, Nupponen H, et al. *Physical fitness, sporting lifestyles and Olympic ideals: cross-cultural studies on youth sport in Europe*. Schorndorf, Germany: Verlag Karl Hofmann, 2002.
- 57 Baquet G, Berthoin S, Padovano C, et al. Effects d'un cycle de course de durée de type intermittent (court-court) sur la condition physique des adolescents. *Rev Educ Physique* 2000;40:51–60.
- 58 Lefèvre J, Bouckaert J, Duquet W, et al. De barometer van de fysieke fitheid van de Vlaamse jeugd 1997. De resultaten. *Sport (Blosso Brussel)* 1998;4:16–22.
- 59 Beunen G, Borms J, Vrijens J, et al. *Fysieke fitheid en sportbeoefening van de Vlaamse jeugd. Volume 1: Fysieke fitheid van de jeugd van 6 tot 18 jaar*. Brussels: Blosso, 1991.
- 60 Poortmans J, Vlaeminck M, Collin M, et al. Estimation indirecte de la puissance aérobie maximale d'une population Bruxelloise masculine et féminine âgée de 6 à 23 ans. Comparaison avec une technique directe de la mesure de la consommation maximale d'oxygène. *J Physiol (Paris)* 1986;81:195–201.
- 61 Cazorla G, Gouthon P, Arémou, et al. *État de développement des capacités physiques des jeunes béninois 8-18 ans et plus. Rapport pour le Ministère Béninois de la Jeunesse et Sports et pour le Ministère Français de la Coopération*. Paris, France: Ministère Béninois de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1987.
- 62 Ribeiro RR, Santos KD, Carvalho WCG, et al. Aerobic fitness and biological and sociodemographic indicators in female school children. *Rev Bras Cineantropom Desempenho Hum* 2013;15:448–57.
- 63 Hobold E. *Indicadores de aptidão física relacionada à saúde de crianças e adolescentes do município de Marechal Cândido Rondon* [dissertation]. Paraná, Brasil: Universidade Federal de Santa Catarina, 2003.
- 64 Pieta S. *Estudio de la aptitud física de una muestra de la población escolar del estado de Paraná mediante la batería Eurofit* [dissertation]. León, Brasil: Universidad de León, 2000.
- 65 Veldhuizen S, Cairney J, Hay J, et al. Relative age effects in fitness testing in a general school sample: how relative are they? *J Sports Sci* 2014;33:109–15.
- 66 Voss C, Sandercock G, Higgins JW, et al. A cross-cultural comparison of body composition, physical fitness, and physical activity between regional samples of Canadian and English children and adolescents. *Can J Public Health* 2014;104:e245–50.
- 67 Leone M, Kalinova E, Comtois AS. *Global motor skill assessment from the UQAC-UQAM test battery: Canadian normative values by age and gender*. Québec: Université du Québec à Chicoutimi, 2011.
- 68 Reed KE, Warburton DER, Whitney CL, et al. Secular changes in shuttle-run performance: a 23-year retrospective comparison of 9- to 11-year-old children. *Pediatr Exerc Sci* 2006;18:364–73.
- 69 Massicotte D. *Partial curl-ups, push-ups and multistage 20 meter shuttle run, national norms for 6 to 17 year-olds. Final report submitted to Canadian Association for Health, Physical Education and Recreation (CAHPER) and Fitness and Amateur Sport Canada*. Montréal: University of Quebec at Montréal, 1990.
- 70 Garber MD, Sajuria M, Lobelo F. Geographical variation in health-related physical fitness and body composition among Chilean 8th graders: a nationally representative cross-sectional study. *PLoS ONE* 2014;9:e108053.
- 71 Kain J, Uauy R, Albala, et al. School-based obesity prevention in Chilean primary school children: methodology and evaluation of a controlled study. *Int J Obes* 2004;28:483–93.
- 72 Wong TW, Yu TS, Wang XR, et al. Predicted maximal oxygen uptake in normal Hong Kong Chinese schoolchildren and those with respiratory diseases. *Pediatr Pulmonol* 2001;31:126–32.
- 73 Wang P-G, Gong G, Wang SQ, et al. Relationship of body fat and cardiorespiratory fitness with cardiovascular risk in Chinese children. *PLoS ONE* 2011;6:e27896.
- 74 Cubides RC, Alarcón LGA, Galvi ARG. Diferencias en la actividad física y la condición física entre los escolares de secundaria de dos programas curriculares oficiales de Bogotá, Colombia. *Nutrición Hospitalaria* 2015;32:2228–34.
- 75 Ortega FJA. Estudio transversal de las cualidades funcionales de los escolares bogotanos: Valores de potencia aeróbica, potencia muscular, velocidad de desplazamiento y velocidad de reacción, de los siete a los dieciocho años. *Med Deporte Trab* 2013;32:1151–70.
- 76 Mojica GT, Poveda JG, Pinilla MI, et al. Sobre peso, inactividad física y baja condición física en un colegio de Bogotá, Colombia. *Arch Latinoam Nutr* 2008;58:265–73.
- 77 Cazorla G, Dudal J, Dieu JL. *État de développement des capacités physiques des jeunes ivoiriens 7-18 ans et plus. Rapport pour le Ministère Ivoirien de la Jeunesse et Sports et pour le Ministère Français de la Coopération*. Paris, France: Ministère Ivoirien de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1985.
- 78 Shahin A. Ying and yang of body composition assessment. *Chin J Integr Med* 2011;17:675–9.
- 79 Tinaz C, Emiroğlu O. Assessment of physical fitness levels, gender and age differences of rural and urban elementary school children. *Türkiye Klinikleri J Med Sci* 2010;30:1–7.
- 80 Gajda V. *The applications of the UNIFIT test battery at basic schools (age groups 7-14): summary report*. Ostrava: Pdf OU, 1994.
- 81 Nielsen GA, Andersen LB. The association between high blood pressure, physical fitness, and body mass index in adolescents. *Prev Med* 2003;36:229–34.
- 82 Cazorla G, Dudal J, Lefrançois L. *État de développement des capacités physiques des jeunes djiboutiens 11-18 ans et plus. Rapport pour le Ministère Djiboutien de la Jeunesse et Sports et pour le Ministère Français de la Coopération*. Paris, France: Ministère Djiboutien de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1986.
- 83 Jürimäe T, Volbekiene V, Jürimäe J. Changes in Eurofit test performance of Estonian and Lithuanian children and adolescents (1992–2002). In: Tomkinson GR, Olds TS, eds. *Pediatric fitness: secular trends and geographic variability*. Basel: Karger, 2007:129–42.
- 84 Jürimäe T, Saar M. Self-perceived and actual indicators of motor abilities in children and adolescents. *Percept Mot Skills* 2003;97:862–6.
- 85 Jürimäe T, Volbekiene V. Eurofit test results in Estonian and Lithuanian 11 to 17-year-old children: a comparative study. *Eur J Phys Educ* 1998;3:178–84.
- 86 Raudsepp L, Jürimäe T. Relationships of physical activity and somatic characteristics with physical fitness and motor skill in prepubertal girls. *Am J Hum Biol* 1997;9:513–21.
- 87 Raudsepp L, Jürimäe T. Relationships between somatic variables, physical activity, fitness and fundamental motor skills in prepubertal boys. *Biol Sport* 1996;13:279–89.
- 88 Kull M, Jürimäe T. Using the Eurofit test battery in Estonian 16-18 years old adolescents. *Acta Commentationes Univ Tartuensis* 1994;96:749–52.
- 89 Palomäki S, Heikinaro-Johansson P, Huotari P. Cardiorespiratory performance and physical activity in normal weight and overweight Finnish adolescents from 2003 to 2010. *J Sports Sci* 2015;33:588–96.
- 90 Viljanen T, Taimela S, Kujala U. Kouluikäisten fyysinen aktiivisuus, kestävyyskunto ja ponnistuskorkeus. *Liikunta Tiede* 2000;37:23–6.
- 91 Baquet G, Twisk JWR, Kemper HCG, et al. Longitudinal follow-up of fitness during childhood: interaction with physical activity. *Am J Hum Biol* 2006;18:51–8.
- 92 Baquet G, Berthoin S, Gerbeaux M, et al. High-intensity aerobic training during a 10 week one-hour physical education cycle: effect on physical fitness of adolescents aged 11 to 16. *Int J Sports Exerc Med* 2001;22:295–300.
- 93 Baquet G, Berthoin S, Gerbeaux M, et al. Assessment of the maximal aerobic speed with the incremental running field tests in children. *Biol Sport* 1999;16:23–30.
- 94 Cazorla G, Portes A, James F. *Opération Martinique-Eval. Centre d'Evaluation Sport Santé, Fort de France (Martinique). Rapport pour l'Inspection d'Académie de la Martinique*. Fort de France, Martinique: L'Education d'Académie de la Martinique, 1997.
- 95 Cazorla G. *Batterie France-Eval: Mesures, épreuves et barèmes: Évaluation des qualités physiques des jeunes Français d'âge scolaire: 7-11 ans. Rapport pour le Secrétariat d'Etat auprès du Premier Ministre Chargé de la Jeunesse et de Sports*. Paris: Ministère de la Jeunesse et de Sports, 1987.
- 96 Brunet J, Van Praagh E. *Batterie expérimentale de tests moteurs Eurofit: Rapport d'Activité de la Région Auvergne—1984-1985*. Clermont-Ferrand, France: Université de Clermont-Ferrand, 1985.
- 97 Tambalis KD, Panagiotakos DB, Psarra G, et al. Physical fitness normative values for 6-18-year-old Greek boys and girls, using the empirical distribution and the lambda, mu, and sigma statistical method. *Eur J Sport Sci* 2015:1–11.
- 98 Tambalis K, Panagiotakos D, Sidossis L. Greek children living in rural areas are heavier but fitter compared to their urban counterparts: a comparative, time-series (1997–2008) analysis. *J Rural Health* 2011;27:270–7.
- 99 Tokmakidis SP, Kasambalis A, Christodoulos AD. Fitness levels of Greek primary schoolchildren in relationship to overweight and obesity. *Eur J Pediatr* 2006;165:867–74.
- 100 Manios Y, Yiannakouris N, Papoutsakis C. Behavioral and physiological indices related to BMI in a cohort of primary schoolchildren in Greece. *Am J Hum Biol* 2004;16:639–47.

- 101 Georgiadis G. *Evaluation of physical fitness of Greek youth aged 6-18 years* [dissertation]. Athens, Greece: University of Athens, 1993.
- 102 Welk GJ, Saint-Maurice PF, Csányi T. Health-related physical fitness in Hungarian youth: age, sex, and regional profiles. *Res Q Exerc Sport* 2015;86:S45–57.
- 103 Barabás A. Measurement of aerobic power by field tests. In: Coudert J, Van Praagh E, eds. *Pediatric work physiology: children and exercise XVI*. Paris: Masson, 1992:39–41.
- 104 Gunnarsson HG, Sigriksson R. *Eurofit: physical fitness of Icelandic pupils at age of 6-15 years old*. Reykjavik, Iceland, 1999.
- 105 Grassi GP, Turci M, Sforza C. Aerobic fitness and somatic growth in adolescents: a cross sectional investigation in a high school context. *J Sports Med Phys Fitness* 2006;46:412–18.
- 106 Cilia G, Bazzano C, Bellucci M, et al. I risultati dei test Eurofit nella scuola Matteucci di Roma. *Alcmeone* 1998;2:16–20.
- 107 Cilia G, Bellucci M, Bazzano C, et al. Eurofit 1997: Banche dati per la scuola. *Alcmeone* 1997;3:13–32.
- 108 Cilia G, Bellucci M, Riva M. *Eurofit 1995*. Roma: Istituto Superiore Statale di Educazione Fisica, 1996.
- 109 Cilia G, Bellucci M. *Eurofit: tests Europei di attitudine fisica*. Roma: Istituto Superiore Statale di Educazione Fisica, 1993.
- 110 Council of Europe. *Evaluation de l'aptitude physique: Eurofit batterie expérimentale*. Rome: Council of Europe, 1986.
- 111 Ministry of Education, Culture, Sports, Science and Technology. *Report book on the survey of physical fitness and athletic ability*. Tokyo: Ministry of Education, Culture, Sports, Science and Technology, 1999–2015.
- 112 Sauka M, Priedite IS, Artjuhova L, et al. Physical fitness in northern European youth: reference values from the Latvian physical health in youth study. *Scand J Public Health* 2011;39:35–43.
- 113 Cazorla G, Dudal J, Garinet P. État de développement des capacités physiques des jeunes mauriciens 7-18 ans et plus. Rapport pour le Ministère Mauricien de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1986.
- 114 Galaviz KI, Tremblay MS, Colley R, et al. Associations between physical activity, cardiorespiratory fitness, and obesity in Mexican children. *Salud Pública Méx* 2012;54:463–9.
- 115 Dadouchi MF, Boudiab K, Yahia J, et al. *De la détermination du profil athlétique marocain au contenu et à l'évaluation*. Ministère de l'Éducation nationale et de la jeunesse. École normale supérieure de Marrakech, 2003.
- 116 Brouwer SI, Stol RP, Liem ET, et al. The role of fitness in the association between fatness and cardiometabolic risk from children to adolescence. *Pediatr Diabetes* 2013;14:57–65.
- 117 Slinger J, van Breda E, Kuipers H. Aerobic fitness data for Dutch adolescents (2002-2005). *Pediatr Exerc Sci* 2009;21:10–18.
- 118 van Mechelen W, van Lier WH, Hlobil H, et al. *Eurofit: Handleiding met referentieschalen voor 12- tot en met 16-jarige jongens en meisjes in Nederland*. Haarlem: Uitgeverij de Vrieseborch, 1991.
- 119 Haugen T, Haugaard R, Seiler S. Normative data of BMI and physical fitness in a Norwegian sample of early adolescents. *Scand J Public Health* 2014;42:67–73.
- 120 Cossio-Bolaños MA, Arruda M. Propuesta de valores normativos para la evaluación de la aptitud física en niños de 6 a 12 años de Arequipa, Perú. *Rev Med Herediana* 2009;20:206–12.
- 121 Gonzalez-Suarez CB, Grimmer-Sommers K. The association of physical activity and physical fitness with pre-adolescent obesity: an observational study in Metro Manila, Philippines. *J Phy Act Health* 2011;8:804–10.
- 122 Bronikowski M, Bronikowska M. Salutogenesis as a framework for improving health resources of adolescent boys. *Scand J Public Health* 2009;37:525–31.
- 123 Pilicz S, Przeweda R, Dobosz J, et al. *Punkcja sprawności fizycznej młodzieży Polskiej wg międzynarodowego testu sprawności fizycznej: Kryteria pomiaru wydolności organizmu testem Cooper*. Warszawa: Akademia Wychowania Fizycznego Józefa Piłsudskiego, 2003.
- 124 Maciaszek J, Osinski W. Poziom sprawności fizycznej u chłopców i dziewcząt Poznańskich w wieku 10-14 lat. *Rocz Naukowe AWF Poznaniu* 2001;50:3–17.
- 125 Mleczo E, Ozimek M. *Rozwój somatyczny i motoryczny młodzieży Krakowskiej między 15 a 19 rokiem życia z uwzględnieniem czynników środowiskowych*. Kraków: Akademia Wychowania Fizycznego, 2000.
- 126 Santos R, Mota J, Santos DA, et al. Physical fitness percentiles for Portuguese children and adolescents aged 10-18 years. *J Sports Sci* 2014;32:1510–81.
- 127 Coelho-Silva MJ, Ronque ERV, Cyrino ES, et al. Nutritional status, biological maturation and cardiorespiratory fitness in Azorean youth aged 11-15 years. *BMC Public Health* 2013;13:495.
- 128 Rodrigues LP, Leitao R, Lopes VP. Physical fitness predicts adiposity longitudinal changes over childhood and adolescence. *J Sci Med Sport* 2013;16:118–23.
- 129 Marta CC, Marinho DA, Barbosa TM. Physical fitness differences between prepubescent boys and girls. *J Strength Cond Res* 2012;26:1756–66.
- 130 Mota J, Guerra S, Leandro C. Association of maturation, sex, and body fat in cardiorespiratory fitness. *Am J Hum Biol* 2002;14:707–12.
- 131 Kim H-B, Stebbins CL, Chai J-H, et al. Taekwondo training and fitness in female adolescents. *J Sports Sci* 2011;29:133–8.
- 132 Cazorla G, Dudal J, Faye J. État de développement des capacités physiques des jeunes sénégalais 7-18 ans et plus. Rapport pour le Ministère Sénégalais de la Jeunesse et Sports et pour le Ministère Français de la Coopération. Paris, France: Ministère Sénégalais de la Jeunesse et Sports et pour le Ministère Français de la Coopération, 1988.
- 133 Bovet P, Auguste R, Burdette H. Strong inverse association between physical fitness and overweight in adolescents: a large school-based survey. *Int J Behav Nutr Phys Act* 2007;4:24.
- 134 Cazorla G, Rousseau G, Dudal J, et al. *Évaluation des capacités motrices de l'enfant, de l'adolescent et du jeune sénégalais : 7-18 ans et plus. Rapport pour le Ministère Seychellois de l'Éducation et pour le Ministère Français de la Coopération*. Paris, France: Ministère Seychellois de l'Éducation et Ministère Français de la Coopération, 1990.
- 135 Krška P, Sedláček J, Hubinák A, et al. General motor fitness and somatic parameters comparison between former population and present primary school girls in Ruzomberok. In: Proceedings of the 9th International Conference in Physical Education, Sport and Physical Therapy, Iași, 13–14 November. Iași, Romania: Faculty of Sports and Physical Education, Alexandru Ioan Cuza University, 2015.
- 136 Krška P, Sedláček J, Hubinák A, et al. Physical development and general motor performance of present primary school boys population in Ruzomberok. In: Zvonář M, Sajdlíková Z, eds. Proceedings of the 10th International Conference of Kinanthropology, Brno, Czech Republic, 18–20 November. Brno: Faculty of Sport Studies, Masaryk University, 2015:289–96.
- 137 Kyselovicová O. Programy aerobiku z aspektu rozvoja telesnej zdatnosti dievcat na gymnáziu. *Acta Facultatis Educationis Physicae Univ Comenianae* 2000;41:50–61.
- 138 Kasa J, Majherová M. Physical and motor development of children by Eurofit. *Stud Psychol (Bratisl)* 1997;39:270–4.
- 139 Moravec R. Eurofit—physique and motor fitness of the Slovak school youth. In: Moravec R, Kampmiller T, Sedláček J, eds. *Eurofit—physique and motor fitness of the Slovak school youth*. Bratislava: Slovak Scientific Society for Physical Education and Sports, 1996:9–22.
- 140 Belej M, Junger J, Mikus M. Výkonnosť žiakov prijatých na stredné školy vo východoslovenskom regióne zisťovaná systémom EUROFIT. *Tel Vych Sport* 1995;5:12–15.
- 141 Pienaar C, Coetzee B, Monyeki AM. The use of anthropometric measurements and the influence of demographic factors on the prediction of $\dot{V}O_{2max}$ in a cohort of adolescents: the PAHL study. *Ann Hum Biol* 2015;42:135–43.
- 142 Du Toit D, Pienaar AE, Truter L. Relationship between physical fitness and academic performance in South African children. *S Afr J Res Sport PH* 2011;33:23–35.
- 143 Pienaar AE, Viljoen A. Physical and motor ability, anthropometrical and growth characteristics of boys in the northwest province of South Africa: a sport talent perspective. *S Afr J Res Sport Phys Educ Recreation* 2010;32:71–93.
- 144 Du Preez SM. *The effect of physical activity on the body composition and health related fitness of 9 to 13-year-old boys* [thesis]. Potchefstroom, South Africa: North-West University, 2008.
- 145 Stadler MC. *The influence of a physical activity intervention program (PAI) on the physical fitness levels, body composition and health risk behaviour of 9 to 13 year old girls* [dissertation]. Potchefstroom, South Africa: North-West University, 2007.
- 146 Du Toit L, Venter RE, Potgieter JR. The relationship between cardiorespiratory fitness, body composition and physical self-perception of adolescent girls. *J Hum Mov Stud* 2005;48:353–64.
- 147 Van Gent M, Malan DD, Pienaar AE. A comparison of the anthropometric, physical and motor growth characteristics of 12-15 year old girls in the North West province with Australian girls. *Afr J Phys Health, Educ Rec Dance* 2002;8:309–20.
- 148 Du Randt R. *The 1996 South African sport talent identification project: report prepared for the Sports Information and Science Agency*. Pretoria: South African Sports Commission, 1996.
- 149 Gullías-González R, Martínez-Vizcaino V, García-Prieto JC, et al. Excess of weight, but not underweight, is associated with poor physical fitness in children and adolescents from Castilla-La. *Eur J Pediatr* 2014;173:727–35.
- 150 Torrijos-Niño C, Martínez-Vizcaino V, Pardo-Guijarro MJ, et al. Physical fitness, obesity, and academic achievement in schoolchildren. *J Pediatr* 2014;165:104–9.
- 151 Chillón P, Ortega FB, Ferrando JA, et al. Physical fitness in rural and urban children and adolescents from Spain. *J Sci Med Sport* 2011;14:417–23.
- 152 Castro-Pinero J, González-Montesinos JL, Mora J, et al. Percentile values for muscular strength field tests in children aged 6 to 17 years: influence of weight status. *J Strength Cond Res* 2009;23:2295–310.
- 153 Ortega FB, Ruiz JR, Castillo MJ, et al. Low level of physical fitness in Spanish adolescents. Relevance for future cardiovascular health (AVENA Study). *Rev Esp Cardiol* 2005;58:898–909.
- 154 García Baena J. *La condición física en la educación secundaria. Trabajo de investigación* [thesis]. Madrid, Spain: Universidad Nacional de Educación a Distancia, 1999.
- 155 Prat JA, Casamort J, Balagué N, et al. *Eurofit: La batería Eurofit en Catalunya*. Barcelona: Secretaria General de l'Esport, 1998.

- 156 Tercedor P, Delgado-Fernandez M. Condición física relacionada con la salud en escolares de 10 años de edad de Granada. *Proceedings of the II congreso internacional sobre la enseñanza de la educación física y el deporte escolar (second international congress about teaching physical education and school sport)*. 1998.
- 157 Sainz RM. *La batería Eurofit en Euskadi*. Vitoria-Gasteiz: Instituto Vasco de Educación Física, 1996.
- 158 Ureña F. *Valoración y baremación de la aptitud física en el alumnado de segundo ciclo de educación secundaria obligatoria de la comunidad autónoma de Murcia. Su utilización según los postulados de la reforma* [dissertation]. Murcia, Spain: Universidad de Murcia, 1996.
- 159 Brito Ojeda EM, Navarro Valdivielso M, García Afonso D, et al. *La condición física en la población escolar de gran Canaria (10-19 años)*. Las Palmas de Gran Canaria, Spain: Excmo. Cabildo Insular de Gran Canaria, 1995.
- 160 Sainz RM. *Aptitudes psíquicas y físicas: Estudio de la aptitud física de los adolescentes de la provincia de Vizcaya y su relación con la personalidad* [dissertation]. Bilbao, Spain: Universidad de Deusto, 1992.
- 161 Rivas FJ. *Valoración de la evolución antropométrica y de las características motrices en la población escolar de un centro de E.G.B. mediante un estudio transversal*. In: *II Congreso Galego da Educación Física e o Deporte*. La Coruña: Escola Galega do Deporte de la Xunta de Galicia, 1987:17–28.
- 162 Lieweld J, Vrijens J, Bouckaert J. Ethnic differences in body structure and physical fitness in Surinamese boys aged 14 years. In: Claessens AL, Lefevre J, Vanden Eynde B, eds. *World-wide variation in physical fitness*. Leuven: Katholieke Universiteit Leuven, 1993:87–92.
- 163 Schmid M, Romann M, Kriemler S, et al. Wie kann die Fitness von Schülern gemessen werden? *Sportmedizin Sporttraumatologie* 2007;55:52–61.
- 164 Cauderay M, Narring F, Michaud P-A. A cross-sectional survey assessing physical fitness of 9- to 19-year-old girls and boys in Switzerland. *Pediatr Exerc Sci* 2000;12:398–412.
- 165 Aandstad A, Berntsen S, Hageberg R, et al. A comparison of estimated maximal oxygen uptake in 9 and 10 year old schoolchildren in Tanzania and Norway. *Br J Sports Med* 2006;40:278–92.
- 166 Çalis M, Ergen E, Turnagöl H, et al. Beden egitimi derslerinin bir öğretim yılı boyunca 15-16 yaş grubu öğrencilerin üzerinde fizyolojik etkilerinin Eurofit test bataryası ile izlenmesi. In: *3 Ulusal Spor Bilimleri Kongresi*. Ankara, Turkey: Hacettepe Üniversitesi, 1992:367–72.
- 167 Richards J, Foster C, Townsend N. Physical fitness and mental health impact of a sport-for-development intervention in a post-conflict setting: randomised controlled trial nested within an observational study of adolescents in Gulu, Uganda. *BMC Public Health* 2014;14:e19.
- 168 Sandercock G, Ogunleye A, Voss C. Six-year changes in body mass index and cardiorespiratory fitness of English schoolchildren from an affluent area. *Int J Obes* 2015;39:1504–7.
- 169 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;30:679–87.
- 170 Boddy L, Fairclough SJ, Atkinson G, et al. Changes in cardiorespiratory fitness in 9- to 10.9-year-old children: Sports Linx 1998-2010. *Med Sci Sports Exerc* 2012;44:481–6.
- 171 Sandercock G, Voss C, McConnell D. Ten year secular declines in the cardiorespiratory fitness of affluent English children are largely independent of changes in body mass index. *Arch Dis Child* 2010;95:46–7.
- 172 Liverpool City Council. *Liverpool Sports Linx Project 01-03: Report on the health and fitness of Liverpool primary and secondary school children*. Liverpool, UK: Liverpool City Council, 2003.
- 173 Boreham C, Twisk J, Murray L, et al. Fitness, fatness, and coronary heart disease risk in adolescents: the Northern Ireland Young Hearts Project. *Med Sci Sports Exerc* 2001;33:270–4.
- 174 Watkins DC. *Ten year trends (1990-2000) in biological and behavioural risk factors for coronary heart disease in northern Irish adolescents* [thesis]. Belfast, UK: The Queen's University of Belfast, 2001.
- 175 Twisk JMR, Boreham C, Cran G, et al. Clustering of biological risk factors for cardiovascular disease and the longitudinal relationship with lifestyle of an adolescent population: the Northern Ireland Young Hearts Project. *J Cardiovasc Risk* 1999;6:355–62.
- 176 Mahoney C. 20-MST and PWC₁₇₀ validity in non-Caucasian children in the UK. *Br J Sports Med* 1992;26:45–7.
- 177 Mahoney CA, Boreham CAG. Eurofit in Belfast primary schools. *Scot J Phys Educ* 1991;19:1–4.
- 178 Riddoch C, Savage JM, Murphy N, et al. Long term health implications of fitness and physical activity patterns. *Arch Dis Child* 1991;66:1426–33.
- 179 Boreham CAG, Paliczka VJ, Nichols AK. Fitness testing of Belfast schoolchildren. In: *5th European research seminar on testing physical fitness*. Strasbourg: Council of Europe, 1987:52–7.
- 180 Lewitt MS, Baker JS, Mooney GP, et al. Pubertal stage and measures of adiposity in British schoolchildren. *Ann Hum Biol* 2012;39:440–7.
- 181 Ranson R, Stratton G, Taylor S. Digit ratio (2D:4D) and physical fitness (Eurofit test battery) in school children. *Early Hum Dev* 2015;91:327–31.
- 182 Liu W, Zillfro TD, Nichols RA. Tracking of health-related physical fitness for middle school boys and girls. *Pediatr Exerc Sci* 2012;24:549–62.
- 183 Welk GJ, Maduro PFDS-M, Laurson KR, et al. Field evaluation of the new FITNESSGRAM criterion-referenced standards. *Am J Prev Med* 2011;41:S131–42.
- 184 Beets MW, Pitetti KH, Cardinal BJ. Progressive aerobic cardiovascular endurance run and body mass index among an ethnically diverse sample of 10–15 year olds. *Res Q Exerc Sport* 2005;76:389–97.
- 185 Beets MW, Pitetti KH. A comparison of shuttle-run performance between mid-western youth and their national and international counterparts. *Pediatr Exerc Sci* 2004;16:94–112.
- 186 Welk GJ, Schaben JA, Shelley M. Physical activity and physical fitness in children schooled at home and children attending public schools. *Pediatr Exerc Sci* 2004;16:310–23.
- 187 Lloyd LK, Bishop PA, Walker JL, et al. The influence of body size and composition on FITNESSGRAM® test performance and the adjustment of FITNESSGRAM® test scores for skinfold thickness in youth. *Meas Phys Educ Exerc Sci* 2003;7:205–26.
- 188 Chun DM, Corbin CB, Pangrazi RP. Validation of criterion-referenced standards for the mile run and progressive aerobic cardiovascular endurance tests. *Res Q Exerc Sport* 2000;71:125–34.
- 189 Wolford N. *The difference in physical fitness levels of fifth graders according to socioeconomic groups and genders* [dissertation]. Ann Arbor, MI: University Microfilms International, 1998.
- 190 Mahar MT, Rowe DA, Parker CR, et al. Criterion-referenced and norm-referenced agreement between the mile run/walk and PACER. *Meas Phys Educ Exerc Sci* 1997;1:245–58.
- 191 <http://data.worldbank.org/about/country-and-lending-groups> (accessed 24 Dec 2015).
- 192 Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. *Arch Dis Child* 1995;73:25–9.
- 193 Cole TJ, Bellizzi MC, Flegal KM, et al. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* 2000;320:1240–3.
- 194 Eisenmann JC. Waist circumference percentiles for 7- to 15-year-old Australian children. *Acta Paediatr* 2005;94:1182–5.
- 195 Fernandez JR, Redden DT, Pietrobelli A, et al. Waist circumference percentiles in nationally representative samples of African-American, European-American, and Mexican-American children and adolescents. *J Pediatr* 2004;145:439–44.
- 196 McCarthy HD, Jarret KV, Crawley HF. The development of waist circumference percentiles in British children aged 5.0-16.9 y. *Eur J Clin Nutr* 2001;55:902–7.
- 197 Shields M. Overweight and obesity among children and youth. *Health Rep* 2006;17:27–42.
- 198 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* 2009;21:1–14.
- 199 Jackson LV, Thalange NKS, Cole TJ. Blood pressure centiles for Great Britain. *Arch Dis Child* 2007;92:298–303.
- 200 Jolliffe CJ, Janssen I. Development of age-specific adolescent metabolic syndrome criteria that are linked to the Adult Treatment Panel III and International Diabetes Federation criteria. *J Am Coll Cardiol* 2007;49:891–8.
- 201 Castro-Piñero J, González-Montesinos JL, Keating XD, et al. Percentile values for running sprint field tests in children ages 6–17 years: influence of weight status. *Res Q Exerc Sport* 2010;81:143–51.
- 202 De Miguel-Etayo P, Gracie-Marco L, Ortega FB, et al. Physical fitness reference standards in European children: the IDEFICS study. *Int J Obes* 2014;38(2):S57–66.
- 203 Eisenmann JC, Laurson KR, Welk GJ. Aerobic fitness percentiles for U.S. adolescents. *Am J Prev Med* 2011;41:S106–10.
- 204 Armstrong N, Barker AR. Endurance training and elite young athletes. *Med Sport Sci* 2011;56:59–83.
- 205 Ruiz JR, Ortega FB, Rizzo NS, et al. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. *Pediatr Res* 2007;61:350–5.
- 206 Lobelo F, Pate RR, Dowda M, et al. Validity of cardiorespiratory fitness criterion-referenced standards for adolescents. *Med Sci Sports Exerc* 2009;41:1222–9.
- 207 Adegboye ARA, Anderssen SA, Froberg K, et al. Recommended aerobic fitness level for metabolic health in children and adolescents: a study of diagnostic accuracy. *Br J Sports Med* 2011;45:722–8.
- 208 Boddy LM, Thomas NE, Fairclough SJ, et al. ROC generated thresholds for field-assessed aerobic fitness related to body size and cardiometabolic risk in schoolchildren. *PLoS ONE* 2012;7:e45755.
- 209 Mesa JL, Ruiz JR, Ortega FB, et al. Aerobic fitness in relation to blood lipids and fasting glycaemia in adolescents: influence of weight status. *Nutr Metab Cardiovasc Dis* 2006;16:285–93.
- 210 Moreira C, Santos R, Ruiz JR, et al. Comparison of different $\dot{V}O_{2max}$ equations in the ability to discriminate the metabolic risk in Portuguese adolescents. *J Sci Med Sport* 2011;14:79–84.

Original article

- 211 Ruiz JR, Huybrechts I, Cuenca-Garcia M, *et al.* Cardiorespiratory fitness and ideal cardiovascular health in European adolescents. *Heart* 2015;101:766–73.
- 212 Tomkinson G. Aerobic fitness thresholds for cardiometabolic health in children and adolescents. *Br J Sports Med* 2011;45:686–7.
- 213 Barnett A, Chan LYS, Bruce IC. A preliminary study of the 20-m multistage shuttle run as a predictor of peak $\dot{V}O_2$ in Hong Kong Chinese students. *Pediatr Exerc Sci* 1993;5:42–50.
- 214 Stickland MK, Petersen SR, Bouffard M. Prediction of maximal aerobic power from the 20-m multi-stage shuttle run test. *Can J Appl Physiol* 2003;28:272–82.
- 215 Flouris AD, Metsios GS, Koutedakis Y. Enhancing the efficacy of the 20 m multistage shuttle run test. *Br J Sports Med* 2005;39:166–70.
- 216 Matsuzaka A, Takahashi Y, Yamazoe M, *et al.* Validity of the multistage 20-m shuttle-run test for Japanese children, adolescents, and adults. *Pediatr Exerc Sci* 2004;16:113–25.
- 217 Mahar MT, Welk GJ, Rowe DA, *et al.* Development and validation of a regression model to estimate $\dot{V}O_{2peak}$ from PACER 20-m shuttle run performance. *J Phys Act Health* 2006;3:534–46.
- 218 Ruiz JR, Ramirez-Lechuga J, Ortega FB, *et al.* Artificial neural network-based equation for estimating $\dot{V}O_{2max}$ from the 20 m shuttle run test in adolescents. *Artif Intell Med* 2008;44:233–45.
- 219 Oliveira J. *Validação directa do teste de vaivém em 20 metros, de Luc- Léger, em adolescentes portugueses.* Cruz-Quebrada: FMH-UTL, 1998.
- 220 Pitetti KH, Fernhall B, Figoni S. Comparing two regression formulas that predict $\dot{V}O_{2peak}$ using the 20-m shuttle run for children and adolescents. *Pediatr Exerc Sci* 2002;14:125–34.
- 221 McVeigh SK, Payne AC, Scott S. The reliability and validity of the 20-meter shuttle test as a predictor of peak oxygen uptake in Edinburgh school children, age 13 to 14 years. *Pediatr Exerc Sci* 1995;7:69–79.
- 222 Moreno LA, González-Gross M, Kersting M, *et al.* Assessing, understanding and modifying nutritional status, eating habits and physical activity in European adolescents: the HELENA (Healthy lifestyle in Europe by nutrition in adolescence) study. *Public Health Nutr* 2008;11:288–99.
- 223 Andersen KL, Seliger V, Rutenfranz J, *et al.* Physical performance capacity of children in Norway. Part IV. The rate of growth in maximal aerobic power and the influence of improved physical education of children in a rural community—population parameters in a rural community. *Eur J Appl Physiol Occup Physiol* 1976;35:49–58.
- 224 Armstrong N, Welsman JR. Assessment and interpretation of aerobic fitness in children and adolescents. *Exerc Sport Sci Rev* 1994;22:435–76.
- 225 Krahenbuhl GS, Skinner JS, Kohrt WM. Developmental aspects of maximal aerobic power in children. *Exerc Sport Sci Rev* 1985;13:503–38.
- 226 Rowland TW. Evolution of maximal oxygen uptake in children. *Med Sport Sci* 2007;50:200–9.
- 227 Rowland TW. *Children's exercise physiology.* 2nd edn. Champaign, IL: Human Kinetics, 2005:89–112.
- 228 Bar-Or O, Rowland TW. Physiologic and perceptual responses to exercise in the healthy child. In: Bar-Or O, Rowland TW, eds. *Pediatric exercise medicine: from physiologic principles to health care application.* Champaign, IL: Human Kinetics, 2004:3–59.
- 229 Tomkinson G, Macfarlane DJ, Noi S, *et al.* Temporal changes in long-distance running performance of Asian children between 1964 and 2009. *Sports Med* 2012;42:267–79.
- 230 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* 2010;13:584–8.
- 231 Tomkinson G, Olds T, Borms J. Who are the Eurofittest? *Med Sport Sci* 2007;50:104–28.
- 232 Castro-Piñero J, Artero EG, España-Romero V, *et al.* Criterion-related validity of field-based fitness tests in youth: a systematic review. *Br J Sports Med* 2010;44:934–43.
- 233 Mayorga-Vega D, Aguilar-Soto P, Vidiána J. Criterion-related validity of the 20-M Shuttle run test for estimating cardiorespiratory fitness: a meta-analysis. *J Sports Sci Med* 2015;14:536–47.
- 234 Lussier L, Buskirk ER. Effects of an endurance training regimen of assessment of work capacity in prepubertal children. *Ann N Y Acad Sci* 1977;301:734–47.
- 235 Krahenbuhl GS, Pangrazi RP, Chomokos EA. Aerobic responses of young boys to submaximal running. *Res Q* 1979;50:413–21.
- 236 Péronnet F, Thibault G. Mathematical analysis of running performance and world records. *J Appl Physiol* 1989;67:453–65.
- 237 Sjödin B. The relationship among running economy, aerobic power, and onset of blood lactate accumulation in young boys (11–15 years). In: Komi PV, ed. *Exercise and sport biology.* Champaign, IL: Human Kinetics, 1982:57–60.
- 238 Mayers N, Gutin B. Physiological characteristics of elite prepubertal cross-country runners. *Med Sci Sports* 1979;11:172–6.
- 239 Cureton KJ, Boileau RA, Lohman TG, *et al.* Determinants of distance running performances in children: analysis of a path model. *Res Q* 1977;42:270–9.
- 240 Voss C, Sandercock G. Does the twenty meter shuttle-run test elicit maximal effort in 11- to 16-year-olds? *Pediatr Exerc Sci* 2009;21:55–62.
- 241 Longmuir PE, Colley RC, Wherley VA, *et al.* Canadian Society for Exercise Physiology position stand: benefit and risk for promoting childhood physical activity. *Appl Physiol Nutr Metab* 2014;39:1271–9.
- 242 Meusel D, Ruiz JR, Ortega FB, *et al.* Assessing levels of physical activity in the European population—the ALPHA project. *Selección* 2007;16:9–12.
- 243 Longmuir PE, Boyer C, Lloyd M, *et al.* The Canadian Assessment of Physical Literacy: methods for children in grades 4 to 6 (8 to 12 years). *BMC Public Health* 2015;15:767.
- 244 Council of Europe. *Eurofit: handbook for the EUROFIT tests of physical fitness.* Rome: Secretariat of the Committee for the Development of Sport within the Council of Europe, 1998.
- 245 Welk GJ, Going SB, Morrow Jr JR, *et al.* Development of new criterion-references fitness standards in the FITNESSGRAM® program: rationale and conceptual overview. *Am J Prev Med* 2011;41:S63–7.
- 246 Ortega FB, Cadenas-Sánchez C, Sánchez-Delgado G, *et al.* Systematic review and proposal of a field-based physical fitness-test battery in preschool children: the PREFIT battery. *Sports Med* 2015;45:533–55.
- 247 Cadenas-Sánchez C, Alcántara-Moral F, Sánchez-Delgado G, *et al.* [Assessment of cardiorespiratory fitness in preschool children: adaptation of the 20 metre shuttle run test]. *Nutr Hosp* 2014;30:1333–43.

Supplement 1. Database search strategies.

*the number in brackets represents the number of studies identified.

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R)
<1946 to Present>

Search Strategy:

-
- 1 shuttle run.mp. (415)
 - 2 beep test.mp. (3)
 - 3 physical endurance/ (13383)
 - 4 Physical Fitness/ (19679)
 - 5 multi-stage test.mp. (4)
 - 6 1 or 2 or 3 or 4 or 5 (31546)
 - 7 6 and (child\$ or adolescent).mp. (7253)

Database: Embase Classic+Embase <1947 to 2014>

Search Strategy:

-
- 1 shuttle run.mp. (513)
 - 2 beep test.mp. (4)
 - 3 multi-stage test.mp. (5)
 - 4 endurance/ (14399)
 - 5 aerobic capacity/ (5560)
 - 6 1 or 2 or 3 or 4 or 5 (19750)
 - 7 6 and (child\$ or adolescent).mp. (2430)

Database: PsycINFO <1806 to 2014>

Search Strategy:

-
- 1 shuttle run.mp. (89)
 - 2 beep test.mp. (1)
 - 3 multi-stage test.mp. (4)

- 4 Physical Endurance/ (403)
- 5 Physical Fitness/ (2674)
- 6 Aerobic Exercise/ (960)
- 7 1 or 2 or 3 or 4 or 5 or 6 (3843)
- 8 7 and (child\$ or adolescent).mp. (570)

Database: EBM Reviews - Cochrane Database of Systematic Reviews <2005 to April 2014>

Search Strategy:

-
- 1 shuttle run.mp. [mp=title, abstract, full text, keywords, caption text] (2)
 - 2 physical endurance.mp. [mp=title, abstract, full text, keywords, caption text] (19)
 - 3 physical fitness.mp. [mp=title, abstract, full text, keywords, caption text] (94)
 - 4 aerobic fitness.mp. [mp=title, abstract, full text, keywords, caption text] (14)
 - 5 exercise test.mp. [mp=title, abstract, full text, keywords, caption text] (60)
 - 6 1 or 2 or 3 or 4 or 5 (139)

Database: SPORTDiscus with Full Text (1965-2014)

(shuttle run OR beep test OR multi stage test OR fitness test OR endurance OR aerobic capacity)
AND (child\$ OR adolescence)

Supplement 3. Sex- and age-specific LMS values for 20 m shuttle run (speed [km.h⁻¹] at the last complete stage).

Age (y)	L	M	S
<i>boys</i>			
9	1.004	10.033	0.085
10	1.004	10.128	0.091
11	1.004	10.249	0.097
12	1.004	10.467	0.106
13	1.004	10.734	0.113
14	1.004	10.964	0.117
15	1.004	11.129	0.120
16	1.004	11.268	0.122
17	1.004	11.413	0.124
<i>girls</i>			
9	1.045	9.717	0.072
10	1.045	9.752	0.079
11	1.045	9.785	0.085
12	1.045	9.826	0.090
13	1.045	9.862	0.092
14	1.045	9.888	0.094
15	1.045	9.908	0.095
16	1.045	9.931	0.096
17	1.045	9.958	0.097

L, skew; M, median; S, coefficient of variation.

The LMS values depict the nature of the age-related distributions for boys and girls and can be used to calculate z-scores, and hence percentile values by looking up a z-table, using the following formula:

$$z = \frac{\left(\frac{x}{M}\right)^L - 1}{L \times S}$$

where, z is the z score, x is the 20mSRT running speed at the last complete stage, L is the sex- and age-specific L-value, M is the sex- and age-specific M-value, and S is the sex- and age-specific S-value.

SUPPLEMENTAL MANUSCRIPT 2

This manuscript has been written and formatted following the publication standards of the peer-reviewed journal, *British Journal of Sports Medicine*.

Title:

Temporal trends in the cardiorespiratory fitness of 965,264 children and youth representing 19 countries since 1981.

Status:

Under review

Title

Temporal trends in the cardiorespiratory fitness of 965,264 children and youth representing 19 countries since 1981

Authors

Grant R. TOMKINSON *PhD**

Department of Kinesiology and Public Health Education, University of North Dakota,
2751 2nd Avenue North, Stop 8235, Grand Forks, ND, 58202, USA

Alliance for Research in Exercise, Nutrition and Activity (ARENA), School of Health
Sciences & Sansom Institute for Health Research, University of South Australia, GPO
Box 2471, Adelaide, SA, Australia

Justin J. LANG *MSc*

Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of
Eastern Ontario Research Institute, 401 Smyth Rd Ottawa, ON, K1H 8L1, Canada

Mark S. TREMBLAY *PhD*

Healthy Active Living and Obesity (HALO) Research Group, Children's Hospital of
Eastern Ontario Research Institute, 401 Smyth Rd Ottawa, ON, K1H 8L1, Canada

Running title

Temporal trends in cardiorespiratory fitness of children

Corresponding author

Dr. Grant R. TOMKINSON*

✉ Department of Kinesiology and Public Health Education
University of North Dakota
2751 2nd Avenue North, Stop 8235
Grand Forks, ND, 58202, USA

☎ +1 701-777-4041

📧 grant.tomkinson@und.edu

Keywords: Time changes; 20 m shuttle run; youth; aerobic fitness; endurance;
international.

Word count: 3330

Abstract

Objective: To estimate and compare international temporal trends in the CRF of children and youth.

Methods: Data were obtained from a systematic search of studies that explicitly reported 20 m shuttle run test descriptive data (a validated measure of CRF) on apparently healthy 9–17 year-olds. Following the estimation of relative peak oxygen uptake (mL/kg/min), sample-weighted temporal trends were estimated at the country-sex-age level using best-fitting linear or polynomial regression models relating the year of testing to mean CRF. Post-stratified population-weighted mean changes in absolute and percent CRF were estimated.

Results: Temporal trends were estimated from 965,264 children and youth from 19 countries (high- and middle-income) between 1981 and 2014, using data from 137 studies. Collectively, mean CRF declined by 3.3 mL/kg/min (95%CI: –3.5 to –3.1) or 7.3% (95%CI: –7.8 to –6.7). This decline was not uniform across the entire period, with the rate of decline diminishing with each decade and recently stabilising near zero. The rate of decline was larger for boys than girls, and similar for children and youth. Temporal trends in CRF differed between countries.

Conclusions: There has been a substantial decline in CRF since 1981, which is suggestive of a meaningful decline in population health and well-being. However, the trend in CRF has not followed the anticipated trajectory, by diminishing since 2000 and stabilising with almost no change recently. This recent trend may reflect numerous campaigns and interventions promoting increased physical activity and sport participation, at least among high-income countries.

Introduction

Cardiorespiratory fitness (CRF) is an important marker of good health^{1,2} and is the fourth-leading risk factor for cardiovascular disease.³ In adults, CRF is strongly and independently associated with cardiovascular and all-cause mortality and morbidity,^{4,5} stroke,⁶ cancer,⁷ diabetes and many other risk factors and co-morbidities.⁸ A meta-analysis by Kodama and colleagues⁹ reported that every metabolic equivalent (MET; 3.5 mL/kg/min) decrease in CRF was associated with a 15% and 13% increase in cardiovascular and all-cause mortality, respectively. While meaningful associations between CRF and cardiovascular disease risk have been reported in young people,^{1,10} until recently, the association between CRF and mortality had not been examined directly. Two studies, one which followed-up 1.3 million 18 year-old Swedish men for 29 years,¹¹ and another which followed-up 510 16 year-old Japanese girls for 64 years,¹² found that low CRF in late adolescence was significantly linked with all-cause mortality in adulthood. This evidence highlights the importance of measuring CRF when trying to screen and track the health and well-being of children and youth.

While the popular belief is that children's CRF has declined in recent decades, informed debate is divided.^{13–15} Contributing to this divide is that researchers have not always been clear on their definition of CRF and the operationalisation of measurement. Gas analysed peak oxygen uptake ($\dot{V}O_{2\text{peak}}$) is the criterion measure of children's CRF,¹⁵ but it is not practical for mass testing. Despite data indicating little change in gas analysed CRF over time,^{15–18} our understanding is confounded by pooled data acquired from different

ergometers and measurement protocols, and small samples of children who may be athletically biased.

Properly conducted field-based CRF testing offers a simple, feasible, practical, reliable and valid alternative to gas analysis,¹⁹ and has been widely used on children and youth. Hitherto, the most comprehensive summaries of temporal trends in field-based CRF indicate international declines from the mid-1970s to the early 2000s.^{13,20} Data on 129,882 children and youth aged 6–19 years from 11 countries indicated a decline in mean CRF (operationalised as performance on the 20 m shuttle run test [20mSRT]²¹—a progressive exercise test involving continuous running between two lines 20 m apart in time to recorded beeps) of 4.3% per decade between 1981 and 2000.²⁰ In a follow-up study, mean CRF (operationalised as distance running and 20mSRT performance) was reported to have declined by 3.6% per decade in 25.4 million 6–19 year olds from 27 countries between 1958 and 2003.¹³ Unfortunately, little is known about current international and national trends, because the data coverage in these large summary studies extends only to 2003. Furthermore, studies to date have been conducted on samples without the use of systematic search strategies to acquire all available data.

The aim of this study was to systematically gather and analyse international and national temporal trends in the CRF of children and youth. CRF was operationalised as relative $\dot{V}O_{2\text{peak}}$ (mL/kg/min) estimated from the 20mSRT, a reliable, valid, and popular field test of CRF.^{13, 19–21} Examination of time trends in CRF: (i) provides meaningful insight into trends in population health and well-being, (ii) provides an international context from

which national trends can be compared, and subsequently (iii) informs population health programs, policies and priorities.

Methods

Data sources

A systematic literature search was conducted to identify studies that reported descriptive 20mSRT data on apparently healthy (free from known disease/injury) 9–17 year-olds (details described elsewhere).²² Briefly, a search of five databases, including MEDLINE (1946–2015), PsycINFO (1806–2014), EMBASE (1947–2014), SPORTDiscus (1949–2014), and the Cochrane Central Register of Controlled Trials (2005–2014), was conducted (October, 2015). The search strategy included the terms: shuttle run*, OR beep test, OR multi-stage, OR aerobic, OR cardio*, OR endurance; with child*, adolescen*, pubescen* boy, girl, young, and youth as search term modifiers. No search restriction was placed on language. Studies were included if descriptive data were provided for 20mSRT performance (1-min stage protocol) at the country-sex-age level. The reference lists of all included studies and personal libraries of the authors were reviewed. When needed, included study authors were invited to clarify study details, provide raw data, and/or asked if they knew of other relevant studies.

Inclusion/exclusion criteria

Of the 979 studies identified as relevant, 802 studies were excluded as described in Figure 1 (details described elsewhere).²² Descriptive data from the remaining 177 studies were extracted by a single author into a spreadsheet using a standardised template, with

data checked and corrected by two authors using procedures reported elsewhere.²² Only country-sex-age groups that included data across at least two time points spanning a minimum of five years were examined. A minimum of five country-sex-age groups (e.g., 9 year-old Australian boys, 9 year-old Australian girls) was required per country for inclusion. This step resulted in the further exclusion of 40 studies, yielding a final dataset of 137 studies (Figure 1; Supplement 1).

<<< *Figure 1* >>>

Data treatment and statistical analysis

All 20mSRT data were expressed using a common protocol and result metric (running speed at the last completed stage, km.h⁻¹) using a detailed data standardisation procedure reported elsewhere.^{20,22} The accuracy of this data standardisation procedure is excellent, with differences in means calculated using this standardisation procedure, and direct calculation from individual final running speeds, <0.5% at the country-sex-age level.²⁰ CRF was estimated by the widely used Léger²¹ prediction equation:

$$\dot{V}O_{2\text{peak}} \text{ (mL/kg/min)} = 31.025 + 3.238 \text{ Speed} - 3.248 \text{ Age} + 0.1536 \text{ Speed} \cdot \text{Age}$$

where *Speed* is the running speed of the last completed stage (km.h⁻¹) and *Age* is age at last birthday. In a sample of Québec children and youth, this equation had a standard error of estimate of 5.9 mL/kg/min or 12.1%.²¹

Temporal trends in mean CRF were analysed using an update of the procedure described in detail elsewhere.^{13,23} Temporal trends at the country-sex-age level were analysed using best-fitting sample-weighted linear or polynomial (quadratic or cubic) regression models relating the year of testing to CRF. Changes in mean CRF were expressed as absolute and percent changes. Absolute changes (mL/kg/min per year) were calculated as the slope of

the regression line, and percent changes (% per year) as the slope of the regression line expressed as a percentage of the sample-weighted mean value (a series of slopes on the polynomial curves were calculated by differentiation for every included year of testing). To interpret the magnitude of change, standardised effect sizes (ES) of 0.2, 0.5 and 0.8 were used as thresholds for small, moderate and large respectively, with effects <0.2 considered to be negligible and ≥ 0.2 considered to be substantial.²⁴ Positive changes indicated increases in mean CRF and negative changes indicated declines.

The temporal trends were described as follows: starting with the first year (Y_1) covered by any study-country-sex-age group, every group including Y_1 in its span of years was located, with every change (dx_1 , mL/kg/min or % per year) recorded.^{13,23} This process was applied to all years for which change data were available ($Y_1 \dots Y_n$), yielding a series of yearly changes. The post-stratified population-weighted mean yearly change was calculated for year (Y_1),²⁵ which was repeated for $Y_2, Y_3, Y_4 \dots$ until the last year covered by any study, Y_n . This process yielded a series of population-weighted mean yearly changes ($dx_1, \dots, dx_n$) that describe the temporal patterns of change; a process that was performed for all children and for different sex, age and country groups. Population estimates were standardised to the mean measurement year of 2000 using United Nations data.²⁶ The post-stratification population-weighting procedure helped to correct the trends for systematic bias associated with over- and under-sampling, and to standardise the trends to underlying country-sex-age-specific demographics. Temporal trends were estimated using best-fitting population-weighted linear or polynomial regression models relating the year of testing to the yearly changes, and were graphically illustrated using an iterative procedure described elsewhere.¹³

Results

Temporal trends in CRF were estimated from 965,264 children and youth aged 9–17 years from 19 countries (1,408 country-sex-age-year groups) between 1981 and 2014 (Table 1). These 19 countries represented 6 of 7 continents, approximately 15% of the world's population, and included 17 high-income and two upper-middle-income countries.²⁷ National surveys provided 64% of all data points, state/provincial studies 24%, and community studies 12% (Supplement 1). Trends were calculated for 262 country-sex-age groups (boys: 125; girls: 137; children [9–12 year-olds]: 107; youth [13–17 year-olds]: 155), with a median sample size of 1,138 (range: 100 to 54,567) and a median span of 16 years (range: 5 to 30).

<<< *Table 1* >>>

Between 1981 and 2014, mean CRF collectively declined by 3.3 mL/kg/min (95%CI: –3.5 to –3.1; ES: moderate) at 1.0 mL/kg/min per decade (95%CI: –1.1 to –0.9), equivalent to a decline of 7.3% (95%CI: –7.8 to –6.7) at 2.2% per decade (95%CI: –2.4 to –2.0) (Table 2; Figure 2a). CRF declined in boys by 4.3 mL/kg/min (95%CI: –4.7 to –3.9; ES: large) at 1.3 mL/kg/min per decade (95%CI: –1.4 to –1.2), which was larger than the decline in girls (ES difference: small), who declined by 2.3 mL/kg/min (95%CI: –2.6 to –2.0; ES: small) at 0.7 mL/kg/min per decade (95%CI: –0.8 to –0.6) (Table 2; Figure 2a). The collective decline in CRF in children and youth was similar (ES difference: negligible; Table 2; Figure 2a).

<<< *Table 2 and Figure 2* >>>

Internationally, the decline in mean CRF was not uniform over time, with the rate of

decline slowing from 1981 to 2000 and stabilising near zero thereafter (Figure 2b). Since 1981, the rate of decline in CRF has diminished from -2.2 mL/kg/min per decade (95%CI: -2.4 to -2.0 ; ES: small) in the 1980s, -1.6 mL/kg/min per decade (95%CI: -1.7 to -1.5 ; ES: small) in the 1990s, -0.4 mL/kg/min per decade (95%CI: -0.5 to -0.3 ; ES: negligible) in the 2000s, to -0.2 mL/kg/min per decade (95%CI: -0.4 to 0.0 ; ES: negligible) in the 2010s. This trend was consistent across different sex and age groups (Figure 2b).

Temporal trends in mean CRF differed by country, ranging from a decline of 5.6 mL/kg/min per decade (95%CI: -6.2 to -5.0 ; ES: large) in Poland (1992–2007) to an increase of 1.4 mL/kg/min per decade (95%CI: 1.3 to 1.5 ; ES: small) in Japan (1998–2014) (Table 2). Most countries experienced declines in CRF (17 of 19 or 89%), ranging from large declines (ES <-0.8) in Australia, Canada, France, Hungary, Lithuania, Poland, and Seychelles, to negligible declines (ES <-0.2) in Belgium, Estonia, and Greece. The increases in Brazil and Japan were small (ES >0.2). While most country trends were approximately uniform (linear), Figure 3 indicates non-uniform (curvilinear) changes evidenced by a slowing or stabilising of the rate of decline near the year 2000 for several countries (Australia, Canada, Greece, South Africa and Spain), a shift from declines to increases in Belgium, and a shift from increases to declines in the UK. Cross-country correlations were very strong between boys and girls ($r=0.96$ [95% CI: 0.89 to 0.99]) and children and youth ($r=0.86$ [95% CI: 0.66 to 0.95]).

<<< *Figure 3* >>>

Discussion

This analysis indicates that since 1981: (i) there has been a substantial international decline in CRF, which has slowed markedly since 2000 and stabilised with almost no change recently; (ii) despite most countries experiencing a decline in CRF, both the rate and pattern of change differed between countries; and (iii) the decline in CRF has been larger for boys than for girls. As low CRF in late adolescence and adulthood is significantly associated with all-cause mortality later in life,^{11,12} these temporal declines in CRF are therefore suggestive of meaningful declines in population health and well-being. This highlights a cause for concern. For example, the observed international decline in CRF of 3.3 mL/kg/min, approximately 1 MET, has been associated with an increase in cardiovascular and all-cause mortality of 15% and 13%, respectively.⁹

Slowing international decline in CRF

While the declines in CRF observed in the 1980s and 1990s are consistent with those reported in several other comprehensive summaries of temporal trends in CRF,^{13,20} this study provides evidence that the decline in CRF has diminished and stabilised with almost no change recently, contradicting the popular belief that CRF is still declining in children and youth. Temporal trends in field-based CRF are probably caused by trends in a network of environmental, social, behavioral, physical, psychosocial and physiological factors.^{13,15,23} Causes of a temporal decline in CRF are not only a function of changes in $\dot{V}O_{2\text{peak}}$ but also changes in mechanical efficiency and fractional utilisation of oxygen, with psychosocial issues (e.g., motivation) also important.^{28–30} Unfortunately, there are no compelling trend data for these factors. Physiological changes are affected by physical

changes, such as increased fatness. An increase in fatness could lead to a decline in CRF either directly through the increased energy demand associated with moving a heavier body mass through space, or indirectly through the likely effect of reduced regular moderate-to-vigorous physical activity.^{15,31} There is overwhelming evidence of an international increase in children's fatness over the past 50 years,^{32–33} which may have stabilised since the mid 1990s in high-income countries.³⁴ This increase and recent stabilisation in fatness has coincided with the decline and recent stabilisation in CRF. Direct analysis of this causal connection indicates that increases in fatness explain 35–70% of the declines in CRF.^{31,35} Other factors, such as a decline in physical activity or an increase in sedentary behaviours, are very likely involved.¹³ Unfortunately, temporal trends in these behaviors are not clear, because of the difficulty in obtaining accurate measurements and temporal differences in sampling and methodology.³⁶ Recent reviews of temporal trends in self-report and objective measures of physical activity in children suggested that physical activity levels and sport participation have not changed in recent decades.^{36,37} Given that physical fitness assesses an individual's physiological response to their total physical activity profile, temporal trends in CRF could therefore suggest that physical activity levels have declined in the 1980s and 1990s and stabilised thereafter, at least in the predominantly high-income countries in this dataset.

Furthermore, it is not clear whether the observed decline in mean CRF has been uniform or skewed across the distribution. Several studies have reported that temporal trends are skewed in the distribution of CRF, with declines in distance running performance larger in children with low CRF than in children with average or high CRF.^{38,39} However,

temporal trends in skewness are probably an artifact of expressing CRF in running time, and are possibly attenuated when expressed in mL/kg/min, which is typically normally distributed and more appropriate for parametric statistical analysis.

Temporal differences in CRF between countries

Our analysis was largely representative of high-income countries that are well along the physical activity and epidemiological transition,⁴⁰ which may help explain why CRF has recently stabilised. Many of these countries have placed substantial effort into developing campaigns, policies, programs, and interventions to promote increased sport participation and daily physical activity.^{41,42} The cumulative effect of these efforts appears to have resulted in the recent stabilisation of overweight and obesity rates seen in many high-income countries (as noted above).³⁴ These are promising findings that may provide a ‘lesson learned’ for low- and middle-income countries that are likely to experience (or are experiencing) substantial declines in CRF similar to those experienced in high-income countries in the 1980s and 1990s. Evidence from trends in CRF⁴³ and overweight and obesity rates⁴⁴ in these countries supports this theory. This highlights the need for low- and middle-income countries to collect CRF data to track trends in population health and fitness in order to help inform policies to prevent declines like those seen in high-income countries in the 1980s and 1990s. The 20mSRT is an excellent tool for population surveillance because it is valid and reliable for use with children, is low cost, requires minimal equipment and personnel, can be used across multiple locations (e.g., indoors, outdoors, smaller spaces) and is practical for mass testing.^{13,19,45} Lang and colleagues recently suggested establishing an international 20mSRT data repository to help monitor

CRF.⁴⁶ This type of data repository would be excellent for tracking trends in CRF in low- and middle- income countries with minimal finances to support other surveillance efforts.

Temporal differences in CRF between sexes

It is not clear why declines in CRF were larger for boys than for girls. Olds³² indicated that increases in percentage body fat were larger for boys (1.0% per decade) than for girls (0.7% per decade) in 30 high-income countries between 1952 and 2003. Cross-sectionally, the negative relationship between fatness (sum of skinfolds) and CRF (20mSRT performance) is stronger for boys than for girls,⁴⁷ suggesting that any given increase in fatness will result in a greater decline in boys' CRF. The temporal difference in CRF between sexes may also be a result of a temporal difference in physical activity participation between sexes, which is apparent in elite sport. For example, the 2012 Olympic Games marked the first time a female athlete was part of every competing country's delegation, with females comprising 44% of all athletes, increasing from 23% in 1984 and 11% in 1960.⁴⁸ Although this highlights the diminishing gender gap in elite sport, it may reflect that sport participation for girls of all ages has increased internationally. Perhaps the recent promotion of gender equality policies and programs⁴⁹ has helped narrow the gender gap in CRF through sport participation and other physical activity areas.

Strengths and limitations

Using a systematic search and analytical strategy, this study cumulated data from 137 studies on 965,264 children and youth from 19 countries to estimate and compare

international temporal trends in CRF. Although not the first study examining international trends in CRF, it does update previous comprehensive studies^{13,20} by: (i) extending the end of the temporal picture from 2003 to 2014; (ii) reporting absolute and percent trends for a single validated measure, (iii) reporting trends in $\dot{V}O_{2peak}$ which reflect underlying trends in the energetic cost of exercise, and (iv) describing trends using weighted regression and a post-stratification population weighting procedure.

There are however several limitations. First, the reported trends are representative of high-income and upper-middle-income countries, and it is not known whether similar trends exist for low-income countries that may be experiencing a physical activity transition.⁴⁰ Second, this study pooled data from studies that used different sampling methods (probability and non-probability sampling) and sampling frames (national, state/provincial, and community levels), which raises the issue of representativeness. However, because trends in population-representative samples could not always be estimated, trend estimates across state/provincial and community samples of children were included because they were considered to provide meaningful insight into country-level trends. Third, differences in testing conditions (e.g., climate, altitude, practice, and running surfaces) and measurement errors (e.g., methodological drift and diurnal variation) might have occurred, although the large number of included data points should have minimised these methodological issues. Fourth, international, regional, and country temporal trends were estimated from available country-sex-age-specific trend data, and therefore may not be representative of all country-sex-age groups or all measurement years. Finally, temporal trends in CRF will be affected by concurrent trends in biological

maturation,⁵⁰ which generally favor children of the same chronological age in more recent years and may underestimate the true temporal declines.²⁰

Conclusion

This study indicates that there has been a substantial decline in CRF in high-income countries since 1981, which has slowed over time and recently stabilised near zero. While declines in CRF have typically been observed at the national level, variability exists in both the rate and pattern of change with boys experiencing larger declines than girls. This study highlights a gap in CRF trend data for low- and middle-income countries, highlighting a need to improve the international monitoring of CRF across populations because temporal declines are indicative of declines in population health and well-being. The recent diminishing and stabilising of the decline in CRF in high-income countries is a promising new finding and may be the result of the numerous campaigns and interventions promoting increased physical activity and sport participation at the community, state, national and international levels. To most accurately track trends in child and youth health, it is recommended that large-scale, harmonised national health and fitness surveys be regularly conducted, especially among low- and middle-income countries.

Acknowledgements

We would like to thank the authors of the included studies for generously clarifying details of their studies and/or for providing raw and/or additional data.

Funding

Nil.

Declaration of interest

All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Contributors

GRT, JYL, and MST developed the research question and designed the study. GRT and JYL had full access to all the data in the study and take responsibility for the integrity of the data. GRT led the statistical analysis, synthesis of results, and writing of the report. All authors contributed to the interpretation of results, editing and critical reviewing of the final report, and approved the final report.

References

- 1 Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: a powerful marker of health. *Int J Obes*. 2008; 32(1): 1–11.
- 2 Ross R, Blair SN, Arena R, Church TS, Després JP, Franklin BA, Haskell WL, Kaminsky LA, Levine BD, Lavie CJ, Myers J, Niebauer J, Sallis R, Sawada SS, Sui X, Wisløff U. Importance of assessing cardiorespiratory fitness in clinical practice: A case for fitness as a clinical vital sign. A scientific statement from the American Heart Association. *Circulation*. 2016; 134(24): e653–e699.
- 3 Fletcher GF, Blair SN, Blumenthal J, Caspersen C, Chaitman B, Epstein S, Falls H, Froelicher ES, Froelicher VF, Pina IL. Statement on exercise: benefits and recommendations for physical activity programs for all Americans: a statement for health professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology, American Heart Association. *Circulation*. 1992;86:340–344.
- 4 Myers J, Prakash B, Groelicher V, Partington S, Atwood JE. Exercise capacity and mortality among men referred for exercise testing. *N Engl J Med*. 2002; 346: 793–801.
- 5 Gulati M, Pandey DK, Arnsdorf MF, Lauderdale DS, Thisted RA, Wicklund RH, Al-Hani AJ, Black HR. Exercise capacity and the risk of death in women: the St James Women Take Heart Project. *Circulation*. 2003;108:1554–1559. doi: 10.1161/01.CIR.0000091080.57509.E9.
- 6 Lee CD, Blair SN. Cardiorespiratory fitness and stroke mortality in men. *Med Sci Sports Exerc*. 2002; 34(4): 592–5.
- 7 Sawada SS, Lee IM, Naito H, Kakigi R, Goto S, Kanazawa M, Okamoto T, Tsukamoto K, Muto T, Tanaka H, Blair SN. Cardiorespiratory fitness, body mass index, and cancer mortality: a cohort study of Japanese men. *BMC Public Health*. 2014;14:1012. doi: 10.1186/1471-2458-14-1012.
- 8 Katzmarzyk P, Church TS, Janssen I, Ross R, Blair SN. Metabolic syndrome, obesity, and mortality: impact of cardiorespiratory fitness. *Diabetes Care*. 2005; 28(2): 391–7.
- 9 Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, Sugawara A, Totsuka K, Shimano H, Ohashi Y, Yamada N, Sone H. 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.
- 10 Eisenmann JC, Wickel EE, Welk GJ, Blair SN. Relationship between adolescent fitness and fatness and cardiovascular disease risk factors in adulthood: the Aerobics Center Longitudinal Study (ACLS). *Am Heart J*. 2005; 149(1): 46–53.

- 11 Högström G, Nordström A, Nordström. Aerobic fitness in late adolescence and the risk of early death: a prospective cohort study of 1.3 million Swedish men. *Int J Epi*. doi: 10.1093/ije/dyv321.
- 12 Sato M, Kodama S, Sugawara A, Saito K, Sone H. Physical fitness during adolescence and adult mortality. *Epidemiology*. 2009; 20(3): 463–4.
- 13 Tomkinson GR, Olds TS. Secular changes in pediatric aerobic fitness test performance: The global picture. *Med Sci Sport Exerc*. 2007; 50: 46–66.
- 14 Rowland T. Declining cardiorespiratory fitness in youth: Fact or supposition? *Pediatr Exerc Sci*. 2002; 14(1): 1–8.
- 15 Armstrong N, Tomkinson G, Ekelund U. Aerobic fitness and its relationship to sport, exercise training and habitual physical activity during childhood and adolescence. *Br J Sports Med*. 2011; 45(11): 849–58.
- 16 Eisenmann JC, Malina RM. Secular trend in peak oxygen consumption among United States youth in the 20th century. *Am J Hum Biol*. 2002; 14(6): 699–706.
- 17 Lin WS, Ji CY, Zhang YQ, Miroslav Z, Shao LX, Guan MJ. Maximal aerobic power in children and adolescents of Beijing, China. *Am J Hum Biol*. 1996; 8: 497–503.
- 18 Armstrong N, Williams J, Balding J, Gentile P, Kirby B. The peak oxygen uptake of British children with reference to age, sex, and sexual maturity. *Eur J Appl Physiol*. 1991; 62(5): 369–75.
- 19 Tomkinson G, Olds T. Field tests of fitness. In: Armstrong N, Van Mechelen W, eds. *Paediatric exercise science and medicine*. 2nd ed. Oxford, United Kingdom: Oxford University Press; 2008: 109–28.
- 20 Tomkinson G, Léger L, Olds T, Cazorla G. Secular trends in the performance of children and adolescents (1980–2000): An analysis of 55 studies of the 20 m shuttle run in 11 countries. *Sports Med*. 2003; 33(4): 285–300.
- 21 Léger 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.
- 22 Tomkinson GR, Lang JJ, Tremblay MS, Dale M, LeBlanc AG, Belanger K, Ortega FB, Léger L. International normative 20 m shuttle run values from 1,142,026 children and youth representing 50 countries. *Brit J Sports Med*. doi:10.1136/bjsports-2016-095987.
- 23 Tomkinson G, Macfarlane DJ, Noi S, Kim DY, Wang Z, Hong R. Temporal changes in long-distance running performance of Asian children between 1964 and 2009. *Sports Med*. 2012; 42(4): 267–79.
- 24 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. New Jersey, USA: Lawrence Erlbaum; 1988.
- 25 Levy PS, Lemeshow S. Stratification random sampling: further issues. In: Levy PS, Lemeshow S, eds. *Sampling of Populations: Methods and Application*. Hoboken, NJ: John Wiley & Sons, Inc.; 2008: 143–88.

- 26 United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects: The 2015 Revision, Key Findings and Advance Tables; 2015. Working Paper No. ESA/P/WP.241.
- 27 The World Bank. *World Bank country and lending groups*. Available from: <http://data.worldbank.org/about/country-and-lending-groups>. Accessed December 24, 2015.
- 28 Krahenbuhl GS, Pangrazi RP, Chomokos EA. Aerobic responses of young boys to submaximal running. *Res Q*. 1979; 50(3): 413–21.
- 29 Lussier L, Buskirk ER. Effects of an endurance training regimen of assessment of work capacity in prepubertal children. *Ann NY Acad Sci*. 1977; 301: 734–47.
- 30 Sjödin B, Svedenhag J. Applied physiology of marathon running. *Sports Med*. 1985; 2(2): 83–99.
- 31 Olds TS, Dollman J. Are changes in distance-run performance of Australian children between 1985 and 1997 explained by changes in fatness? *Pediatr Exerc Sci*. 2004; 16(3): 201–9.
- 32 Olds T. One million skinfolds: secular trends in the fatness of young people 1951–2004. *Eur J Clin Nutr*. 2008; 63(8): 934–46.
- 33 Wang Y, Lobstein T. Worldwide trends in childhood overweight and obesity. *Int J of Pediatr Obes*. 2006; 1(1): 11–25.
- 34 Olds T, Maher C, Zumin S, Péneau S, Lioret S, Castetbon K, Bellisle, de Wilde J, Hohepa M, Maddison R, Lissner L, Sjöberg A, Zimmermann M, Aeberli I, Ogden C, Flegal K, Summerbell C. Evidence that the prevalence of childhood overweight is plateauing: data from nine countries. *Int J Pediatr Obes*. 2011; 6(5-6): 342–60.
- 35 Albon HM, Hamlin MJ, Ross JJ. Secular trends and distributional changes in health and fitness performance variables of 10–14-year-old children in New Zealand between 1991 and 2003. *Br J Sports Med*. 2010; 44(4): 263–9.
- 36 Ekelund U, Tomkinson GR, Armstrong N. What proportion of youth are physically active? Measurement issues, levels and recent time trends. *Br J Sports Med*. 2011; 45(11): 859–65.
- 37 Booth VM, Rowlands AV, Dollman J. Physical activity temporal trends among children and adolescents. *J Sci Med Sport*. 2015; 18: 418–25.
- 38 Dollman J, Olds T, Norton K, Stuart D. The evolution of fitness and fatness in 10–11-year-old Australian schoolchildren: changes in distributional characteristics between 1985 and 1997. *Pediatr Exerc Sci*. 1999; 11(2): 108–21.
- 39 McNaughton L, Morgan R, Smith P, Hannan G. An investigation into the fitness levels of Tasmanian primary schoolchildren. *ACHPER Healthy Lifestyles J*. 1996; 43: 4–10.
- 40 Katzmarzyk PT, Mason C. The physical activity transition. *J Phys Act Health*. 2009; 6(3): 269–80.

- 41 Heath GW, Parra DC, Sarmiento OL, Andersen LB, Owen N, Goenka S, Montes F, Brownson RC. Evidence-based intervention in physical activity: Lessons from around the world. *Lancet*. 2012; 380: 272–81.
- 42 Reis RS, Salvo D, Ogilvie D, Lambert EV, Goenka S, Brownson RC. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. *Lancet*. doi:10.1016/S0140-6736(16)30728-0.
- 43 Dos Santos FK, Prista A, Gomes TN, Daca T, Katzmarzyk PT, Maia JA. Secular trends in physical fitness of Mozambican school-aged children and adolescents. *Am J Hum Biol*. 2015; 27(2): 201–6.
- 44 Muthuri SK, Wachira LM, Leblanc AG, Francis CE, Sampson M, Onywera VO, Tremblay MS. Temporal trends and correlates of physical activity, sedentary behaviour, and physical fitness among school-aged children in sub-Saharan Africa: a systematic review. *Int J Environ Res Public Health*. 2014; 11(3): 3327–59.
- 45 Ruiz JR, Silva G, Oliveira N, Ribeiro JC, Oliveira JF, Mota J. Criterion related validity of the 20m shuttle run test in adolescents aged 13–19 years. *J Sports Sci*. 2009; 27(9): 899–06.
- 46 Lang JJ, Tremblay MS, Léger L, Olds TS, Tomkinson GR. International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results. *Brit J Sports Med*; doi:10.1136/bjsports-2016-096224.
- 47 Olds T, Ridley K, Tomkinson G. Declines in aerobic fitness: Are they only due to increasing fatness? *Med Sport Sci Exerc*. 2007; 50: 226–40.
- 48 International Olympic Committee. IOC factsheet women in the Olympic movement—updated January 2016. Available from http://www.olympic.org/Documents/Reference_documents_Factsheets/Women_in_Olympic_Movement.pdf. Accessed August 28, 2016.
- 49 UN Development Goals and Beyond 2015: Goal 5 Achieve gender equality and empower all women and girls. <http://www.un.org/sustainabledevelopment/gender-equality/>. Accessed August 28, 2016.
- 50 Malina RM. Secular trends in growth, maturation and physical performance: A review. *Anthropological Review*. 2004; 67: 3–31.

Table 1 Summary of the included studies by country.

Country	Sex	Ages (years)	Sample size	Country-sex-age- year groups	Span of years
Australia ^{w1-w12}	Boys+Girls	9–16	25,961	101	1988–2010
Belgium ^{w13-w23}	Boys+Girls	9–17	23,358	117	1985–2010
Brazil ^{w24-w26}	Girls	9–17	2,826	24	1998–2006
Canada ^{w27-w32}	Boys+Girls	9–17	29,009	91	1981–2011
Estonia ^{w21,w33-w38}	Boys+Girls	10–17	7,906	46	1992–2002
France ^{w18,w39-w45}	Boys+Girls	9–14	5,763	32	1985–2007
Greece ^{w18,w46-w50}	Boys+Girls	9–17	268,382	55	1991–2014
Hungary ^{w18,w21,w51,w52}	Boys+Girls	12–17	2,455	22	1990–2013
Italy ^{w18,w53-w58}	Boys+Girls	12–17	5,218	31	1986–2007
Japan ^{w59-w75}	Boys+Girls	9–17	279,732	306	1998–2014
Lithuania ^{w34,w35}	Boys+Girls	11–17	6,479	28	1992–2002
Netherlands ^{w76-w78}	Boys+Girls	12–16	3,457	22	1987–2006
Poland ^{w79-w82}	Boys+Girls	13–17	22,895	19	1992–2007
Portugal ^{w83-w89}	Boys+Girls	9–17	29,857	71	1999–2011
Seychelles ^{w90,w91}	Boys+Girls	11–16	7,913	24	1990–2004
South Africa ^{w92-w99}	Boys+Girls	11–16	2,668	30	1996–2012
Spain ^{w18,w100-w112}	Boys+Girls	9–17	25,950	151	1985–2010
UK ^{w113-w126}	Boys+Girls	9–16	48,535	116	1986–2014
USA ^{w127-w137}	Boys+Girls	9–17	166,900	122	1995–2013

Note: UK=United Kingdom; USA=United States of America.

Table 2 National and international temporal trends in mean CRF (per decade) from 965,264 children aged 9–17 years from 19 countries between 1981 and 2014.

Group	Absolute changes in means per decade mL/kg/min (95%CI)	Percent changes in means per decade (95%CI)
Australia	−1.7 (−2.0 to −1.4)	−3.9 (−4.6 to −3.2)
Belgium	−0.4 (−0.7 to −0.1)	−1.0 (−1.7 to −0.3)
Brazil	1.2 (0.7 to 1.7)	3.0 (1.8 to 4.2)
Canada	−1.4 (−1.6 to −1.2)	−3.0 (−3.4 to −2.6)
Estonia	−0.2 (−0.4 to 0.0)	−0.3 (−0.6 to 0.0)
France	−2.7 (−3.0 to −2.4)	−5.7 (−6.3 to −5.1)
Greece	−0.4 (−0.5 to −0.3)	−0.8 (−1.1 to −0.5)
Hungary	−1.8 (−2.0 to −1.6)	−4.4 (−4.8 to −4.0)
Italy	−0.6 (−0.8 to −0.4)	−1.6 (−2.2 to −1.0)
Japan	1.4 (1.3 to 1.5)	2.9 (2.8 to 3.0)
Lithuania	−5.0 (−5.2 to −4.8)	−11.2 (−11.6 to −10.8)
Netherlands	−0.7 (−0.9 to −0.5)	−1.6 (−2.0 to −1.2)
Poland	−5.6 (−6.2 to −5.0)	−12.9 (−14.4 to −11.4)
Portugal	−0.9 (−1.2 to −0.6)	−1.9 (−2.6 to −1.2)
Seychelles	−3.1 (−3.2 to −3.0)	−7.3 (−7.5 to −7.1)
South Africa	−2.5 (−3.2 to −1.8)	−5.8 (−7.3 to −4.3)
Spain	−0.8 (−0.9 to −0.7)	−1.8 (−2.1 to −1.5)
United Kingdom	−1.0 (−1.2 to −0.8)	−2.2 (−2.6 to −1.8)
United States	−0.9 (−1.2 to −0.6)	−2.2 (−2.8 to −1.6)
Boys	−1.3 (−1.4 to −1.2)	−2.9 (−3.1 to −2.7)
Girls	−0.7 (−0.8 to −0.6)	−1.6 (−1.8 to −1.4)
Children (9–12 year-olds)	−1.0 (−1.1 to −0.9)	−2.1 (−2.3 to −1.9)
Youth (13–17 year-olds)	−1.0 (−1.1 to −0.9)	−2.4 (−2.6 to −2.2)
All	−1.0 (−1.1 to −0.9)	−2.2 (−2.4 to −2.0)

Note: positive changes in means indicate trend increases in mean CRF and negative values indicate trend declines; 95%CI=95% confidence interval.

Figure captions

Figure 1 Flow chart outlining the identification of the included studies.

Figure 2 International temporal trends in mean CRF (mL/kg/min) between 1981 and 2014. Note: Figure 2a shows the absolute differences in CRF, with the grey child-like figure indicating mean CRF in 1981 and the white child-like figure indicating mean CRF in 2014. Figure 2b shows the time-related patterns of change in CRF. Data are standardised to the year 2000=0 mL/kg/min, with positive values indicating better CRF and negative values poorer CRF; the solid lines represent the international change in mean CRF and the shaded areas the 95% CIs, with upward sloping lines indicating increases over time and downward sloping lines declines; mean (95%CI) absolute changes (per decade) are shown at the top of each panel.

Figure 3 National temporal trends in mean CRF (mL/kg/min) between 1981 and 2014. Note: data are standardised to the year 2000=0 mL/kg/min, with positive values indicating better CRF and negative values poorer CRF; the solid lines represent the national changes in mean CRF and the shaded areas the 95% CIs, with upward sloping lines indicating increases over time and downward sloping lines declines; mean (95%CI) absolute changes (per decade) are shown at the top of each panel.

Figure 1

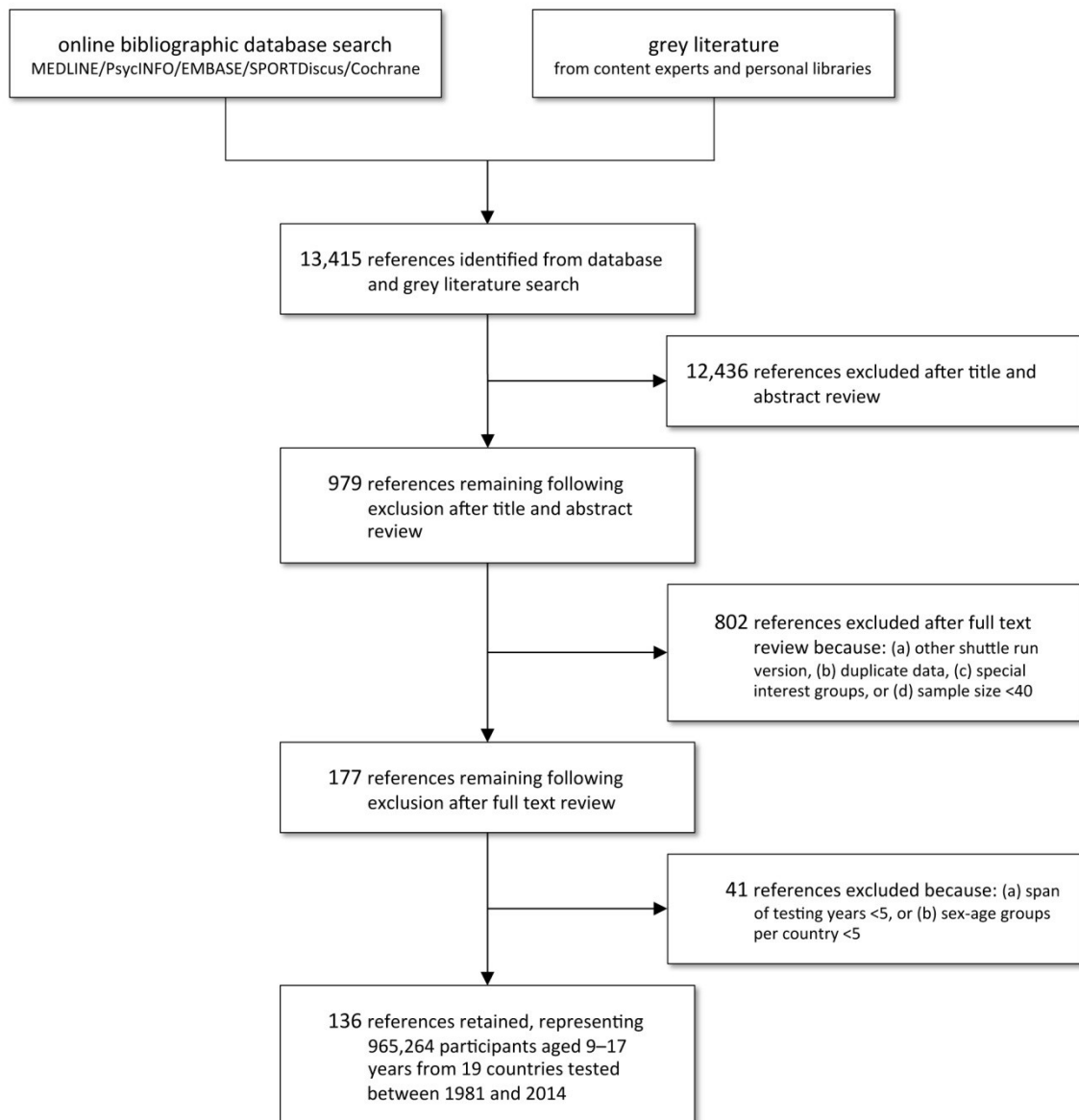


Figure 2a

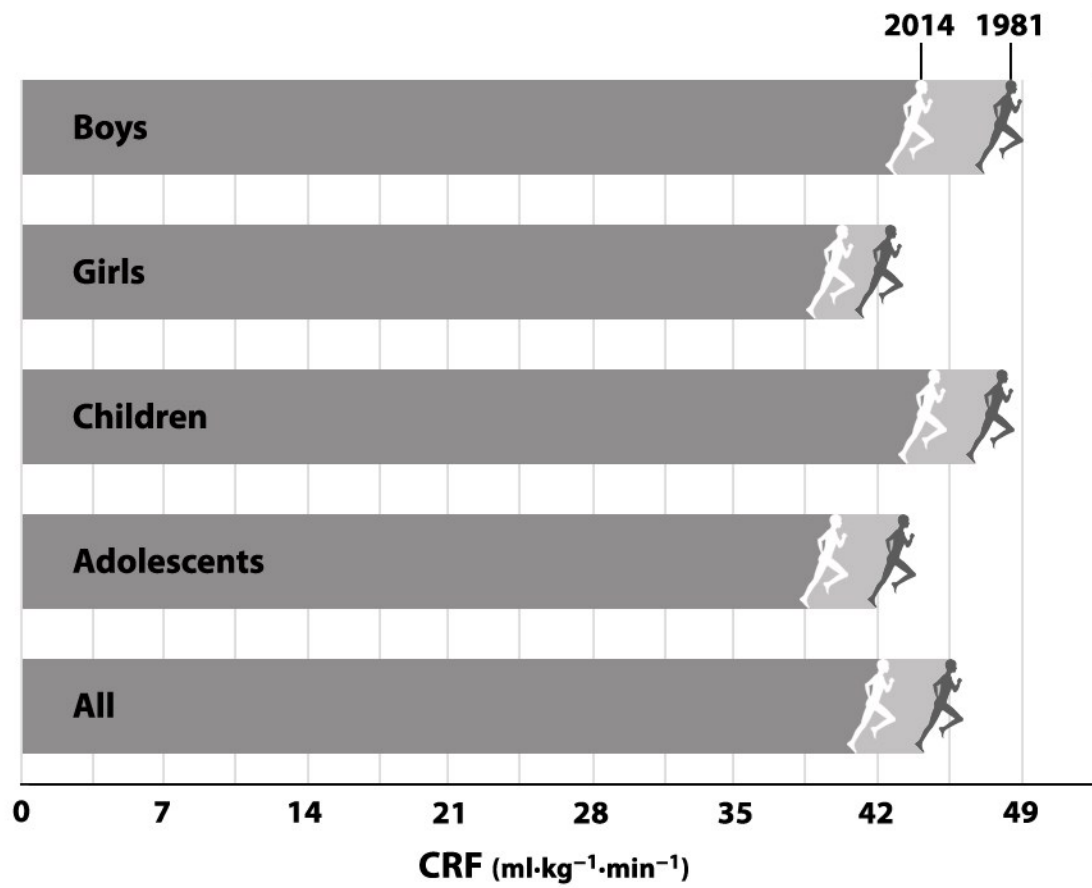


Figure 2b

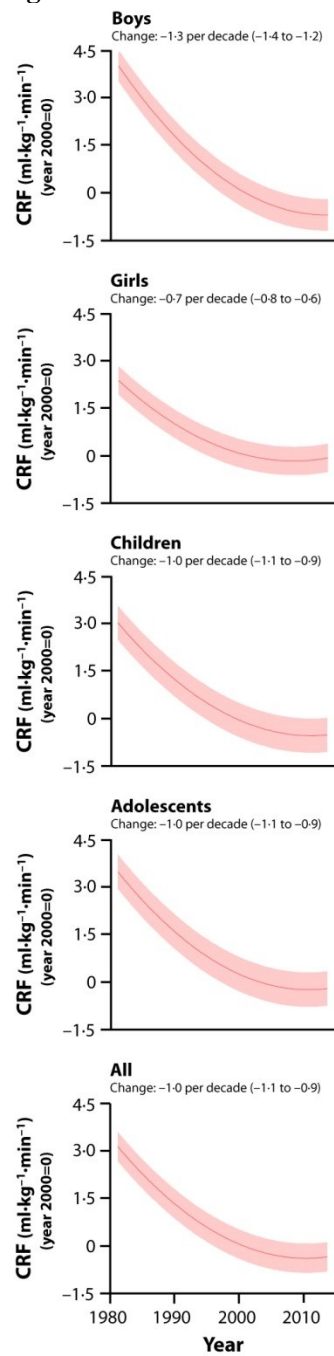
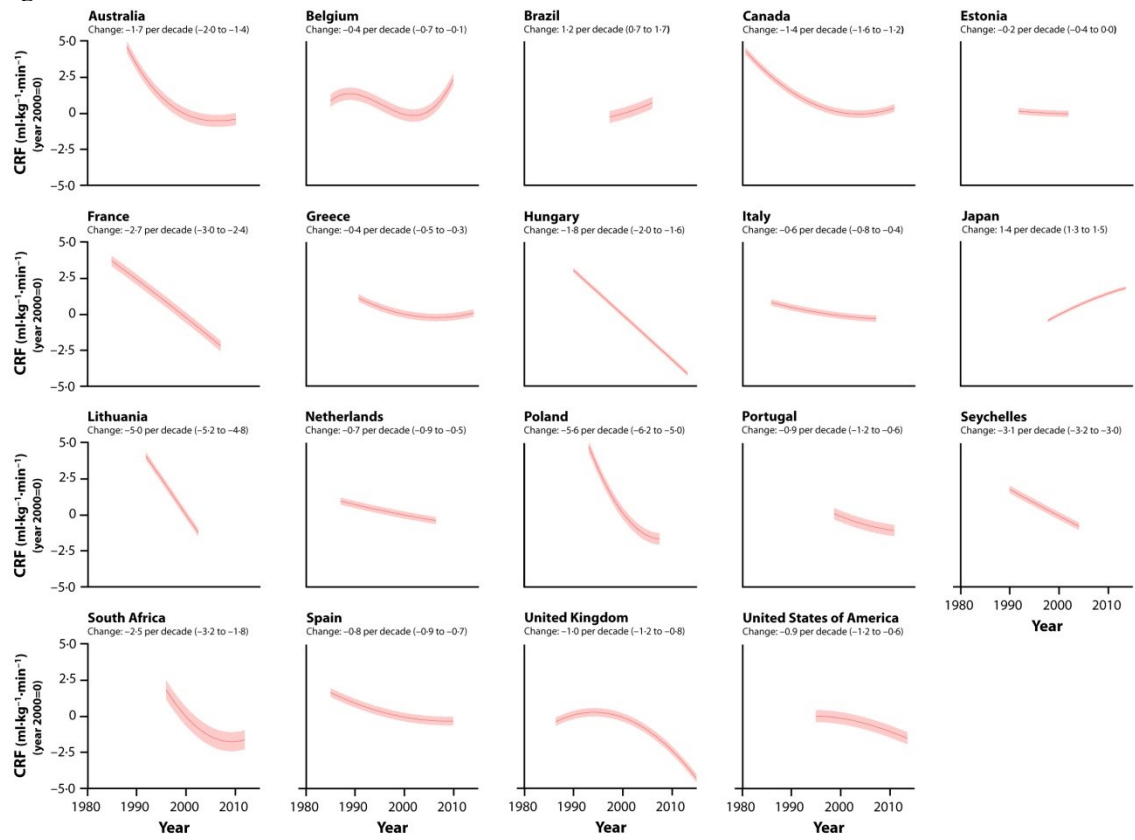


Figure 3



Appendix A: Study Registration

Description: Approved PROSPERO registration for the systematic review found in Chapter 3.

PROSPERO International prospective register of systematic reviews

A systematic review and meta-analysis of the relationships between 20 m shuttle run test performance and health markers in children and youth

Justin Lang, Veronica Poitras, Kevin Belanger, Ian Janssen, Grant Tomkinson, Mark Tremblay

Citation

Justin Lang, Veronica Poitras, Kevin Belanger, Ian Janssen, Grant Tomkinson, Mark Tremblay. A systematic review and meta-analysis of the relationships between 20 m shuttle run test performance and health markers in children and youth. PROSPERO 2015:CRD42015032250 Available from http://www.crd.york.ac.uk/PROSPERO_REBRANDING/display_record.asp?ID=CRD42015032250

Review question(s)

What are the relationships between 20 m shuttle run test (20mSRT) performance and health markers (psychosocial and biophysical) in children and youth aged 5 to 17 years?

Searches

Study design: Any. For longitudinal studies the exposure must be measured on at least one time point within the 5 to 17 years age range;

Sample size (minimum): 40 human participants;

Date limits: Published after 1981;

Language: English language only;

Study type: Peer-reviewed original or review articles only;

Database: MEDLINE (OVID), PsycINFO (OVID), EMBASE (OVID), SPORTDiscus (EBSCO), Cochrane Central Register of Controlled Trials (CENTRAL);

Mean age: Children (child* OR adolescen* OR youth OR boy* OR girl* OR teen* OR paediatric* OR pediatric*) between 5 and 17 years.

The reference lists of included studies will be examined for potentially relevant studies not captured by the search. Content experts will also be contacted to request relevant papers that were not captured in the systematic review of literature.

Types of study to be included

No restrictions will be placed on study design. However, the sample size for all studies must be over 40 participants.

Condition or domain being studied

The 20mSRT is the most popular field-based assessment of cardiorespiratory fitness among children and youth. Although results from lab-based measures of cardiorespiratory fitness in children and youth have remained relatively stable over time, performance on field-based assessments, such as the 20mSRT, have declined at a rate of 0.46% per annum since 1970. Recent reviews report a moderate relationship between lab-based measures of cardiorespiratory fitness and health markers in children and youth. However, no review has attempted to examine the relationships between 20mSRT performance and health markers in children and youth.

Participants/ population

This study will include apparently healthy school-aged children and youth aged 5 to 17 years and/or grades 1 to 12 (or equivalent grades if studies conducted outside of Canada and the US).

Intervention(s), exposure(s)

Studies will be included if they report:

- Main exposure - performance results for the 20mSRT with the 1-minute stage protocol;
- Human participants;
- Mean age between 5 and 17 years old and/or grades 1 to 12;

Comparator(s)/ control

There are no comparators.

Outcome(s)

Primary outcomes

Studies will be included if they report at least one of the following health outcome measurements:

Physiological outcomes:

1. Adiposity measures (i.e., body mass index (BMI), BMI z-score, waist-height ratio, waist circumference, skinfolds, bio-impedance analysis (BIA), dual-energy x-ray absorptiometry (DEXA), computed tomography (CT), magnetic resonance imaging (MRI), densitometry (e.g., underwater weighting, BodPod), etc.);
2. Cardiometabolic biomarkers (i.e., cardiovascular and/or metabolic disease risk factors, lipid profile, insulin resistance, inflammatory biomarkers, blood pressure, etc.);
3. Bone health (i.e., bone mineral density, bone mineral content, bone area, etc.);
4. Musculoskeletal fitness (strength, muscular endurance, flexibility, balance, etc.);

Psychological and cognitive outcomes:

5. Psychosocial health measures (i.e., behaviour conduct (pro-social behaviour, aggression), child social development, self-efficacy, etc.);
6. Cognition (academic achievement (i.e., overall grade point average and/or specific subject grades), concentration, attention, memory, executive function, etc.);
7. Quality of life and well-being;
8. Self-esteem and physical self-perception (i.e., perceived sport competence);
9. Mental health (depressed mood, risk of mental illness, suicidal tendencies, anxiety/stress symptoms, etc.).

Motor skill outcomes:

10. Motor skill development (i.e., gross/fine motor skills, coordination);

Adverse events:

11. Injury (i.e., fractures, acute injury)

Secondary outcomes

None

Data extraction, (selection and coding)

Bibliographic records will be extracted as text files from the Ovid and EBSCO interfaces and imported into Reference Manager Software version 11 for de-duplication. The final database will be exported from Reference Manager

Page removed to protect personal information.

Research Institute

Dr Ian Janssen, School of Kinesiology and Health Studies, and Department of Public Health Sciences, Queen's University

Dr Grant Tomkinson, Department of Kinesiology and Public Health Education, University of North Dakota

Dr Mark Tremblay, Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute

Details of any existing review of the same topic by the same authors

Normative health-related fitness values for the 20 m shuttle run in children and youth. Registration no.

CRD42013003622 – status: ongoing.

The current review application will include a different population in addition to health outcomes which are different from the previous PROSPERO review application.

Anticipated or actual start date

01 February 2016

Anticipated completion date

31 December 2016

Funding sources/sponsors

Healthy Active Living and Obesity Research Group, Children's Hospital of Eastern Ontario Research Institute

Conflicts of interest

None known

Language

English

Country

Canada

Subject index terms status

Subject indexing assigned by CRD

Subject index terms

Adolescent; Child; Exercise Test; Health Status; Humans; Physical Fitness; Running

Stage of review

Ongoing

Date of registration in PROSPERO

14 December 2015

Date of publication of this revision

13 September 2016

Stage of review at time of this submission	Started	Completed
Preliminary searches	Yes	Yes
Piloting of the study selection process	Yes	Yes
Formal screening of search results against eligibility criteria	Yes	Yes
Data extraction	Yes	Yes
Risk of bias (quality) assessment	Yes	Yes
Data analysis	Yes	No

PROSPERO

International prospective register of systematic reviews

The information in this record has been provided by the named contact for this review. CRD has accepted this information in good faith and registered the review in PROSPERO. CRD bears no responsibility or liability for the content of this registration record, any associated files or external websites.

Appendix B: Chapter 3 Supplemental Materials

Description: Supplement 1 is the systematic search strategy. Supplement 2 are the additional references. Supplement 2 is a summary of the 142 studies included in Chapter 3.

Supplement 1

Search strategies

MEDLINE, Embase, PsycINFO, CENTRAL

1. ((Shuttle adj2 (run* or test)) or SRT).tw.
2. ((Beep or Bleep or bip) adj3 test*).tw.
3. (PACER or Progressive Aerobic Cardiovascular Endurance).tw.
4. (FITNESSGRAM or Eurofit test battery).tw.
5. (Leger test or course navette or test Leger navette or test navette).tw.
6. ((Multi-stage or multistage) adj2 (test* or fitness)).tw.
7. Prefit.tw.
8. ((20 m or 20m or 20 metre) adj3 (test* or run)).tw.
9. or/1-8
10. 9 and (boy OR boys OR girl* OR child* OR adolescen* OR juvenil* OR youth* OR teen* OR pediatric* OR paediatric* OR school*).mp.
11. remove duplicates from 10

SportDiscus

(Shuttle N2 run* or shuttle n2 test or SRT) OR (Beep n3 test or Bleep N3 test or bip N3 test*) OR (PACER or Progressive Aerobic Cardiovascular Endurance) OR (FITNESSGRAM or Eurofit test battery or Prefit) OR (Leger test or course navette or test Leger navette or test navette) OR (Multi-stage N2 test* or multistage N2 test* or Multistage N2 fitness or multistage N2 fitness) OR (20 m test* or 20 m run or 20m test* or 20m run or 20 metre test* or 20 metre run)

(health* OR wellness OR well being OR wellbeing OR “quality of life” OR social behav* OR prosocial behav* OR behavio* conduct or conduct disorder* OR aggression OR delinquen* OR academic achievement OR educational achievement OR grade point average OR GPA OR drop out OR dropout OR attention or concentrat* OR psychomotor performance OR motor development OR motor skill* OR motor activit* OR movement OR physical competenc* OR physical* litera* OR postural balance OR depression OR depressive OR depressed OR mood disorder* or anxiety OR distress OR worry OR burnout OR burn out OR psycho-social adjustment or psychosocial adjustment OR self esteem OR self concept OR (bone N2 density) OR (bone N2 mass) OR hypertens* OR high blood pressure OR cholesterol OR hypercholester* OR hyperlipid* OR dyslipid* OR lipids* OR lipoprotie* OR obese OR obesity OR overweight OR body mass index OR BMI OR waist OR adipos* OR fat OR metabolic syndrome OR insulin resistance OR diabet* OR glucose OR HBA1C OR fitness or physical condition* or (Quality N2 life) or QOL or Suicid* or physical self-perception or physical self concept or perceived sport competen*)

(boy OR boys OR girl* OR child* OR schoolchild* OR adolescen* OR juvenil* OR youth* OR teen* OR pediatric* OR paediatric* or school*)

S1 and S2 and S3

Supplement 2

Additional References

- S1. Baranowski T, Chen TA, Mendoza JA, et al. Prospective BMI category change associated with cardiovascular fitness change. *Med Sci Sports Exerc.* 2013; 45(2):294–298.
- S2. Rodrigues LP, Leitaó R, Lopes VP. Physical fitness predicts adiposity longitudinal changes over childhood and adolescence. *J Sci Med Sport.* 2013; 16(2):118–123.
- S3. Brouwer SI, Stolk RP, Liem ET, et al. The role of fitness in the association between fatness and cardiometabolic risk from childhood to adolescence. *Pediatr Diabetes.* 2013;14(1):57–65.
- S4. He QQ, Wong TW, Du L, et al. Physical activity, cardiorespiratory fitness, and obesity among Chinese children. *Prev Med.* 2011; 52(2):109–113.
- S5. Kim J, Must A, Fitzmaurice GM, et al. Relationship of physical fitness to prevalence and incidence of overweight among schoolchildren. *Obes Res.* 2005; 13(7):1246–1254.
- S6. Savva SC, Tornaritis MJ, Kolokotroni O, et al. High cardiorespiratory fitness is inversely associated with incidence of overweight in adolescence: a longitudinal study. *Scand J Med Sci Sports.* 2014; 24(6):982–989.
- S7. Aires L, Mendonça D, Silva G, et al. A 3-year longitudinal analysis of changes in Body Mass Index. *Int J Sports Med.* 2010; 31(2):133–137.
- S8. de Araujo SS, Miguel-dos-Santos R, Silva RJS, et al. Association between body mass index and cardiorespiratory fitness as predictor of health status in schoolchildren. *Revista Andaluza de Medicina del Deporte.* 2015; 8(2):73–78.
- S9. Williams SM, Phelps D, Laurson KR, et al. Fitness knowledge, cardiorespiratory endurance and body composition of high school students. *Biomed Hum Kinetics.* 2013; 5(1):17–21.
- S10. Borrás PA, Vidal J, Ponseti X, et al. Predictors of quality of life in children. *J Hum Sport Exerc.* 2011; 6(4):649–656.
- S11. Silva G, Aires L, Martins C, et al. Cardiorespiratory fitness associates with metabolic risk independent of central adiposity. *Int J Sports Med.* 2013; 34(10):912–916.
- S12. Klincarov I, Aceski A, Stojanovic T. Basic anthropometric and motor differences in 14 years old male pupils with different cardio respiratory endurance level. Proceedings of the Faculty of Physical Education, University of Banja Luka 2010; (2):56–64.
- S13. Colquitt G, Walker A, Langdon JL, et al. Exploring student attitudes toward physical education and implications for policy. *Sport Scientific & Practical Aspects.* 2012; 9(2):5–12.
- S14. Thomas NE, Baker JS, Graham MR, et al. C-reactive protein in schoolchildren and its relation to adiposity, physical activity, aerobic fitness and habitual diet. *Brit J Sports Med.* 2008; 42(5):357–360.
- S15. Geok SK, Lam SK, Fauzee M, et al. Body fat percentage and level of aerobic capacity among students from upper primary school. *Asian Journal of Physical Education & Recreation.* 2009; 15(2):60–64.
- S16. Hussey J, Bell C, Bennett K, et al. Relationship between the intensity of physical activity, inactivity, cardiorespiratory fitness and body composition in 7–10-year-old Dublin children. *Brit J Sports Med.* 2007; 41(5):311–316.
- S17. Aguiar Santos D, Mónica Silva A, Santa-Clara H, et al. Determinant factors of cardiorespiratory fitness in Portuguese adolescents of different ethnicities. *Braz J Kinesiol Hum Perform.* 2011; 13(4):243–249.

- S18. Marshall JD, Bouffard M. The effects of quality daily physical education on movement competency in obese versus nonobese. *Adapt Phys Activ Q*. 1997; 14(3):222–237.
- S19. Raudsepp L, Jurimae T. Physical activity, fitness, and adiposity of prepubertal girls. *Pediatr Exerc Sci*. 1996; 8(3):259–267.
- S20. Grassi GP, Turci M, Sforza C. Aerobic fitness and somatic growth in adolescents: a cross sectional investigation in a high school context. *J Sports Med Phys Fitness*. 2006; 46(3):412–418.
- S21. Lloyd LK, Bishop PA, Walker JL, et al. The influence of body size and composition on FITNESSGRAM® test performance and the adjustment of FITNESSGRAM® test scores for skinfold thickness in youth. *Meas Phys Educ Exerc Sci*. 2003; 7(4):205–226.
- S22. Raudsepp L, Jurimae T. Relationships between somatic variables, physical activity, fitness and fundamental motor skills in prepubertal boys. *Biol Sport*. 1996; 13(4):279–289.
- S23. Saar M, Jurimae T. Influence of anthropometric parameters on motor ability test results in 13–15 year-old-girls. *Pap Antropol*. 2001; 10:260–267.
- S24. Raudsepp L, Jurimae T. Physique and physical fitness in prepubertal children. *Acta Kinesiologiae Universitatis Tartuensis*. 1996; 1:49–58.
- S25. Agha-Alinejad H, Farzad B, Salari M, et al. Prevalence of overweight and obesity among Iranian preschoolers: Interrelationship with physical fitness. *J Res Med Sci*. 2015; 20(4):334–341.
- S26. Carrel AL, Sledge JS, Ventura SJ, et al. GPS suggests low physical activity in urban Hispanic school children: a proof of concept study. *Int J Pediatr Endocrinol*. 2014; 2014(1):25.
- S27. Bovet P, Auguste R, Burdette H. Strong inverse association between physical fitness and overweight in adolescents: a large school-based survey. *Int J Behav Nutr Phys Act*. 2007; 4:24.
- S28. Albarwani S, Phil D, Al Hashmi K, et al. Effects of overweight and leisure-time activities on aerobic fitness in urban and rural adolescents. *Metab*. 2008; 7(4):369–374.
- S29. Palomaki S, Heikinaro-Johansson P, Huotari P. Cardiorespiratory performance and physical activity in normal weight and overweight Finnish adolescents from 2003 to 2010. *J Sports Sci*. 2015; 33(6):588–596.
- S30. Pienaar C, Coetzee B, Monyeki AM. The use of anthropometric measurements and the influence of demographic factors on the prediction of VO₂(max) in a cohort of adolescents: the PAHL study. *Ann Hum Biol*. 2015; 42(2):134–142.
- S31. Arango CM, Parra DC, Gomez LF, et al. Screen time, cardiorespiratory fitness and adiposity among school-age children from Monteria, Colombia. *J Sci Med Sport*. 2014; 17(5):491–495.
- S32. Andrade S, Ochoa-Aviles A, Lachat C, et al. Physical fitness among urban and rural Ecuadorian adolescents and its association with blood lipids: a cross sectional study. *BMC Pediatr*. 2014; 14:106.
- S33. Scudder MR, Lambourne K, Drollette ES, et al. Aerobic capacity and cognitive control in elementary school-age children. *Med Sci Sports Exerc*. 2014; 46(5):1025–1035.
- S34. Diez-Fernandez A, Sanchez-Lopez M, Mora-Rodriguez R, et al. Obesity as a mediator of the influence of cardiorespiratory fitness on cardiometabolic risk: a mediation analysis. *Diabetes Care*. 2014; 37(3):855–862.
- S35. Tambalis KD, Panagiotakos DB, Arnaoutis G, et al. Endurance, explosive power, and muscle strength in relation to body mass index and physical fitness in Greek children aged 7–10 years. *Pediatr Exerc Sci*. 2013; 25(3):394–406.
- S36. Buchan DS, Young JD, Boddy LM, et al. Fitness and adiposity are independently associated with cardiometabolic risk in youth. *Biomed Res Int*. 2013; 2013:261698.

- S37. Coe DP, Peterson T, Blair C, et al. Physical fitness, academic achievement, and socioeconomic status in school-aged youth. *J Sch Health*. 2013; 83(7):500–507.
- S38. Jimenez-Pavon D, Ruiz JR, Ortega FB, et al. Physical activity and markers of insulin resistance in adolescents: role of cardiorespiratory fitness levels--the HELENA study. *Pediatr Diabetes*. 2013; 14(4):249–258.
- S39. Burns R, Hannon JC, Brusseau TA, et al. Indices of abdominal adiposity and cardiorespiratory fitness test performance in middle-school students. *J Obes*. 2013; 2013:912460.
- S40. Martins C, Freitas I, Jr., Pizarro A, et al. Cardiorespiratory fitness, but not central obesity or C-reactive protein, is related to liver function in obese children. *Pediatr Exerc Sci*. 2013; 25(1):3–11.
- S41. Gonzalez-Suarez CB, Caralipio N, Gambito E, et al. The association of physical fitness with body mass index and waist circumference in Filipino preadolescents. *Asia Pac J Public Health*. 2013; 25(1):74–83.
- S42. Coe DP, Pivarnik JM, Womack CJ, et al. Health-related fitness and academic achievement in middle school students. *J Sports Med Phys Fitness*. 2012; 52(6):654–660.
- S43. Lewitt MS, Baker JS, Mooney GP, et al. Pubertal stage and measures of adiposity in British schoolchildren. *Ann Hum Biol*. 2012; 39(5):440–447.
- S44. Galaviz KI, Tremblay MS, Colley R, et al. Associations between physical activity, cardiorespiratory fitness, and obesity in Mexican children. *Salud Publica Mex*. 2012; 54(5):463–469.
- S45. Musa DI, Williams CA. Cardiorespiratory fitness, fatness, and blood pressure associations in Nigerian youth. *Med Sci Sports Exerc*. 2012; 44(10):1978–1985.
- S46. Llorente-Cantarero FJ, Perez-Navero JL, de DB-S, et al. Non-traditional markers of metabolic risk in prepubertal children with different levels of cardiorespiratory fitness. *Public Health Nutr*. 2012; 15(10):1827–1834.
- S47. Duncan MJ, Mota J, Vale S, et al. Comparisons between inverted body mass index and body mass index as proxies for body fatness and risk factors for metabolic risk and cardiorespiratory fitness in Portuguese adolescents. *Am J Human Biol*. 2012; 24(5):618–625.
- S48. Machado-Rodrigues AM, Coelho-E-Silva MJ, Mota J, et al. Cardiorespiratory fitness, weight status and objectively measured sedentary behaviour and physical activity in rural and urban Portuguese adolescents. *J Child Health Care*. 2012; 16(2):166–177.
- S49. Joshi P, Bryan C, Howat H. Relationship of body mass index and fitness levels among schoolchildren. *J Strength Cond Res*. 2012; 26(4):1006–1014.
- S50. Ebenegger V, Marques-Vidal P, Kriemler S, et al. Differences in aerobic fitness and lifestyle characteristics in preschoolers according to their weight status and sports club participation. *Obes Facts*. 2012; 5(1):23–33.
- S51. Niederer I, Kriemler S, Zahner L, et al. BMI group-related differences in physical fitness and physical activity in preschool-age children: a cross-sectional analysis. *Res Q Exercise Sport*. 2012; 83(1):12–19.
- S52. 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*. 2011; 12(8):704–712.
- S53. Aires L, Pratt M, Lobelo F, et al. Associations of cardiorespiratory fitness in children and adolescents with physical activity, active commuting to school, and screen time. *J Phys Act Health*. 2011; 8(Suppl 2):S198–S205.

- S54. Castro-Pineiro J, Ortega FB, Keating XD, et al. Percentile values for aerobic performance running/walking field tests in children aged 6 to 17 years: influence of weight status. *Nutr Hosp*. 2011; 26(3):572–578.
- S55. Gonzalez-Suarez CB, Grimmer-Somers K. The association of physical activity and physical fitness with pre-adolescent obesity: an observational study in Metromanila, Philippines. *J Phys Act Health*. 2011; 8(6):804–810.
- S56. Jimenez-Pavon D, Castillo MJ, Moreno LA, et al. Fitness and fatness are independently associated with markers of insulin resistance in European adolescents; the HELENA study. *Int J Pediatr Obes*. 2011; 6(3–4):253–260.
- S57. Moliner-Urdiales D, Ruiz JR, Vicente-Rodriguez G, et al. Associations of muscular and cardiorespiratory fitness with total and central body fat in adolescents: the HELENA study. *Brit J Sports Med*. 2011; 45(2):101–108.
- S58. Pino-Ortega J, De la Cruz-Sánchez E, Martinez-Santos R. Health-related fitness in school children: compliance with physical activity recommendations and its relationship with body mass index and diet quality. *Arch Latinoam Nutr*. 2010; 60(4):374–379.
- S59. Marques-Vidal P, Marcelino G, Ravasco P, et al. Increased body fat is independently and negatively related with cardiorespiratory fitness levels in children and adolescents with normal weight. *Eur J Cardiovasc Prev Rehabil*. 2010; 17(6):649–654.
- S60. Ostojic SM, Stojanovic MD. High aerobic fitness is associated with lower total and regional adiposity in 12-year-old overweight boys. *J Sports Med Phys Fitness*. 2010; 50(4):443–449.
- S61. Ostojic SM, O'Neil M, Calleja J, et al. Cardiorespiratory fitness and adiposity in overweight and nonoverweight 8-year-old school children. *Minerva Pediatr*. 2010; 62(6):537–543.
- S62. Houwen S, Hartman E, Visscher C. The relationship among motor proficiency, physical fitness, and body composition in children with and without visual impairments. *Res Q Exerc Sport*. 2010; 81(3):290–299.
- S63. Martinez-Gomez D, Eisenmann JC, Warnberg J, et al. Associations of physical activity, cardiorespiratory fitness and fatness with low-grade inflammation in adolescents: the AFINOS Study. *Int J Obes (Lond)*. 2010; 34(10):1501–1507.
- S64. Jago R, Drews KL, McMurray RG, et al. Fatness, fitness, and cardiometabolic risk factors among sixth-grade youth. *Med Sci Sports Exerc*. 2010; 42(8):1502–1510.
- S65. Vale S, Santos R, Soares-Miranda L, et al. The relationship of cardiorespiratory fitness, birth weight and parental BMI on adolescents' obesity status. *Eur J Clin Nutr*. 2010; 64(6):622–627.
- S66. Aires L, Silva P, Silva G, et al. Intensity of physical activity, cardiorespiratory fitness, and body mass index in youth. *J Phys Act Health*. 2010; 7(1):54–59.
- S67. Martinez-Gomez D, Eisenmann JC, Moya JM, et al. The role of physical activity and fitness on the metabolic syndrome in adolescents: effect of different scores. The AFINOS Study. *J Physiol Biochem*. 2009; 65(3):277–289.
- S68. Stigman S, Rintala P, Kukkonen-Harjula K, et al. Eight-year-old children with high cardiorespiratory fitness have lower overall and abdominal fatness. *Int J Pediatr Obes*. 2009; 4(2):98–105.
- S69. Sakuragi S, Abhayaratna K, Gravenmaker KJ, et al. Influence of adiposity and physical activity on arterial stiffness in healthy children: the lifestyle of our kids study. *Hypertension*. 2009; 53(4):611–616.
- S70. Ortega FB, Labayen I, Ruiz JR, et al. Are muscular and cardiovascular fitness partially programmed at birth? Role of body composition. *J Pediatr*. 2009; 154(1):61–66.

- S71. Hoekstra T, Boreham CA, Murray LJ, et al. Associations between aerobic and muscular fitness and cardiovascular disease risk: the Northern Ireland young hearts study. *J Phys Act Health*. 2008; 5(6):815–829.
- S72. Downs SM, Marshall D, Ng C, et al. Central adiposity and associated lifestyle factors in Cree children. *Appl Physiol Nutr Metab*. 2008; 33(3):476–482.
- S73. Cairney J, Hay JA, Faught BE, et al. Generalized self-efficacy and performance on the 20-metre shuttle run in children. *Am J Human Biol*. 2008; 20(2):132–138.
- S74. Buck SM, Hillman CH, Castelli DM. The relation of aerobic fitness to stroop task performance in preadolescent children. *Med Sci Sports Exerc*. 2008; 40(1):166–172.
- S75. 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.
- S76. Burke V, Beilin LJ, Durkin K, et al. Television, computer use, physical activity, diet and fatness in Australian adolescents. *Int J Pediatr Obes*. 2006; 1(4):248–255.
- S77. Ng C, Marshall D, Willows ND. Obesity, adiposity, physical fitness and activity levels in Cree children. *Int J Circumpolar Health*. 2006; 65(4):322–330.
- S78. Tokmakidis SP, Kasambalis A, Christodoulos AD. Fitness levels of Greek primary schoolchildren in relationship to overweight and obesity. *Eur J Pediatr*. 2006; 165(12):867–874.
- S79. Nassis GP, Psarra G, Sidossis LS. Central and total adiposity are lower in overweight and obese children with high cardiorespiratory fitness. *Eur J Clin Nutr*. 2005; 59(1):137–141.
- S80. Gonzalez-Gross M, Ruiz JR, Moreno LA, et al. Body composition and physical performance of Spanish adolescents: the AVENA pilot study. *Acta Diabetol*. 2003; 40(Suppl 1):S299–S301.
- S81. Cieslack TJ, Frost G, Klentrou P. Effects of physical activity, body fat, and salivary cortisol on mucosal immunity in children. *J Appl Physiol*. 2003; 95(6):2315–2320.
- S82. Deforche B, Lefevre J, De B, I, et al. Physical fitness and physical activity in obese and nonobese Flemish youth. *Obes Res*. 2003; 11(3):434–441.
- S83. Mota J, Guerra S, Leandro C, et al. Association of maturation, sex, and body fat in cardiorespiratory fitness. *Am J Human Biol*. 2002; 14(6):707–712.
- S84. Coledam DHC, Batista Junior JP, Glaner MF. Low agreement between the fitnessgram criterion references for adolescents. *Rev Paul Pediatr*. 2015; 33(2):181–186.
- S85. Saavedra JM, Escalante Y, Dominguez AM, et al. Prediction of correlates of daily physical activity in Spanish children aged 8–9 years. *Scand J Med Sci Sports*. 2014; 24(3):e213–e219.
- S86. Coelho-E-Silva, Vaz Ronque ER, Cyrino ES, et al. Nutritional status, biological maturation and cardiorespiratory fitness in Azorean youth aged 11–15 years. *BMC Public Health*. 2013; 13(pp 495).
- S87. Vale S, Soares-Miranda L, Santos R, et al. Influence of cardiorespiratory fitness and parental lifestyle on adolescents' abdominal obesity. *Ann Hum Biol*. 2011; 38(5):531–536.
- S88. Artero EG, Espana-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.
- S89. Fogelholm M, Stigman S, Huisman T, et al. Physical fitness in adolescents with normal weight and overweight. *Scand J Med Sci Sports*. 2008; 18(2):162–170.
- S90. Puder JJ, Schindler C, Zahner L, et al. Adiposity, fitness and metabolic risk in children: a cross-sectional and longitudinal study. *Int J Pediatr Obes*. 2011; 6(2–2):e297–e306.

- S91. Telford RD, Cunningham RB, Shaw JE, et al. Contrasting longitudinal and cross-sectional relationships between insulin resistance and percentage of body fat, fitness, and physical activity in children-the LOOK study. *Pediatr Diabetes*. 2009; 10(8):500–507.
- S92. Telford RD, Cunningham RB, Telford RM, et al. Effects of changes in adiposity and physical activity on preadolescent insulin resistance: the Australian LOOK longitudinal study. *PLoS ONE*. 2012; 7(10):e47438.
- S93. Gostinis Sobrinho CA, Moreira CM, Mota JA, et al. C-reactive protein, physical activity and cardiorespiratory fitness in Portuguese adolescents: a cross-sectional study. *Cad Saude Publica*. 2015; 31(9):1907–1915.
- S94. Buchan DS, Boddy LM, Young JD, et al. Relationships between cardiorespiratory and muscular fitness with cardiometabolic risk in adolescents. *Res Sports Med*. 2015; 23(3):227–239.
- S95. Curtis VA, Carrel AL, Eickhoff JC, et al. Gender and race influence metabolic benefits of fitness in children: a cross-sectional study. *Int J Pediatr Endocrinol*. 2012; 2012(1):4.
- S96. Varness T, Carrel AL, Eickhoff JC, et al. Reliable prediction of insulin resistance by a school-based fitness test in middle-school children. *Int J Pediatr Endocrinol*. 2009; 2009:487804.
- S97. Benser J, Valtuena J, Ruiz JR, et al. Impact of physical activity and cardiovascular fitness on total homocysteine concentrations in European adolescents: The HELENA study. *J Nutr Sci Vitaminol (Tokyo)*. 2015; 61(1):45–54.
- S98. Machado-Rodrigues AM, Leite N, Coelho-E-Silva MJ, et al. Relationship between metabolic syndrome and moderate-to-vigorous physical activity in youth. *J Phys Act Health*. 2015; 12(1):13–19.
- S99. Artero EG, Espana-Romero V, Jimenez-Pavon D, et al. Muscular fitness, fatness and inflammatory biomarkers in adolescents. *Pediatr Obes*. 2014; 9(5):391–400.
- S100. Machado-Rodrigues AM, Leite N, Coelho-E-Silva MJ, et al. Independent association of clustered risk factors with cardiorespiratory fitness in youth aged 11–17 years. *Ann Hum Biol*. 2014; 41(3):271–276.
- S101. Buchan DS, Young JD, Boddy LM, et al. Independent associations between cardiorespiratory fitness, waist circumference, BMI, and clustered cardiometabolic risk in adolescents. *Am J Human Biol*. 2014; 26(1):29–35.
- S102. Ogunleye AA, Sandercock GR, Voss C, et al. Prevalence of elevated mean arterial pressure and how fitness moderates its association with BMI in youth.[Erratum appears in Public Health Nutr. 2015 Jun;18(8):1522; PMID: 25307851]. *Public Health Nutr*. 2013; 16(11):2046–2054.
- S103. Moreira C, Santos R, Moreira P, et al. Cardiorespiratory fitness is negatively associated with metabolic risk factors independently of the adherence to a healthy dietary pattern. *Nutr Metab Cardiovasc Dis*. 2013; 23(7):670–676.
- S104. Marcelino G, Melich-Cerveira J, Paccaud F, et al. Obese and fit adolescents have lower blood pressure levels than obese and unfit counterparts. *J Sports Med Phys Fitness*. 2012; 52(6):639–646.
- S105. Naidoo T, Konkol K, Biccand B, et al. Elevated salivary C-reactive protein predicted by low cardio-respiratory fitness and being overweight in African children. *Cardiovasc*. 2012; 23(9):501–506.
- S106. Martinez-Gomez D, Eisenmann JC, Gomez-Martinez S, et al. Associations of physical activity and fitness with adipocytokines in adolescents: the AFINOS Study. *Nutr Metab Cardiovasc Dis*. 2012; 22(3):252–259.
- S107. Neto SA, Sasaki JE, Mascarenhas LP, et al. Physical activity, cardiorespiratory fitness, and metabolic syndrome in adolescents: a cross-sectional study. *BMC Public Health*. 2011; 11:674.

- S108. 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 (Engl)*. 2007; 60(6):581–588.
- S109. Ruiz JR, Sola R, Gonzalez-Gross M, et al. Cardiovascular fitness is negatively associated with homocysteine levels in female adolescents. *Arch Pediatr Adolesc Med*. 2007; 161(2):166–171.
- S110. Nielsen GA, Andersen LB. The association between high blood pressure, physical fitness, and body mass index in adolescents. *Prev Med*. 2003; 36(2):229–234.
- S111. Mesa JL, Ruiz JR, Ortega FB, et al. Aerobic physical fitness in relation to blood lipids and fasting glycaemia in adolescents: influence of weight status. *Nutr Metab Cardiovasc Dis*. 2006; 16(4):285–293.
- S112. Guerra S, Ribeiro JC, Costa R, et al. Relationship between cardiorespiratory fitness, body composition and blood pressure in school children. *J Sports Med Phys Fitness*. 2002; 42(2):207–213.
- S113. Boreham C, Twisk J, Murray L, et al. Fitness, fatness, and coronary heart disease risk in adolescents: the Northern Ireland Young Hearts Project. *Med Sci Sports Exerc*. 2001; 33(2):270–274.
- S114. Niederer I, Kriemler S, Gut J, et al. Relationship of aerobic fitness and motor skills with memory and attention in preschoolers (Ballabeina): a cross-sectional and longitudinal study. *BMC Pediatr*. 2011; 11:34.
- S115. du Toit D, Pienaar AE, Truter L. Relationship between physical fitness and academic performance in South African children. *S Afr J Res Sport Phys Educ Recreation*. 2011; 33(3):23–35.
- S116. Joshi P, Howat H, Bryan C. Relationship between fitness levels and academic performance. *J Phys Educ Sport*. 2011; 11(44):376–382.
- S117. Wittberg R, Cottrell LA, Davis CL, et al. Aerobic fitness thresholds associated with fifth grade academic achievement. *J Health Educ*. 2010; 41(5):284–291.
- S118. Hansen DM, Herrmann SD, Lambourne K, et al. Linear/nonlinear relations of activity and fitness with children’s academic achievement. *Med Sci Sports Exerc*. 2014; 46(12):2279–2285.
- S119. Sardinha LB, Marques A, Martins S, et al. Fitness, fatness, and academic performance in seventh-grade elementary school students. *BMC Pediatr*. 2014; 14:176.
- S120. Rauner RR, Walters RW, Avery M, et al. Evidence that aerobic fitness is more salient than weight status in predicting standardized math and reading outcomes in fourth- through eighth-grade students. *J Pediatr*. 2013; 163(2):344–348.
- S121. Bass RW, Brown DD, Laurson KR, et al. Physical fitness and academic performance in middle school students. *Acta Paediatr*. 2013; 102(8):832–837.
- S122. Telford RD, Cunningham RB, Telford RM, et al. Schools with fitter children achieve better literacy and numeracy results: evidence of a school cultural effect.[Erratum appears in *Pediatr Exerc Sci*. 2012 Nov;24(4):ii Note: Abharatna, Walter P [corrected to Abhayaratna, Walter P]]. *Pediatr Exerc Sci*. 2012; 24(1):45–57.
- S123. Padilla-Moledo C, Castro-Pinero J, Ortega FB, et al. Positive health, cardiorespiratory fitness and fatness in children and adolescents. *Eur J Public Health*. 2012; 22(1):52–56.
- S124. Rieck T, Jackson A, Martin SB, et al. Health-related fitness, body mass index, and risk of depression among adolescents. *Med Sci Sports Exerc*. 2013; 45(6):1083–1088.
- S125. Greenleaf CA, Petrie TA, Martin SB. Psychosocial variables associated with body composition and cardiorespiratory fitness in middle school students. *Res Q Exerc Sport*. 2010; 81(3 Suppl):S65–S74.

- S126. Zhu X, Chen S. The impacts of adolescent beliefs in performing a cardiorespiratory fitness test. *Int J Sport Exerc Psychol*. 2015; 13(2):182–192.
- S127. Gao Z, Xiang P, Lochbaum M, et al. The impact of achievement goals on cardiorespiratory fitness: does self-efficacy make a difference? *Res Q Exerc Sport*. 2013; 84(3):313–322.
- S128. Goa Z, Lodewyk KR, Zhang T. The role of ability beliefs and incentives in middle school students' intention, cardiovascular fitness, and effort. *Journal of teaching in physical education*. 2009; 28:3–20.
- S129. Morales PF, Sanchez-Lopez M, Moya-Martinez P, et al. Health-related quality of life, obesity, and fitness in schoolchildren: the Cuenca study. *Qual Life Res*. 2013; 22(7):1515–1523.
- S130. Olive LS, Byrne DG, Cunningham RB, et al. Effects of physical activity, fitness and fatness on children's body image: The Australian LOOK longitudinal study. *Ment Health Phys Act*. 2012; 5(2):116–124.
- S131. Gao Z. Perceived competence and enjoyment in predicting students' physical activity and cardiorespiratory fitness. *Percept Mot Skills*. 2008; 107(2):365–372.
- S132. Jurimae T, Saar M. Self-perceived and actual indicators of motor abilities in children and adolescents. *Percept Mot Skills*. 2003; 97(3 Pt 1):862–866.
- S133. Riiser K, Ommundsen Y, Smastuen MC, et al. The relationship between fitness and health-related quality of life and the mediating role of self-determined motivation in overweight adolescents. *Scand J Public Health*. 2014; 42(8):766–772.
- S134. Shen B, McCaughtry N, Martin J. The influence of self-determination in physical education on leisure-time physical activity behavior. *Res Q Exerc Sport*. 2007; 78(4):328–338.
- S135. Castro-Pinero J, Padilla-Moledo C, Ortega FB, et al. Cardiorespiratory fitness and fatness are associated with health complaints and health risk behaviors in youth. *J Phys Act Health*. 2012; 9(5):642–649.
- S136. Mota J, Santos RM, Silva P, et al. Associations between self-rated health with cardiorespiratory fitness and obesity status among adolescent girls. *J Phys Act Health*. 2012; 9(3):378–381.
- S137. Vincente-Rodriguez G, Urzanqui A, Mesana MI, et al. Physical fitness effect on bone mass is mediated by the independent association between lean mass and bone mass through adolescence: a cross-sectional study. *J Bone Miner Metab*. 2008; 26(3):288–294.
- S138. Reeves L, Broeder CE, Kennedy-Honeycutt L, et al. Relationship of fitness and gross motor skills for five- to six-yr.-old children. *Percept Mot Skills*. 1999; 89(3 Pt 1):739–747.
- S139. Burns RD, Brusseau TA, Fu Y, et al. Predictors and trends of gross motor skill performance in at-risk elementary school-aged children. *Percept Mot Skills*. 2015; 121(1):284–299.
- S140. Okely AD, Booth ML, Patterson JW. Relationship of cardiorespiratory endurance to fundamental movement skill proficiency among adolescents. *Pediatr Exerc Sci*. 2001; 13(4):380–391.
- S141. Martin-Diener E, Wanner M, Kriemler S, et al. Associations of objectively assessed levels of physical activity, aerobic fitness and motor coordination with injury risk in school children aged 7–9 years: a cross-sectional study. *BMJ Open*. 2013; 3(8).
- S142. Clark EM, Tobias JH, Murray L, et al. Children with low muscle strength are at an increased risk of fracture with exposure to exercise. *J Musculoskelet Neuronal Interact*. 2011; 11(2):196–202.

Supplement 3

Summary of included studies

Supplement Table 1. 20mSRT performance and adiposity in children and youth

Note: β indicates standardized Beta; B indicates unstandardized Beta

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1368 Baranowski (2012); USA	Longitudinal (1 yr follow- up)	n = 3998 T1: Grade 6 T2: Grade 8 Baseline age = 11.3 yr	VO2max Leger	Weight status (CDC-cut points) determined from BMI (calculated from objectively measured height and weight)	Adiposity	Boys that remained obese at follow-up had lower 20mSRT performance at baseline compared to all other groups $\beta = 2.79$, SE = 0.67 Girls that remained overweight or at healthy BMI at follow-up had higher baseline 20mSRT performance than all groups, except for the increased BMI group OW $\beta = 3.54$, SE = 0.62 Healthy $\beta = 3.71$, SE = 0.59	Baseline 20mSRT Baseline BMI
1379 Rodrigues (2013); Portugal	Longitudinal (9 yr follow- up)	n = 518 T1 = 6 yr T2 = 15 yr	n/a	SSF (mm) were calculated at two sites (subscapular, triceps).	Adiposity	A positive change in 20mSRT performance predicted positive changes in adiposity over time $\beta = -0.08$, SE = 0.01, $p < 0.01$	SSF Age Sex Flexed arm hang Standing long jump Sit-ups 50m dash

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1387 Brouwer (2013); Netherlands	Longitudinal (5 yr follow- up)	n = 565 range = 11-16 yr T1 age = 11.4 yr T2 age = 16.1 yr	VO2max Leger	WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance at baseline was negatively associated with WC at follow-up. Boys $r = -0.289$, $p < 0.01$ Girls $r = -0.229$, $p < 0.01$	n/a
1617 He (2011); China	Longitudinal (18 month follow-up)	n = 1795 range = 8-10 yr age = 10.2 yr	VO2max Leger	BMI was calculated from measured height and weight	Adiposity	20mSRT performance at baseline was inversely associated with changes in BMI at T2 $p < 0.001$ Boys with low 20mSRT performance at T1 had higher odds of being ow/obese at T2 OR 8.71, 95 % CI 2.59-29.26 No significant odds were observed in girls at follow-up OR 6.87, 95 % CI 0.96-49.09	District School Age PA level BMI at T1 20mSRT difference between T1-T2
1992 Kim (2005); USA	Longitudinal (1 yr follow- up)	n = 6297 range = 5-14 yr	PACER HFZ	BMI was calculated from measured height and weight	Adiposity	Not attaining the HFZ at baseline was associated with incidence of overweight at follow-up in girls but not in boys Girls OR 1.96, 95 % CI 1.10- 3.50 Boys OR 0.98, 95 % CI 0.58- 1.67	Age group Race Lunch program Baseline BMI
2519 Savva (2014); Cyprus	Longitudinal	n = 4878 T1 range = 10-13.5 yr T2 range = 14.4- 17.9 yr	VO2max Leger	BMI was calculated from measured height and weight	Adiposity	Boys and girls in the highest quartile of 20mSRT performance at baseline had the lowest odds of becoming ow/obese at follow- up	Baseline age Follow-up Living area Air temp

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
		T1 age = 11.4 yr T2 age = 16.1 yr				Boys OR 0.40, 95 % CI 0.26-0.61 Girls OR 0.57, 95 % CI 0.33-0.99 Boys $p < 0.001$; Girls $p < 0.05$	
3150 Aires (2010); Portugal	Longitudinal (3.6 to 6.1 yr follow- up)	n = 345 range = 11-16 yr Baseline age = 11.4 yr Follow-up age = 16.1 yr	PACER HFZ	BMI was calculated from measured height and weight	Adiposity	Low 20mSRT performance at baseline was positively associated with BMI at follow-up $\beta = 0.76$, 95 % CI 0.29-1.24, $p = 0.002$	Age Sex
5 de Araujo (2015); Brazil	Cross- sectional	n = 288 range = 9.5-14.5 yr age = 12.2 yr	VO2max Leger	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was inversely associated with BMI for boys and girls Boys $R^2 = 0.85$, $p = 0.03$ Girls $R^2 = 0.91$, $p = 0.01$	n/a
211 Williams (2013); USA	Cross- sectional	n = 171 range = 17-18 yr age = 17.3 yr	VO2max Fitnessgram	BF% was measured using BIA	Adiposity	20mSRT performance was negatively related with BF% $r = -0.44$, $p < 0.001$	Age Sex
220 Borras (2011); Spain	Cross- sectional	n = 302 range = 11-12 yr Boy age = 11.6 yr Girl age = 10.7 yr	Decimal stages	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related with BMI $r = -0.241$, $p < 0.001$	n/a
238 Silva (2013); Portugal	Cross- sectional	n = 633 range = 10-18 yr age = 14.3 yr	VO2max Fitnessgram	WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was negatively related to WC $r = -0.50$, $p < 0.001$ WC was lower in the fit group compared to the low fit group	Age Sex

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						$F = 212.80, p < 0.001$	
292 Klincarov (2011); Macedonia	Cross- sectional	n = 156 range = 14 yr age = 14.0 yr	Stages	Triceps skinfold thickness (mm) was measured according to standardized procedures BMI was calculated from measured height and weight	Adiposity	Triceps skinfold was different between low, medium and high 20mSRT performance groups $F = 39.6, p < 0.0001$ BMI was different between low, medium and high 20mSRT performance groups $F = 14.5, p < 0.0001$	n/a
298 Colquitt (2012); USA	Cross- sectional	n = 122 range = 11-15 yr age = 12.8 yr	VO2max Fitnessgram	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI $r = -0.43, p < 0.01$	n/a
339 Thomas (2008); Wales	Cross- sectional	n = 164 range = 12-13 yr age = 12.9 yr	Shuttles	BMI was calculated from measured height and weight WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was negatively related to BMI Boys $r = -0.63, p < 0.01$ Girls $r = -0.52, p < 0.01$	n/a
396 Geok (2015); Malaysia	Cross- sectional	n = 240 range = 10-11 yr	VO2max Leger	BF% was calculated from measured triceps and calf skinfold thickness	Adiposity	20mSRT performance was negatively related to BF% $r = -0.28, p < 0.01$	n/a
446 Hussey (2007); Ireland	Cross- sectional	n = 224 range = 7-10 yr	VO2max Leger	BMI z-score calculated from measured height and weight using British norms WC (cm) was measured	Adiposity	20mSRT performance was negatively related to BMI z-score Boys $r = -0.43, p < 0.01$ Girls $r = -0.22, p < 0.01$ 20mSRT performance was	Age Sex WC BMI

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				according to standardized procedures		negatively related to WC Boys $r = -0.50$, $p < 0.01$ Girls $r = -0.33$, $p < 0.01$ 20mSRT performance was a negative predictor of BMI z- score $B = -1.12$, $SE = 0.2$, $p < 0.01$ 20mSRT performance was a positive predictor of WC $B = 0.22$, $SE = 0.04$, $p < 0.01$	
470 Aguiar Santos (2011); Portugal	Cross- sectional	$n = 266$ range = 12-18 yr	VO2max Fitnessgram	BF% was measured using a handheld BIA	Adiposity	20mSRT performance was negatively associated with BF% $\beta = -0.003$, $SE = 0.00$, $p < 0.001$	Age Ethnicity Sex
534 Marshall (1996); Canada	Randomized Factorial	$n = 200$ range = 5-6 & 9-10 yr	Laps	Adiposity status was determined using the Marshall Visual Rating Scale	Adiposity	Normal weight children had higher 20mSRT performance than obese children $F = 21.1$, $p < 0.001$	n/a
594 Raudsepp (1996); Estonia	Cross- sectional	$n = 77$ range = 10-11 yr age = 10.5 yr	Stages	SSF (mm) measured at 5 sites (triceps, biceps, subscapular, abdominal, medial calf); mean of 3 measurements/site used	Adiposity	20mSRT performance was negatively related to SSF $r = -0.51$, $p < 0.05$	n/a
601 Grassi (2006); Italy	Cross- sectional	$n = 290$ range = 14-18 yr	Speed last completed VO2max Leger	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI Boys $r = -0.45$, $p < 0.001$ Girls $r = -0.319$, $p < 0.01$	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was negatively related to BMI Boys $r = -0.41$, $p < 0.001$ Girls $r = -0.34$, $p < 0.001$	
647 Lloyd (2003); USA	Cross-sectional	n = 200 range = 10-12 yr	Laps	SSF (mm) measured at triceps and gastrocnemius BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related with SSF $r = -0.42$, $p < 0.05$ 20mSRT performance was negatively related with BMI $r = -0.49$, $p < 0.05$	n/a
720 Raudsepp (1996); Estonia	Cross-sectional	n = 203 range = 7-10 yr	Stages	SSF (mm) measured at 5 sites (triceps, biceps, subscapular, abdominal, medial calf)	Adiposity	20mSRT performance was negatively related with SSF Age 7 $r = -0.56$, $p < 0.01$ Age 8 $r = -0.36$, $p < 0.05$ Age 9 $r = -0.57$, $p < 0.05$ Age 10 $r = -0.48$, $p < 0.01$	n/a
745 Saar (2001); Estonia	Cross-sectional	n = 80 range = 13-15 yr age = 13.4 yr	Minutes	SSF (mm) measured at 9 sites (triceps, subscapular, biceps, iliac crest, supraspinatus, abdominal, front thigh, medial calf, mix-axilla).	Adiposity	20mSRT performance was a significant predictor of SSF $R^2 = 17.6$, $p < 0.0001$	n/a
847 Raudsepp (1996); Estonia	Cross-sectional	n = 92 range = 7-9 yr Boy age = 8.1 yr Girl age = 14.1 yr	Minutes	BMI was calculated from measured height and weight SSF (mm) measured at 5 sites (triceps, subscapular, supraspinatus, abdomen, thigh); mean of 3	Adiposity	20mSRT performance was not related to BMI Boys $r = -0.39$, $p = ns$ Girls $r = -0.12$, $p = ns$ 20mSRT performance was negatively related to SSF Boys $r = -0.52$, $p < 0.05$	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				measurements/site used		Girls $r = -0.45$, $p < 0.05$	
1083 Agha- Alinejad (2015); Iran	Cross- sectional	n = 381 range = 5-6 yr	VO2max n/a	WC (cm) was measured according to standardized procedures BMI was calculated from measured height and weight Waist to hip ratio was calculated from measured circumferences Waist to height ratio was calculated from measured metrics BF% was calculated from measured skinfolds	Adiposity	20mSRT performance was negatively related to WC Boys $r = -0.27$, $p < 0.001$ Girls $r = -0.16$, $p < 0.05$ 20mSRT performance was negatively related to BMI in girls but not boys Boys $r = -0.08$, $p = ns$ Girls $r = -0.21$, $p < 0.01$ 20mSRT performance was negatively related to waist to hip ratio in boys but not girls Boys $r = -0.160$, $p < 0.05$ Girls $r = 0.001$, $p = ns$ 20mSRT performance was negatively related to waist to height ratio Boys $r = -0.28$, $p < 0.001$ Girls $r = -0.35$, $p < 0.01$ 20mSRT performance was negatively related to BF% in boys but not girls Boys $r = -0.18$, $p < 0.05$ Girls $r = -0.09$, $p = ns$	n/a
1086 Carrel	Cross- sectional	n = 141 range = 10-14 yr	PACER z- score from	BMI z-score was calculated from measured	Adiposity	20mSRT z-score was negatively related to BMI z-score	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
(2014); USA		age = 12.7 yr	laps	height and weight		$r = -0.47$, 95 % CI -0.58 to -0.33, $p < 0.0001$	
1110 Bovet (2007); Seychelles	Cross- sectional	n = 4343 range = 12-15 yr	VO2max n/a	BMI was calculated from measured height and weight. IOTF cut-offs for lean, normal overweight and obese were used	Adiposity	Normal weight children attained healthy 20mSRT levels more than lean, overweight or obese children $p < 0.05$	n/a
1113 Albarwani (2008); Oman	Cross- sectional	n = 529 range = 15-16 yr	VO2max Leger	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related with BMI in urban but not rural children Urban Boys $r = -0.24$, $p = 0.05$ Urban Girls $r = -0.23$, $p = 0.01$ Rural Boys $r = -0.10$, $p = ns$ Rural Girls $r = -0.16$, $p = ns$	n/a
1134 Palomaki (2015); Finland	Cross- sectional	n = 3559 range = 15-16 yr 2003 age: 15.2 yr 2010 age: 15.3 yr	Shuttles	BMI was calculated from self-reported height and weight	Adiposity	Normal weight children completed more shuttles in 2003 and in 2010 than overweight children all $p < 0.001$	n/a
1169 Pienaar (2015); South Africa	Cross- sectional	n = 214 range = 15-16 yr age = 15.8 yr	VO2max n/a	Lean tissue mass (%) was calculated using the Poortman's equation Hip circumference (cm) was measured according to standardized procedures	Adiposity	Lean mass % was a positive predictor of 20mSRT performance $\beta = 0.63$, $p < 0.001$ Hip circumference was a negative predictor of 20mSRT performance $\beta = -0.16$, $p = 0.006$	Sport participation Stature % Lean Mass Ectomorphy Mid-thigh girth Living area
1205 Arango (2013);	Cross- sectional	n = 546 range = 11-18 yr age = 14.4 yr	VO2max Leger	WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was negatively associated with WC in boys but not in girls	Screen use TV viewing TST

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Colombia				BMI z-score was calculated from measured height and weight, with z-scores calculated from WHO data		Boys $\beta = -0.03$, SE = 0.01, $p < 0.01$ Girls $\beta = -0.05$, SE = 0.02, $p = \text{ns}$ 20mSRT performance was negatively associated with BMI in girls but not in boys Boys $\beta = -0.02$, SE = 0.01, $p = \text{ns}$ Girls $\beta = -0.05$, SE = 0.01, $p = 0.01$	Age Location PA level Institution SSB Fast food Fried food
1234 Andrade (2014); Ecuador	Cross-sectional	n = 648 range = 11-15 yr age = 13.6 yr	Laps	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was inversely related to BMI $\beta = -2.39$, $p < 0.01$	Sex BMI SES Cluster
1237 Scudder (2013); USA	Cross-sectional	n = 397 range = 6-9 yr age = 7.6 yr	Laps	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI $r = -0.32$, p not reported	n/a
1253 Diez-Fernandez (2014); Spain	Cross-sectional	n = 1158 range = 8-11 yr age = 9.5 yr	Stages (recorded to half stags)	BMI was calculated from measured height and weight WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was negatively associated with BMI Boys $\beta = -0.56$, $p < 0.001$ Girls $\beta = -0.45$, $p < 0.001$ 20mSRT performance was negatively associated with WC Boys $\beta = -0.56$, $p < 0.001$ Girls $\beta = -0.44$, $p < 0.001$	Age
1306 Tambalis (Greece); 2013	Cross-sectional	n = 141,169 range = 7-10 yr	Stages	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI Boys $r = -0.27$, $p < 0.01$ Girls $r = -0.25$, $p < 0.01$	Age

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1307 Buchan (2013); Scotland	Cross- sectional	n = 192 range = 14-16 yr age = 15.0 yr	Laps	BMI was calculated from measured height and weight WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was negatively related to BMI r = -0.27, p<0.001 20mSRT performance was negatively related to WC r = -0.29, p<0.001	Age Sex Maturity
1335 Coe (2013); USA	Cross- sectional	n = 1701 range = Grade 6 & 9	PACER No units specified	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI Gr. 6 r = -0.30, p<0.01 Gr. 9 r = -0.21, p<0.01	n/a
1342 Jimenez- Pavon (2013); International	Cross- sectional	n = 1053 range = 12.5-17.5 yr age = 14.9 yr	VO2max Leger	WC (cm) was measured according to standardized procedures BF% was calculated from SSF (biceps, triceps, subscapular, suprailiac, thigh, medial calf) BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to WC Boys r = -0.30, p<0.001 Girls r = -0.25, p<0.001 20mSRT performance was negatively related to BF% Boys r = -0.39, p<0.001 Girls r = -0.36, p<0.001 20mSRT performance was negatively related to BMI Boys r = -0.35, p<0.001 Girls r = -0.34, p<0.001	Pubertal status
1359 Burns (2013); USA	Cross- sectional	n = 134 range = Grades 6-8 age = 12.9 yr	Laps	BMI was calculated from measured height and weight SSF (mm) (triceps, calf) were used to measure	Adiposity	20mSRT performance was not related to BMI Boys r = -0.004, p = ns Girls r = -0.19, p = ns 20mSRT performance was	BMI Age

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				BF% BF% was measured using a handheld BIA WC (cm) was measured according to standardized procedures		negatively related to SSF Boys $r = -0.54$, $p < 0.001$ Girls $r = -0.26$, $p < 0.05$ 20mSRT performance was negatively related to BF% in boys but not girls Boys $r = -0.54$, $p < 0.001$ Girls $r = -0.20$, $p = ns$ 20mSRT performance was not related to WC Boys $r = 0.03$, $p = ns$ Girls $r = -0.17$, $p = ns$	
1376 Martins (2013); Portugal	Cross- sectional	n = 79 range = 11-13 yr	VO2max Fitnessgram	BMI was calculated from measured height and weight WC was measured according to standardized procedures	Adiposity	Fit children presented lower BMI values than unfit children $p < 0.05$ Fit children presented lower WC values than unfit children ($p < 0.05$)	n/a
1396 Ganzalez- Suarez (2013); Philippines	Cross- sectional	n = 1021 range = 10-12	VO2max Leger	BMI was calculated from measured height and weight WC (cm) was measured according to standardized procedures	Adiposity	Normal weight children presented higher 20mSRT performance than overweight or obese children $p < 0.0001$ Overweight boys had decreased odds of performing poorly on the 20mSRT compared to normal weight boys. No difference was	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						observed in girls. Boys OR 0.84, 95 % CI 0.78-0.91 Girls OR 1.42, 95% CI 0.94-2.16 Obese children had increased odds of performing poorly on the 20mSRT compared to normal weight children Boys OR 4.54, 95% CI 2.64-7.79 Girls OR 2.51, 95% CI 1.31-4.80 20mSRT performance was negatively related to BMI Boys $r = -0.41$, $p < 0.0001$ Girls $r = -0.28$, $p < 0.0001$ 20mSRT performance was negatively related to WC Boys $r = -0.38$, $p < 0.0001$ Girls $r = -0.27$, $p < 0.0001$	
1407 Coe (2012); USA	Cross-sectional	n = 312 range = 12-14 yr age = 12.1 yr	Laps	BF% was calculated using the Slaughter equation from SSF measures (calf, triceps).	Adiposity	20mSRT performance was negatively related to BF% $r = -0.54$, $p < 0.01$	n/a
1423 Lewitt (2012); Wales	Cross-sectional	n = 247 range = Gr. 8 Boy age = 13.5 yr Girl age = 13.5 yr	Laps	BF% was calculated using the Slaughter equation from SSF measures (triceps, biceps, subscapular, suprailiac).	Adiposity	20mSRT performance was negatively related to BF% but varied across maturity levels Boys Early $r = -0.59$, $p < 0.001$ Boys Late $r = -0.53$, $p < 0.001$ Girls Early $r = -0.55$, $p < 0.001$	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				WC (cm) was calculated according to standardized procedures		Girls Late $r = -0.55$, $p < 0.001$ 20mSRT performance was negatively related to WC Boys Early $r = -0.49$, $p < 0.01$ Boys Late $r = -0.40$, $p < 0.01$ Girls Early $r = -0.55$, $p < 0.01$ Girls Late $r = -0.36$, $p < 0.01$	
1427 Galaviz (2012); Mexico	Cross-sectional	$n = 193$ range = 10-12 yr Boy age = 10.1 yr Girl age = 10.2 yr	VO2max Leger	WC (cm) was measured according to standardized procedures BMI was calculated from measured height and weight SSF (mm) was calculated from site measurements (triceps, biceps, calf).	Adiposity	20mSRT performance was negatively related to WC Boys $r = -0.59$, $p < 0.001$ Girls $r = -0.57$, $p < 0.001$ 20mSRT performance was negatively related to BMI Boys $r = -0.57$, $p < 0.001$ Girls $r = -0.57$, $p < 0.001$ 20mSRT performance was negatively related to SSF Boys $r = -0.64$, $p < 0.001$ Girls $r = -0.58$, $p < 0.001$ 20mSRT performance was a negative predictor for WC $R^2 = 0.23$, $p < 0.01$ 20mSRT performance was a negative predictor for BMI $R^2 = 0.23$, $p < 0.01$	Age Sex PA level

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was a negative predictor of SSF $R^2 = 0.34$, $p < 0.01$	
1432 Musa (2012); Nigeria	Cross-sectional	n = 3243 range = 9-15 yr Boy age = 10.1 yr Girl age = 10.2 yr Male age = 13.5 yr Female age = 13.6 yr	Laps	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was not related to BMI Boys $r = 0.02$ Girls $r = -0.18$ Male $r = 0.002$ Female $r = 0.04$ $p = ns$ for all	n/a
1433 Llorente-Cantarero (2012); Spain	Cross-sectional	n = 141 range = 7-12 yr	Stages	BMI was calculated from measured height and weight WC (cm) was measured according to standardized procedures	Adiposity	Children with low 20mSRT performance presented higher BMI values $p < 0.0001$ Children with low 20mSRT performance presented higher WC values $p < 0.0001$	n/a
1444 Duncan (2012); Portugal	Cross-sectional	n = 445 range = 10-17 yr age = 13.9 yr	VO2max Ruiz	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI Boys $r = -0.70$, $p = 0.0001$ Girls $r = -0.69$, $p = 0.001$	n/a
1461 Machado-Rodrigues (2012); Portugal	Cross-sectional	n = 362 range = 13-16 yr	Laps	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI Rural $r = -0.35$, $p < 0.01$ Urban $r = -0.42$, $p < 0.01$ Higher levels of 20mSRT performance were associated with a lower risk of being	Age Sex SB level

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						overweight or obese Rural $\beta = -1.51$, SE = 0.42 95% CI 0.10-0.50, $p < 0.01$ Urban $\beta = -1.37$, SE = 0.58 95% CI 0.08-0.80, $p = 0.02$	
1480 Joshi (2012); USA	Cross- sectional	n = 6625 range = 5-17 yr	HFZ	BMI was calculated from measured height and weight	Adiposity	Normal weight children were more likely to achieve the healthy levels of 20mSRT performance than obese children Chi-square = 192.6 OR 3.45, $p < 0.001$	n/a
1483 Ebenegger (2012); Switzerland	Cross- sectional	n = 600 range = 4-6 yr age = 5.1 yr	Stages	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was higher in normal weight children than overweight or obese children $p < 0.001$	Sex Age Parental migrant status Education
1484 Niederer (2012); Switzerland	Cross- sectional	n = 501 range = 5-6 yr	Stages	BMI was calculated from measured height and weight	Adiposity	Overweight and obese children had lower 20mSRT performance compared to normal weight children $p < 0.001$, effect size = 0.34	Age BMI Sex Class cluster BMI group- age group x
1526 Artero (2011); International	Cross- sectional	n = 709 range = 12.5-17.5 yr age = 14.9 yr	Stages	WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was negatively related to WC $r = -0.29$, $p < 0.001$	Age Sex Pubertal status Centre
1550 Aires (2011);	Cross- sectional	n = 1708 range = 11-19 yr	VO2max Ruiz	SSF (mm) was measured using three sites (triceps, subscapular and calf).	Adiposity	20mSRT performance was negatively associated with SSF $B = -0.07$, 95 % CI -0.076 to -	Sex Maturation SES

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Portugal						0.060 p = 0.000	PA index TV time PC time School commute
1553 Castro- Pinero (2011); Spain	Cross- sectional	n = 2462 range = 6-17.9 yr	Stages	BMI was calculated from measured height and weight	Adiposity	Normal weight children completed more 20mSRT stages than overweight and obese children. Normal vs Overweight p<0.001 Normal vs Obese p<0.001	Age
1563 Gonzalez- Suarez (2011); Philippines	Cross- sectional	n = 380 range = 11-12 yr	VO2max Leger	BMI was calculated from measured height and weight	Adiposity	Normal weight children had higher 20mSRT performance than overweight children p = 0.001 Children with 20mSRT performance below the median score were more likely to be classified as overweight or obese OW OR 0.5, 95 % CI 0.3-0.8 Obese OR 0.2, 95 % CI 0.1-0.5	n/a
1572 Jimenez- Pavon (2011); International	Cross- sectional	n = 1053 range = 12.5-17.5 yr age = 14.9 yr	VO2max Leger	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI Boys r = -0.45, p<0.01 Girls r = -0.40, p<0.01	n/a
1620 Moliner- Urdiales (2011);	Cross- sectional	n = 363 range = 12.5-17.5 yr Boy age = 14.8 yr	Stages	Total body fat (kg) was quantified by DXA Total body fat (kg) was	Adiposity	20mSRT performance was inversely associated with BF measured by DXA Boys β = -0.54, R^2 = 0.26,	Age Pubertal status PA level

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
International		Girl age = 14.8 yr		measured by BodPod Total body fat was measured by SSF (mm) (biceps, triceps, subscapular, suprailiac, thigh, calf) WC (cm) was measured according to standardized procedures		$p < 0.001$ Girls $\beta = -0.46$, $R^2 = 0.33$, $p < 0.001$ 20mSRT performance was inversely associated with BF measured by BodPod Boys $\beta = -0.55$, $R^2 = 0.28$, $p < 0.001$ Girls $\beta = -0.40$, $R^2 = 0.25$, $p < 0.001$ 20mSRT performance was inversely associated with BF measured by SSF Boys $\beta = -0.54$, $R^2 = 0.34$, $p < 0.001$ Girls $\beta = -0.41$, $R^2 = 0.20$, $p < 0.001$ 20mSRT performance was inversely associated with WC in boys but not girls Boys $\beta = -0.53$, $R^2 = 0.28$, $p < 0.001$ Girls $\beta = 0.12$, $R^2 = 0.07$, $p = 0.143$	
1624 Pino-Ortega (2010); Spain	Cross- sectional	n = 293 range = 9-11 yr age = 9.9 yr	Stages	BMI was calculated from measured height and weight	Adiposity	Normal weight children had increased odds of having higher 20mSRT performance OR 12.14, 95 % CI 4.9-30.5,	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						p<0.000	
1628 Marques- Vidal; (2010); Portugal	Cross- sectional	n = 4689 range = 10-18 yr Boy age = 14.6 yr Girl age = 14.6 yr	VO2max Fitnessgram	BMI was calculated from measured height and weight BF% was measured by handheld BIA	Adiposity	20mSRT performance was negatively related to BMI Boys r = -0.53, p<0.01 Girls r = -0.50, p<0.01 20mSRT performance was negatively related to BF% Boys r = -0.45, p<0.01 Girls r = -0.33, p<0.01	n/a
1630 Ostojic (2010); Serbia	Cross- sectional	n = 84 range = 10-11 yr age = 11.9 yr	VO2max Ruiz	BMI was calculated from measured height and weight WC (cm) was measured according to standardized procedures Abdominal skinfold thickness (mm) was measured in triplicate BF% was measured by BIA Muscle Mass % was measured by Bia	Adiposity	20mSRT performance was negatively related to BMI Overweight r = -0.51, p<0.05 Non-overweight r = -0.40, p<0.05 20mSRT performance was negatively related to WC in overweight children but not in non-overweight children Overweight r = -0.51, p<0.05 Non-overweight r = -0.11, p = ns 20mSRT performance was negatively related to abdominal skinfold thickness Overweight r = -0.54, p<0.05 Non-overweight r = -0.27, p<0.05 20mSRT performance was	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						negatively related to BF% Overweight $r = -0.62$, $p < 0.05$ Non-overweight $r = -0.36$, $p < 0.05$ 20mSRT performance was related to muscle mass in overweight children but not in the non-overweight group Overweight $r = 0.68$, $p < 0.05$ Non-overweight $r = 0.20$, $p = ns$	
1644 Ostojic (2010); Serbia	Cross- sectional	n = 245 Overweight = 8.8 yr Non-overweight = 8.9 yr	VO2max Leger	BMI was calculated from measured height and weight WC (cm) was measured according to standardized procedures SSF (mm) were measured in triplicate (abdominal, biceps, triceps, subscapular, suprailiac) BF% was calculated from the Slaughter equations	Adiposity	20mSRT performance was negatively related to BMI in the non-overweight group but not in the overweight group Overweight $r = -0.15$, $p = ns$ Non-overweight $r = -0.43$, $p < 0.05$ 20mSRT performance was negatively related to WC in the non-overweight group but not in the overweight group Overweight $r = -0.12$, $p = ns$ Non-overweight $r = -0.37$, $p < 0.05$ 20mSRT performance was negatively related to abdominal skinfold thickness Overweight $r = -0.38$, $p < 0.05$	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Non-overweight $r = -0.43$, $p < 0.05$ 20mSRT performance was negatively related to subscapular skinfold thickness Overweight $r = -0.29$, $p < 0.05$ Non-overweight $r = -0.41$, $p < 0.05$ 20mSRT performance was negatively related to suprailiac skinfold thickness Overweight $r = -0.43$, $p < 0.05$ Non-overweight $r = -0.47$, $p < 0.05$ 20mSRT performance was negatively related to BF% Overweight $r = -0.36$, $p < 0.05$ Non-overweight $r = -0.38$, $p < 0.05$	
1653 Houwen (2010); Netherlands	Cross- sectional	n = 60 range = 6-12 yr age = 8.9 yr	Stages	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI $r = -0.39$, 95 % CI -0.60 to 0.13 $p < 0.01$	Age Sex
1657 Martinez- Gomez (2010); Spain	Cross- sectional	n = 192 range = 13-17 yr age = 14.8 yr	Laps	BMI was calculated from measured height and weight SSF (mm) was measured	Adiposity	20mSRT performance was negatively related to BMI $r = -0.35$, $p < 0.001$ 20mSRT performance was	Age Sex Pubertal status

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				according to standardized procedures WC (cm) was measured according to standardized procedures		negatively related to SSF $r = -0.45, p < 0.001$ 20mSRT performance was negatively related to WC $r = -0.30, p < 0.001$	
1677 Jago (2010); USA	Cross- sectional	n = 4955 range = 11-12 yr Boy age = 11.4 yr Girl age = 11.3 yr	Laps	WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was associated with WC Boys $F = 33.5, p < 0.0001$ Girls $F = 36.3, p < 0.0001$	Race, Household education Pubertal stage BMI
1692 Vale (2010); Portugal	Cross- sectional	n = 788 range = 12-18 yr age = 15.1 yr	VO2max Fitnessgram	BMI was calculated from measured height and weight	Adiposity	Overweight children completed fewer 20mSRT laps than normal weight children Boys $p < 0.0001$ Girls $p < 0.0001$ Boys and girls with low 20mSRT performance were at increased odds of being overweight or obese Boys OR 3.75, 95 % CI 2.14- 6.59 Girls OR 2.66, 95 % CI 1.02- 2.67 Both $p < 0.05$	n/a
1709 Aires (2010); Portugal	Cross- sectional	n = 111 range = 11-18 yr age = 14.5 yr	VO2max Fitnessgram	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI $r = -0.20, p < 0.05$	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Children with higher 20mSRT performance had lower odds of being overweight or obese OR 0.97, 95 % CI 0.94-0.998, p = 0.037	
1735 Martinez-Gomez (2009); Spain	Cross-sectional	n = 202 range = 13-17 yr Boy age = 14.7 yr Girl age = 14.9 yr	Laps	BMI was calculated from measured height and weight SSF (mm) was measured on 6 sites on the left side of the body (triceps, biceps, subscapular, suprailiac, thigh, calf) WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was negatively related to BMI r = -0.27, p<0.001 20mSRT performance was negatively related to SSF r = -0.44, p<0.001 20mSRT performance was negatively related to WC r = -0.26, p<0.001	Age Sex Maturity
1769 Stigman (2009); Finland	Cross-sectional	n = 304 range = 8-9 yr age = 8.6 yr	Laps	WC (no units reported) was measured according to standardized procedures Abdominal fat % was quantified by DXA BF% was quantified by DXA	Adiposity	20mSRT performance was negatively related to WC p = 0.001 20mSRT performance was negatively related to abdominal fat % p<0.001 20mSRT performance was negatively related to BF% p<0.001	Age Sex
1781	Cross-	n = 573	Stages	BF% was quantified by	Adiposity	Higher BF% was associated with	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Sakuragi (2008); Australia	sectional	range = 10 yr age = 10.1 yr		DXA		decreased 20mSRT performance $p < 0.0001$	
1784 Ortega (2009); Spain	Cross- sectional	n = 1801 range = 13-18.5 yr Boy age = 15.3 yr Girl age = 15.4 yr	VO2max Leger	BMI was calculated from measured height and weight Fat free mass (kg) was calculated using the Slaughter equations from skinfold thickness BF% was calculated using the Slaughter equations from skinfold thickness WC (cm) was measured according to standardized procedures	Adiposity	20mSRT performance was negatively related to BMI Boy $r = -0.35$, $p < 0.001$ Girl $r = -0.25$, $p < 0.001$ 20mSRT performance was not related to FFM Boy $r = 0.07$, $p = 0.06$ Girl $r = -0.06$, $p = 0.12$ 20mSRT performance was negatively related to BF% Boy $r = -0.47$, $p < 0.001$ Girl $r = -0.31$, $p < 0.001$ 20mSRT performance was negatively related to WC Boy $r = -0.34$, $p < 0.001$ Girl $r = -0.21$, $p < 0.001$	n/a
1797 Hoekstra (2008); Ireland	Cross- sectional	n = 2016 range = 12 & 15 yr	VO2max n/a	SSF (mm) from 4 sites (biceps, triceps, subscapular, suprailiac) were measured according to published methods	Adiposity	20mSRT performance was inversely associated with SSF Boy 12 yr $\beta = -1.54$, 95 % CI - 1.86 to -1.22, $p < 0.001$ Boy 15 yr $\beta = -1.28$, 95 % CI - 1.57 to -0.99, $p < 0.001$ Girl 12 yr $\beta = -0.86$, 95 % CI - 1.15 to -0.56, $p < 0.001$	Maturation SES PA score Muscular fitness

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Girl 15 yr $\beta = -0.92$, 95 % CI -1.21 to -0.63, $p < 0.001$	
1838 Downs (2008); Canada	Cross- sectional	n = 178 range = 9-12 yr	Minutes	WC (cm) was measured according to standardized procedures	Adiposity	Meeting or exceeding the 20 th percentile for 20mSRT was associated with a reduced risk for central adiposity OR 0.11, 95 % CI 0.04-0.33, $p < 0.05$ Children with central adiposity had shorter 20mSRT durations than children without central adiposity $p < 0.0001$	Age Sex
1846 Cairney (2008); Canada	Cross- sectional	n = 2245 range = 9-10 yr age = 9.4 yr	Stages	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was inversely associated with BMI $\beta = -0.19$, SE = -0.42, $p < 0.001$	Sex
1850 Buck (2007); USA	Cross- sectional	n = 74 range = 7-12 yr age = 9.3 yr	Laps	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was negatively related to BMI $r = -0.49$, $p < 0.05$	n/a
1872 Casajus (2007); Spain	Cross- sectional	n = 1068 range = 7-12 yr	Stages	BMI was calculated from measured height and weight	Adiposity	Normal weight children completed more 20mSRT laps than overweight or obese children $F = 46.8$, $p < 0.05$	n/a
1915 Burke (2006); Australia	Cross- sectional	n = 599 range = 11-14 yr Boy age = 12.0 yr Girls age = 12.0 yr	Laps	BMI was calculated from measured height and weight	Adiposity	Risk of overweight/obesity was lower in those children who completed more 20mSRT laps Boy OR 0.64, 95 % CI 0.47-0.86	School cluster Age SES

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Girl OR 0.95, 95 % CI 0.91-0.99	TV viewing Computer use Video games Other SB Walk/cycle to school
1921 Ng (2006); Canada	Cross- sectional	n = 82 range = 9-12 yr	Minutes	BMI was calculated from measured height and weight	Adiposity	Obese children obtained lower 20mSRT performance scores compared to overweight and normal weight children. $F = 21.01, p < 0.0001$	n/a
1928 Tokmakisis (2006); Greece	Cross- sectional	n = 709 range = n/a age = 8.9 yr	VO2max Leger	BMI was calculated from measured height and weight	Adiposity	Normal weight children had higher 20mSRT performance levels than obese children Boys $p < 0.01$ Girls $p < 0.01$ Normal weight boys had higher 20mSRT performance levels than overweight boys; no difference was observed in girls Boys $p < 0.05$ Girls $p = ns$	n/a
1992 Kim (2005); USA	Cross- sectional	n = 6297 range = 5-14 yr	HFZ	BMI was calculated from measured height and weight	Adiposity	Not attaining healthy levels of 20mSRT performance was associated with being overweight Boys OR 2.34, 95 % CI 2.08- 2.64 Girls OR 2.32, 95 % CI 2.02- 2.66	Age Race Lunch program Baseline BMI

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
2015 Nassis (2005); Greece	Cross- sectional	n = 1362 range = 6-13 yr	Stages	SSF (mm) were measured in triplicate at four sites (triceps, biceps, subscapular, suprailiac) BF% was quantified from SSF using the Slaughter equations BMI was calculated from measured height and weight	Adiposity	20mSRT performance favourably affected SSF in fit normal weight and overweight children Both $p < 0.01$ 20mSRT performance favourably affected BF% in fit normal weight and overweight children Both $p < 0.01$ 20mSRT performance was negatively related to BMI $r = -0.27, p < 0.001$	n/a
2049 Gonzalez- Gross (2003); Spain	Cross- sectional	n = 101 range = 15-18 yr	Stages	SSF (mm) was measured according to standardized protocols (biceps, triceps, subscapular, suprailiac)	Adiposity	20mSRT performance was negatively related to SSF Boys $r = -0.65, p < 0.01$ Girls $r = -0.43, p < 0.01$	n/a
2051 Cieslack (2003); Canada	Cross- sectional	n = 61 range = 10-11 yr	VO2max Leger	BF% was measured by BIA	Adiposity	20mSRT performance was not related to BF% $r = 0.11, p = ns$	n/a
2066 Deforche (2003); Belgium	Cross- sectional	n = 3214 range = 12-18 yr	n/a	BMI was calculated from measured height and weight	Adiposity	Non-obese children performed better on the 20mSRT $F = 359.3, p < 0.001$	n/a
2083 Mota (2002);	Cross- sectional	n = 494 range = 8-16 yr	VO2max Leger	SSF (mm) were measured twice at each location (triceps, subscapular) by a	Adiposity	20mSRT performance was not related to SSF Boys $r = 0.13, p = ns$	Maturity Age

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Portugal				trained technician BF% was quantified from SSF using the Slaughter equations		Girls $r = -0.12$, $p = ns$ 20mSRT performance was negatively related to BF% Boys $r = -0.49$, $p < 0.01$ Girls $r = -0.34$, $p < 0.01$ 20mSRT performance was associated with skinfold thickness Boys $F = 23.84$, $R^2 = 0.28$, $p < 0.001$ Girls $F = 13.20$, $R^2 = 0.18$, $p < 0.001$ 20mSRT performance was associated with BF% Boys $F = 34.52$, $R^2 = 0.28$, $p < 0.001$ Girls $F = 20.58$, $R^2 = 0.18$, $p < 0.001$	
2458 Coledam (2015); Brazil	Cross- sectional	$n = 781$ range = 10-18 yr	Criterion Ref. Leger	BMI was calculated from measured height and weight	Adiposity	Meeting the healthy level for BMI was associated with 20mSRT performance PR 1.49, 95 % CI 1.27-1.75, $p < 0.001$	Sex Age SES
2542 Saavedra (2014); Spain	Cross- sectional	$n = 351$ range = 8-9 yr age = 8.7 yr	Stages	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was not associated with BMI $\beta = -0.36$, 95 % CI -0.453 to - 0.263, $p < 0.05$	n/a
2681	Cross-	$n = 1206$	Laps	BMI was calculated from	Adiposity	20mSRT performance was	Age

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Coelho-e-Silva (2013); Portugal	sectional	range = 11-15 yr		measured height and weight BF% was calculated from measured skinfold thickness		negatively related to BMI $r = -0.29$, 95 % CI -0.31 to -0.24 $p < 0.01$ 20mSRT performance was negatively related to BF% $r = -0.35$, 95 % CI -0.40 to -0.30 Fit children were less likely to be classified as overweight Boys PR 0.86, 95 % CI 0.81-0.91 Girls PR 0.87, 95 % CI 0.82-0.91 Fit children were less likely to be classified as overfat Boys PR 0.86, 95 % CI 0.82-0.91 Girls PR 0.88, 95 % CI 0.84-0.93	
2922 Vale (2011); Portugal	Cross-sectional	$n = 779$ range = 12-18 yr Boy age = 14.8 yr Girl age = 15.2 yr	Laps	Waist to height ratio was calculated from measured WC and height	Adiposity	Children in the risk category for waist to height ratio had lower 20mSRT performance Boys $p < 0.001$ Girls $p < 0.001$ Children with low 20mSRT performance were more likely to have abdominal obesity than their fit peers Boys OR 6.43, 95 % CI 3.33-12.39, $p < 0.05$ Girls OR 1.78, 95 % CI 1.14-	n/a

Ref # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						2.78, p<0.05	
3109 Artero (2010); Spain	Cross- sectional	n = 2474 range = 13-18.5 yr Boy age = 15.2 yr Girl age = 15.4 yr	Stages	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was associated with weight status when adjusting for FFM Boys p<0.001; Girls p<0.001 20mSRT performance was not associated with weight status when adjusting for FM Boys p = 0.06; Girls p = 0.053	Age Maturation Fat Mass
3609 Fogelholm (2008); Finland	Cross- sectional	n = 2266 range = 15-16 yr	Laps	BMI was calculated from measured height and weight	Adiposity	20mSRT performance was not associated with weight status in children Boys β = -0.24, p = ns Girls β = -0.24, p = ns	n/a

Note* 20mSRT, 20 meter shuttle run test; BIA, bioelectrical impedance analysis; BF, body fat; BMI, body mass index; DXA, dual-energy x-ray absorptiometry; HFZ, Healthy Fitness Zone; NIZ, Needs Improvement Zone; PACER, Progressive Aerobic Cardiovascular Endurance Run; SSF, sum of skinfolds; WC, waist circumference

“Age” in sample column refers to mean age unless stated otherwise. If “age” is not stated in the sample column it’s due to the authors not reporting mean age of sample.

Supplement Table 2. 20mSRT performance and cardiometabolic biomarkers in children and youth

Note: β indicates standardized Beta; B indicates unstandardized Beta

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1422 Telford (2012); Australia	Longitudinal (4 yr follow- up)	n age 8 yr = 734 n age 10 yr = 586 n age 12 yr = 534	Stages	HOMA-IR was calculated using standard procedures	Cardiometabolic	20mSRT performance was not a within-child predictor of HOMA-IR from age 8 to 12 years Boys β = -0.06, SE = 0.088, p = ns Girls β = -0.118, SE = 0.094, p = ns	Sex School SES Pubertal status Weight BF% PA level
1582 Puder (2011); Switzerland	Longitudinal (9 mo. Follow-up)	n = 83 T1 range = 6-7 yr T1 range = 11-12 yr	Stages	CRP (mg/L) was measured following standard procedures HOMA-IR was calculated using standard procedures	Cardiometabolic	20mSRT performance was not a predictor of 9 month change in CRP β = -0.06, 95 % CI -0.24 to 0.12, p = 0.50 20mSRT performance was not a predictor of 9 month change in HOMA- IR β = -0.05, 95 % CI -0.17 to 0.28, p = 0.60	Age group Sex Pubertal status
1742 Telford (2009); Australia	Longitudinal (2 yr follow- up)	N = 498 T1 age = 8 yr T2 age	Stages	HOMA-IR was calculated from measured fasting insulin and glucose	Cardiometabolic	20mSRT performance was inversely associated with HOMA-IR at follow-up in boys but not	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
		= 10 yr				girls Boys $p = 0.05$ Girls $p = 0.31$	
238 Silva (2013); Portugal	Cross- sectional	n = 633 range = 10-18 yr age = 14.3 yr	VO _{2max} Fitnessgram	TC (mg/dl) was measured using standard procedures HDL (mg/dl) was measured using standard procedures LDL (mg/dl) was measured using standard procedures TG (mg/dl) were measured using standard procedures Glucose (mg/dl) was measured from fasting blood samples SBP (mmHg) was measured using standard procedures DBP (mmHg) was measured using standard procedures	Cardiometabolic	20mSRT performance was negatively related with TC $r = -0.102, p < 0.05$ There was no difference in TC between high and low fit groups ($F = 0.159, p = ns$) 20mSRT performance was not related with HDL $r = -0.035, p = ns$ There was no difference in HDL between high and low fit groups $F = 1.009, p = ns$ 20mSRT performance was not related with LDL $r = -0.05, p = ns$ There was no difference in LDL between high and low fit groups $F = 0.474, p = ns$	Age Sex WC

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				MAP (mmHg) was calculated using standard procedures CMR was calculated from 4 health markers: HDL, TG, Glucose, and MAP		20mSRT performance was negatively related with TG $r = -0.115, p < 0.05$ There was a difference in TG between high and low fit groups $F = 8.711, p < 0.05$ 20mSRT performance was not related with glucose $r = 0.029, p = ns$ There was no difference in glucose levels between high and low fit groups $F = 0.471, p = ns$ 20mSRT performance was not related with SBP $r = -0.009, p = ns$ There was no difference in SBP between high and low fit groups $F = 1.080, p = ns$ 20mSRT performance was not related with DBP $r = -0.064, p = ns$	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						There was no difference in DBP between high and low fit groups $F = 1.874$, $p = ns$ 20mSRT performance was not related with MAP $r = -0.046$, $p = ns$ There was no difference in MAP between high and low fit groups $F = 1.893$, $p = ns$ 20mSRT performance was negatively related with CMR $r = -0.095$, $p < 0.05$ There was a difference in CMR between high and low fit groups $F = 5.999$, $p < 0.05$ Those with low 20mSRT performance had greater odds of having high cardiometabolic risk than those with high 20mSRT performance OR 2.06, 95 % CI 1.24 to 3.43, $p < 0.05$	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
339 Thomas (2008); Wales	Cross- sectional	n = 164 range = 12-13 yr age = 12.9 yr	Shuttles	CRP (mg/l) was measured from fasting blood samples	Cardiometabolic	20mSRT performance was not related with CRP Boys r = -0.02, p = 0.099 Girls r = -0.11, p = 0.301	n/a
999 Sobrinho (2015); Portugal	Cross- sectional	n = 386 range = 12-18 yr age = 14.03 yr	VO _{2max} Léger	CRP (mg/l) was measured from fasting blood samples	Cardiometabolic	20mSRT performance was negatively related with CRP in girls but not boys Boys r = -0.039, p = ns Girls r = -0.215, p<0.001	Age Maturation BMI
1056 Buchan (2015); Scotland	Cross- sectional	n = 192 range = 15-17.5 yr age = 16.7 yr	Shuttles	CMR was calculated from 5 health markers: interleukin-6, PAI-1, CRP, adiponectin (inverted), and fibrinogen	Cardiometabolic	20mSRT performance was a negative predictor of CMR score β = -0.015, p<0.001	Sex Age PA level WC Maturation Muscular fitness
1086 Carrel (2014); USA	Cross- sectional	n = 141 range = grades 5-8 age = 12.7 yr	Laps	HOMA-IR was calculated from measured glucose and insulin	Cardiometabolic	20mSRT performance was negatively related with HOMA-IR r = -0.60, 95 % CI -0.70 to -0.48, p<0.0001	n/a
1098 Curtis (2012); USA	Cross- sectional	n = 203 range = 11-14 yr age = 12.2 yr	Laps	HOMA-IR was calculated from measured glucose and insulin CRP (mg/l) was measured using standard procedures	Cardiometabolic	20mSRT performance was inversely associated with HOMA-IR in boys but not girls Boys p = 0.004 Girls p = ns 20mSRT performance	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						was inversely associated with CRP Boys $p < 0.01$ Girls $p < 0.01$	
1106 Varness (2009); USA	Cross- sectional	n = 82 age = 13.0 yr	Laps and VO_{2max} Léger	HOMA-IR was calculated from measured glucose and insulin	Cardiometabolic	20mSRT performance was inversely associated with HOMA-IR $r_s = -0.60, p < 0.0001$ Predicted VO_{2max} was a predictor of HOMA-IR $\beta = -0.21, SE = 0.04,$ $p < 0.001$ 20mSRT performance was a predictor of HOMA-IR $\beta = -0.012, SE = 0.004, p$ $= 0.05$	Age Sex BMI z-score
1117 Benser (2015); International	Cross- sectional	n = 713 range = 12.5- 17.5 yr age = 14.8 yr	Stages and VO_{2max} Léger	Homocysteine ($\mu\text{mol/l}$) was measured from fasting blood samples	Cardiometabolic	20mSRT performance was positively associated with homocysteine in boys but not girls Boys $\beta = 0.015, p =$ 0.051 Girls $\beta = -0.278, p =$ 0.222 Predicted VO_{2max} was not a predictor of homocysteine Boys $\beta = -0.246, p =$	Age Maturation Folate Cobalamin Creatinine Smoking Supplement use MTHFR 677C/T polymorphism

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						0.100 Girls $\beta = 0.096$, $p = 0.370$	
1157 Machado- Rodrigues (2015); Brazil	Cross- sectional	n = 924 range = 11-17 yr age = 13.7 yr	Laps	CMR was calculated from 5 health markers: insulin, TG, glucose, HDL, BP	Cardiometabolic	20mSRT performance was inversely associated with CMR score Boys $\beta = -0.14$, 95 % CI - 0.03 to -0.01 Girls $\beta = -0.06$, 95 % CI - 0.04 to 0.00	MVPA Age WC
1201 Artero (2013); Europe	Cross- sectional	n = 639 range = 12.5- 17.5 yr age = 14.9 yr	Half stages	CRP (mg/l) was measured from fasting blood samples Complement factor C3(g/l) was measured from fasting blood samples Complement factor C4 (g/l) was measured from fasting blood samples White blood cell count ($10^3/\mu\text{l}$) was measured using standard procedures Inflammatory marker score was calculated from 5 inflammation markers	Cardiometabolic	20mSRT performance was negatively related with CRP $r = -0.141$, $p < 0.001$ 20mSRT performance was negatively related with complement factor C3 $r = -0.198$, $p < 0.001$ 20mSRT performance was negatively related with complement factor C4 $r = -0.177$, $p < 0.001$ 20mSRT performance was negatively related with white blood cell count $r = -0.308$, $p < 0.05$	Sex Age Pubertal status Test centre

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was negatively related with inflammatory marker score $r = -0.308, p < 0.001$	
1234 Andrade (2014); Ecuador	Cross-sectional	n = 648 range = 11-15 yr age = 13.6 yr	Laps and health / unhealthy fitness (Fitnessgram cut-points)	Dyslipidemia was classified when at least one lipid profile reached a risk level Cholesterol (mg/l) was measured from fasting blood samples HDL (mg/l) cholesterol was measures from fasting blood samples LDL (mg/l) cholesterol was measures from fasting blood samples TG were measured from fasting blood samples	Cardiometabolic	20mSRT performance was not a predictor of dyslipidemia $\beta = 0.85, p = 0.20$ 20mSRT performance was not a predictor of cholesterol Urban $\beta\% = -1.54, p = 0.18$ Rural $\beta\% = -2.70, p = 0.11$ 20mSRT performance was not a predictor of HDL $\beta\% = -0.53, p = 0.65$ 20mSRT performance was not associated with LDL Urban $\beta\% = -1.72, p = 0.34$ Rural $\beta\% = -5.70, p = 0.13$ 20mSRT performance	Sex BMI SES Energy intake Residence Cluster

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						was not associated with TG $\beta\% = -0.23, p = 0.95$ Healthy fitness was not a predictor of dyslipidemia $\beta = 0.85, p = 0.46$ Healthy fitness was associated with cholesterol levels in rural but not urban children Rural $\beta\% = -8.91, p < 0.01$ Urban $\beta\% = -5.25, p = 0.46$ Healthy fitness was not associated with HDL $\beta\% = -5.31, p = 0.08$ Healthy fitness was associated with LDL in rural but not urban children Rural $\beta\% = -11.71, p = 0.04$ Urban $\beta\% = -3.31, p = 0.46$ Healthy fitness was not associated with TG in children $\beta\% = -4.03, p = 0.67$	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1241 Machado- Rodrigues (2014); Brazil	Cross- sectional	n = 924 range = 11-17 yr age = 13.7 yr	Laps	CMR was calculated from 5 health markers: insulin, TG, glucose, HDL, and BP	Cardiometabolic	20mSRT performance was a negative predictor of CMR in children B = -0.18, B = -0.02, 95 % CI -0.03 to -0.01, p<0.01	Age Sex BMI MVPA Parental education
1253 Diez- Fernandez (2014); Spain	Cross- sectional	n = 1158 range = 8-11 yr age = 9.49 yr	Half stages	CMR was calculated from 5 health markers: WC, TG/HDL ratio, MAP, insulin Insulin (mg/dl) was measured from fasting blood samples TG/HDL (mg/dl) cholesterol ratio was measured from fasting blood samples MAP (mmHg) was calculated from measured systolic and diastolic blood pressure	Cardiometabolic	20mSRT performance was a negative predictor of CMR Boys β = -0.429, p<0.001 Girls β = -0.321, p<0.001 20mSRT performance was a negative predictor of insulin Boys β = -0.409, p<0.001 Girls β = -0.284, p<0.001 20mSRT performance was a negative predictor of TG/HDL ratio Boys β = -0.308, p<0.001 Girls β = -0.311, p<0.001 20mSRT performance was a negative predictor of MAP Boys β = -0.251, p<0.001 Girls β = -0.159, p<0.001	Age
1267 Buchan (2014);	Cross- sectional	n = 209 range = 15-17.5	Laps	TC/HDL ratio was calculated from measured fasting blood samples	Cardiometabolic	20mSRT performance was not related with TC/HDL ratio	Sex Age PA level

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Scotland		yr age = 16.7 yr		SBP (mmHg) was measured using standard procedures Adiponectin (mg l⁻¹) was measured from fasting blood samples CRP (mg l⁻¹) was measured from fasting blood samples Interleukin-6 (pg ml⁻¹) was measured from fasting blood samples HOMA-IR was calculated from measured insulin and glucose TG (mmol l⁻¹) was measured from fasting blood samples CMR score was calculated from 7 health markers: SBP, TC:HDL ratio, HOMA-IR, TG, CRP, interleukin-6, and Adiponectin		r = 0.093, p = ns 20mSRT performance was not related with SBP r = -0.023, p = ns 20mSRT performance was not related with adiponectin r = 0.079, p = ns 20mSRT performance was not related with CRP r = -0.074, p = ns 20mSRT performance was not related with interleukin-6 r = -0.005, p = ns 20mSRT performance was not related with HOMA-IR r = 0.023, p = ns 20mSRT performance was not related with TG r = -0.085, p = ns 20mSRT performance was not related with CMR r = 0.025, p = ns	Maturation

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was not a predictor of CMR $\beta = -0.004, p = 0.507$	
1292 Ogunleye (2012); England	Cross-sectional	n = 5983 range = 10-16 yr	VO _{2max} Léger Dichotomized using Fitnessgram standards	MAP (mmHg) was calculated from measured blood pressure	Cardiometabolic	Low performance on the 20mSRT was not associated with higher odds of elevated MAP Boys OR 1.02, 95 % CI 0.80 to -1.30 Girls OR 1.03, 95 % CI 0.76 to -1.39	BMI Sex Age
1307 Buchan (2013); Scotland	Cross-sectional	n = 192 range = 14-16 yr age = 15.0 yr	Laps	Adiponectin (mg/l) was measured from fasting blood samples CRP (mg/l) was measured from fasting blood samples Interleukin-6 (pg/ml) was measured from fasting blood samples CMR was calculated from 5 health markers: SBP, HDL, LDL, HOMA-IR, and TG	Cardiometabolic	20mSRT performance was not related with adiponectin $r = 0.08, p = ns$ 20mSRT performance was not related with CRP $r = 0.01, p = ns$ 20mSRT performance was not related with interleukin-6 $r = -0.11, p = ns$ 20mSRT performance was not related with CMR $r = 0.02, p = ns$	Age Sex Maturation

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was a negative predictor of CMR $\beta = -0.078, p < 0.001$	
1337 Moreira (2013); Portugal	Cross-sectional	n = 468 range = 15-18 yr age = 16.5 yr	VO _{2max} Léger Dichotomized using Fitnessgram standards	CMR was calculated from 5 health markers: TC:HDL ratio, TG, HOMA-IR, SBP, Waist-to-height ratio	Cardiometabolic	CMR scores were different between unfit and fit groups p < 0.001	Pubertal status
1342 Jimenez-Pavon (2013); Europe	Cross-sectional	n = 1053 range = 12.5- 17.5 yr age = 14.9 yr	VO _{2max} Léger	Insulin (μ IU/ml) was measured from fasting blood samples QUICKI (insulin sensitivity) was measured from fasting blood samples HOMA-IR was measured from fasting blood samples Glucose (mmol/l) was measured from fasting blood samples	Cardiometabolic	20mSRT performance was negatively related with insulin Boys r = -0.248, p < 0.001 Girls r = -0.29, p < 0.001 20mSRT performance was negatively related with QUICKI Boys r = 0.223, p < 0.001 Girls r = 0.281, p < 0.001 20mSRT performance was negatively related with HOMA-IR Boys r = -0.291, p < 0.001 Girls r = -0.302, p < 0.001 20mSRT performance was negatively related with glucose in boys but	Pubertal status

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						not girls Boys $r = -0.15$, $p < 0.05$ Girls $r = -0.11$, $p = ns$	
1376 Martins (2013); Portugal	Cross- sectional	n = 79 range = 11-13 yr	VO _{2max} Léger Dichotomized using Fitnessgram standards	CRP (mg/l) was measured from fasting blood samples Serum aminotransferase (U/l) was measured from fasting blood samples	Cardiometabolic	CRP levels did not differ between fit and unfit groups $p = ns$ Serum aminotransferase was different between fit and unfit groups $p < 0.01$ Fit individuals had lower odds of having high serum aminotransferase compared to low fit individuals OR 0.03, 95 % CI 0.01- 0.17	Sex Maturation Weight
1387 Brouwer (2013); Netherlands	Cross- sectional	n = 565 range = 11-16 yr T1 age = 11.4 yr T2 age = 16.1 yr	VO _{2max} Léger	CMR was calculated from 5 health markers: WC, glucose, TG, MAP, and HDL CMR was calculated from 4 health markers, excluding WC MAP (mmHg) was calculated according to standard procedures	Cardiometabolic	20mSRT performance was negatively related with CMR Boys $r = -0.253$, $p < 0.01$ Girls $r = -0.174$, $p < 0.01$ 20mSRT performance was a negative predictor of CMR independent of BMI and BMI change in boys but not a predictor in girls	Age Test location Tanner stage Pubertal status

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				DBP (mmHg) was measured according to standard procedures SBP (mmHg) was measured according to standard procedures TG (mmol/l) was measured from fasting blood samples HDL (mmol/l) was measured from fasting blood samples LDL (mmol/l) was measured from fasting blood samples Insulin (mU/l) was measured from fasting blood samples HOMA-IR was calculated using standard procedures		Boys $\beta = -0.014$, 95 % CI -0.023 to -0.005, $p < 0.05$ Girls $\beta = -0.009$, 95 % CI -0.019 to 0.001, $p = ns$ 20mSRT performance was negatively related with CMR in boys but not girls Boys $r = -0.175$, $p < 0.01$ Girls $r = -0.103$, $p = ns$ 20mSRT performance was negatively related with MAP in boys but not girls Boys $r = -0.120$, $p < 0.05$ Girls $r = 0.065$, $p = ns$ 20mSRT performance was negatively related with DBP in boys but not girls Boys $r = -0.156$, $p < 0.01$ Girls $r = 0.071$, $p = ns$ 20mSRT performance was not related with SBP Boys $r = -0.026$, $p = ns$ Girls $r = 0.061$, $p = ns$ 20mSRT performance was negatively related	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						with TG in boys but not girls Boys $r = -0.119$, $p < 0.05$ Girls $r = -0.086$, $p = ns$ 20mSRT performance was positively related with HDL in girls but not boys Boys $r = 0.007$, $p = ns$ Girls $r = 0.159$, $p < 0.01$ 20mSRT performance was not related with LDL Boys $r = -0.107$, $p = ns$ Girls $r = 0.045$, $p = ns$ 20mSRT performance was not related with TC Boys $r = -0.086$, $p = ns$ Girls $r = 0.089$, $p = ns$ 20mSRT performance was negatively related with glucose in boys but not girls Boys $r = -0.166$, $p < 0.01$ Girls $r = -0.042$, $p = ns$ 20mSRT performance was negatively related with insulin Boys $r = -0.314$, $p < 0.01$	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Girls $r = -0.180$, $p < 0.01$ 20mSRT performance was negatively related with HOMA-IR Boys $r = -0.314$, $p < 0.01$ Girls $r = -0.180$, $p < 0.01$ 20mSRT performance was a negative predictor of HOMA-IR Boys $B = -0.021$, 95 % CI -0.03 to -0.012, $p < 0.01$ Girls $B = -0.012$, 95 % CI -0.021 to -0.002, $p < 0.05$	
1408 Marcelino (2012); Portugal	Cross-sectional	n = 4036 range = 10-18 yr	VO _{2max} Dichotomized using Fitnessgram standards	Blood pressure (mmHg) was measured using standard procedures	Cardiometabolic	20mSRT performance was associated with SBP across BMI categories in boys but not girls Boys $p < 0.05$ Girls $p = ns$ 20mSRT performance was not associated with DBP across BMI categories $p = ns$ 20mSRT performance was associated with SBP in girls but not boys	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						across %BF categories Boys p = ns Girls p<0.05 20mSRT performance was associated with DBP in boys and girls across %BF categories Boys p<0.05 Girls p<0.01	
1418 Naidoo (2012); South Africa	Cross-sectional	n = 170 range = grades 3-7 age = 9.4 yr	VO _{2max} Léger Dichotomized by the 50 th percentile	CRP concentration (pg/ml) was calculated as the secretion rate of salivary CRP Salivary CRP rate (pg/min) was calculated by multiplying absolute salivary CRP by salivary flow rate	Cardiometabolic	Those with low 20mSRT performance had greater odds of having high CRP compared to those with high 20mSRT performance OR 3.9, 95 % CI 1.7-8.9, p = 0.001 Those with low 20mSRT performance had greater odds of having a higher salivary CRP rate compared to those with high 20mSRT performance OR 2.7, 95 % CI 1.2-6.1, p = 0.02	BMI BF%
1432 Musa (2012); Nigeria	Cross-sectional	n = 3243 range = 9-15 yr	Laps	Resting SBP and DBP (mmHg) were measured using an automated digital blood pressure	Cardiometabolic	20mSRT performance was not a predictor of SBP in children Boys β = -0.056, p = ns	Age Blood Pressure BMI

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				monitor		Girls $\beta = 0.006$, $p = \text{ns}$ 20mSRT performance was not a predictor of DBP in children Boys $\beta = 0.025$, $p = \text{ns}$ Girls $\beta = 0.033$, $p = \text{ns}$ 20mSRT performance was not a predictor of SBP in adolescents Boys $\beta = -0.012$, $p = \text{ns}$ Girls $\beta = -0.058$, $p = \text{ns}$ 20mSRT performance was not a predictor of DBP in adolescents Boys $\beta = -0.002$, $p = \text{ns}$ Girls $\beta = 0.067$, $p = \text{ns}$	
1433 Llorente-Cantarero (2012); Spain	Cross-sectional	n = 141 range = 7-12 yr	Stages Dichotomized into low 20mSRT performance and high 20mSRT performance groups	SBP and DBP (mmHg) were measured using standardized procedures HDL cholesterol (mg/dl) was measured from fasting blood samples TAG (mg/dl) was measured from fasting blood samples Glucose (mg/dl) was	Cardiometabolic	There were no differences in SBP between low and high 20mSRT performance groups $p = 0.15$ There were no differences in DBP between low and high 20mSRT performance groups $p = 0.82$	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				measured from fasting blood samples Insulin (mU/l) was measured from fasting blood samples HOMA-IR was calculated from measured insulin and glucose Aspartate aminotransferase (U/l) was measured from fasting blood samples CRP (mg/dl) was measured from fasting blood samples LDL (mg/dl) cholesterol was measured from fasting blood samples		HDL cholesterol was higher in the high 20mSRT performance group compared with the low 20mSRT performance group p = 0.001 TAG was higher in the low 20mSRT performance group when compared with the high fitness group p = 0.004 20mSRT performance was negatively related with TAG r = -0.196, p = 0.023 Glucose levels were higher in the low 20mSRT performance group when compared with the high 20mSRT performance group p = 0.033 Insulin levels were higher in the low 20mSRT performance group when compared with the high	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance group p<0.001 HOMA-IR was higher in the low 20mSRT performance group when compared with the high 20mSRT performance group p<0.001 20mSRT performance was negatively related with HOMA-IR r = -0.240, p = 0.05 Aspartate aminotransferase was higher in the high 20mSRT performance group when compared with the low 20mSRT performance group p = 0.001 CRP was higher in the low 20mSRT performance group when compared with the high 20mSRT performance group p = 0.02	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was negatively related with LDL cholesterol $r = -0.227$, $p = 0.008$	
1487 Martinez-Gomez (2012); Spain	Cross-sectional	n = 198 range = 13-17 yr age = 14.8 yr	Laps VO_{2max} Léger Dichotomized using Fitnessgram standards	Adiponectin ($\mu\text{g/ml}$) was measured from fasting blood samples	Cardiometabolic	Adiponectin was higher in the low 20mSRT performance group when compared with the high 20mSRT performance group $p = 0.014$ 20mSRT performance was a negative predictor of adiponectin levels $\beta = -0.193$, $p = 0.028$	Age Sex Pubertal status WC
1526 Artero (2011); Europe	Cross-sectional	n = 709 range = 12.5- 17.5 yr age = 14.9 yr	Stages	SBP and DBP (mmHg) were measured using an automated blood pressure monitor HOMA-IR was calculated from measured insulin and glucose TG (mg/dl) was measured from fasting blood samples TC/HDL-c ratio was calculated from measured values using standard procedures	Cardiometabolic	20mSRT performance was not related with SBP $r = -0.046$, $p = \text{ns}$ 20mSRT performance was negatively related with HOMA-IR $r = -0.265$, $p < 0.001$ 20mSRT performance was negatively related with TG $r = -0.168$, $p < 0.001$ 20mSRT performance was negatively related with TC/HDL ratio	Age Sex Pubertal status Center

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				CMR was measured from 5 health markers: WC, SBP, HOMA-IR, TG and TC:HDL ratio		$r = -0.208, p < 0.001$ 20mSRT performance was negatively related with CMR $r = -0.313, p < 0.001$ 20mSRT performance was a negative predictor of CMR $\beta = -0.264, p < 0.001$ Those in the least fit 20mSRT performance quartile had increased odds of having higher CMR scores compared to the high 20mSRT performance quartile OR 4.3, 95 % CI 2.0-9.3, $p < 0.001$	
1545 Neto (2011); Brazil	Cross-sectional	n = 456 range = 10-18 yr	VO _{2max} Léger	CMR was determined using the age-modified standards of the ATP III MetS criteria which uses 5 health markers: WC, SBP, TG, HDL, and hyperglycemia	Cardiometabolic	A higher prevalence of high CMR was observed in those with low 20mSRT performance when compared with 20mSRT performance groups $p < 0.05$ There were increased odds of having higher	Age Sex

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						MRS in those with low 20mSRT performance compared with those with high 20mSRT performance OR 3.0, 95 % CI 1.13-7.94	
1572 Jimenez-Pavon (2011); Europe	Cross-sectional	n = 1053 range = 12.5-17.5 yr age = 14.9 yr	VO _{2max} Léger	Insulin (μIU/ml) was measured from fasting blood samples Glucose (mmol/l) was measured from fasting blood samples HOMA-IR was calculated from measured insulin and glucose levels	Cardiometabolic	20mSRT performance was negatively related with insulin levels Boys r = -0.23, p<0.01 Girls r = -0.21, p<0.01 20mSRT performance was negatively related with glucose levels in boys but not in girls Boys r = -0.13, p<0.01 Girls r = -0.02, p = ns 20mSRT performance was negatively related with HOMA-IR Boys r = -0.23, p<0.01 Girls r = -0.18, p<0.01	Test center Age Pubertal status Body composition
1582 Puder (2011); Switzerland	Cross-sectional	n = 83 T1 range = 6-7 yr T1 range = 11-12 yr	Stages	CRP (mg/l) was measured from a blood sample. HOMA-IR was calculated using standard procedures	Cardiometabolic	20mSRT performance was a negative predictor of CRP β = -0.16, 95 % CI -0.28 to -0.03, p = 0.02	Age Sex Pubertal status

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was a negative predictor of HOMA-IR $\beta = -0.08$, 95 % CI -0.14 to -0.01, $p = 0.02$	
1657 Martinez-Gomez (2010); Spain	Cross-sectional	n = 192 range = 13-17 yr age = 14.8 yr	Laps	Glucose (mmol/l) was measured from fasting blood samples Insulin (mU/dl) was measured from fasting blood samples HOMA-IR was calculated from measured insulin and glucose levels CRP (mg/l) was measured from fasting blood samples Interleukin-6 (pg/ml) was measured from fasting blood samples Complement factor C3 and complement factor C4 (g/l) were measured from fasting blood samples	Cardiometabolic	20mSRT performance was not related with glucose levels $r = -0.125$, $p = ns$ 20mSRT performance was negatively related with insulin levels $r = -0.151$, $p < 0.05$ 20mSRT performance was negatively related with HOMA-IR $r = -0.158$, $p < 0.05$ 20mSRT performance was negatively related with CRP $r = -0.148$, $p < 0.05$ 20mSRT performance was not a predictor of CRP $\beta = -0.121$, $p = 0.227$ 20mSRT performance	Age Sex Pubertal status

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						was not related with interleukin-6 $r = 0.019$, $p = ns$ 20mSRT performance was not a predictor of interleukin-6 $\beta = 0.045$, $p = 0.691$ 20mSRT performance was negatively related with complement factor C3 $r = -0.323$, $p < 0.001$ 20mSRT performance was a negative predictor of complement factor C3 $\beta = -0.227$, $p = 0.022$ 20mSRT performance was negatively related with complement factor C4 $r = -0.250$, $p < 0.001$ 20mSRT performance was not a predictor of complement factor C4 $\beta = -0.163$, $p = 0.113$	
1677	Cross-	n =	Laps	TC (mg/dl) was	Cardiometabolic	20mSRT performance	Race,

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Jago (2010); USA	sectional	4955 range = 11-12 yr		measured from fasting blood samples HDL cholesterol (mg/dl) was measured from fasting blood samples LDL cholesterol (mg/dl) was measured from fasting blood samples TG (mg/dl) was measured from fasting blood samples SBP and DBP (mmHg) were measured using an automated blood pressure analyzer Glucose (mg/dl) was measured from fasting blood samples Insulin (μU/ml) was measured from fasting blood samples		was not associated with TC Boys $F = 2.00$, $p = 0.09$ Girls $F = 1.43$, $p = 0.22$ 20mSRT performance was not associated with HDL cholesterol Boys $F = 1.85$, $p = 0.12$ Girls $F = 1.05$, $p = 0.38$ 20mSRT performance was associated with LDL cholesterol in boys but not girls Boys $F = 3.48$, $p = 0.008$ Girls $F = 1.96$, $p = 0.10$ 20mSRT performance was associated with TG in girls but not boys Boys $F = 1.10$, $p = 0.36$ Girls $F = 3.15$, $p = 0.01$ 20mSRT performance was associated with SBP Boys $F = 6.17$, $p < 0.0001$ Girls $F = 2.11$, $p = 0.08$ 20mSRT performance was associated with DBP in girls but not boys Boys $F = 0.82$, $p = 0.51$	Household education Pubertal stage BMI

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Girls $F = 2.54$, $p = 0.04$ 20mSRT performance was not associated with glucose Boys $F = 0.60$, $p = 0.66$ Girls $F = 0.50$, $p = 0.74$ 20mSRT performance was associated with insulin Boys $F = 7.35$, $p < 0.0001$ Girls $F = 6.50$, $p < 0.0001$	
1735 Martinez-Gomez (2009); Spain	Cross-sectional	n = 202 range = 13-17 yr	Laps	SBP and DBP (mmHg) were measured using an automated blood pressure monitor TC (mg/dl) was measured from fasting blood samples HDL cholesterol (mg/dl) was measured from fasting blood samples TG (mg/dl) was measured from fasting blood samples Glucose (mg/dl) was measured from fasting	Cardiometabolic	20mSRT performance was not related with SBP $r = -0.014$, $p = \text{ns}$ 20mSRT performance was not related with DBP $r = -0.129$, $p = \text{ns}$ 20mSRT performance was not related with TC $r = -0.019$, $p = \text{ns}$ 20mSRT performance was not related with HDL cholesterol $r = 0.085$, $p = \text{ns}$ 20mSRT performance was negatively related	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				blood samples Insulin (mU/dl) was measured from fasting blood samples CMR was calculated from using methods from three different studies		with TG $r = -0.179$, $p < 0.05$ 20mSRT performance was not related with glucose levels $r = -0.098$, $p = ns$ 20mSRT performance was negatively related with insulin $r = -0.194$, $p < 0.05$ 20mSRT performance was inversely associated with all three CMR scores Corpus Christi Study $\beta = -0.322$, $p = 0.002$ Aerobic Center Study $\beta = -0.216$, $p = 0.016$ European Youth Heart Study $\beta = -0.357$, $p = 0.001$	
1742 Telford (2009); Australia	Cross-sectional	N = 498 T1 age = 8 yr	Stages	HOMA-IR was calculated from measured fasting insulin and glucose	Cardiometabolic	20mSRT performance was inversely associated with HOMA-IR at T1 Boys $p = 0.02$ Girls $p < 0.001$	n/a
1781 Sakuragi (2008);	Cross-sectional	n = 573 range = 10 yr	Stages	Pulse wave velocity (m/s) was assessed non-invasively using the	Cardiometabolic	20mSRT performance was negatively related with pulse wave velocity	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Australia		age = 10.1yr		Sphygmocor system		$r = -0.233, p < 0.0001$ 20mSRT performance was a negative predictor of pulse wave velocity $\beta = -0.047, 95\% \text{ CI } -0.07 \text{ to } -0.024, p < 0.0001$	
1797 Hoekstra (2008); Ireland	Cross-sectional	n = 2016 range = 12 and 15 yr	VO _{2max}	Resting SBP and DBP (mmHg) were measured manually TC (mmol/l) was measured from fasting blood samples HDL cholesterol (mmol/l) was measured from fasting blood samples	Cardiometabolic	20mSRT performance was a negative predictor of SBP in 12 year old boys but not girls Boys $\beta = -0.291, 95\% \text{ CI } -0.478 \text{ to } -0.104, p = 0.002$ Girls $\beta = -0.127, 95\% \text{ CI } -0.305 \text{ to } 0.051, p = 0.162$ 20mSRT performance was a negative predictor of SBP in 15 year olds Boys $\beta = -0.242, 95\% \text{ CI } -0.450 \text{ to } -0.034, p = 0.027$ Girls $\beta = -0.199, 95\% \text{ CI } -0.374 \text{ to } -0.023, p = 0.027$ 20mSRT performance was a negative predictor of DBP in 12 year olds Boys $\beta = -0.201, 95\% \text{ CI } -0.346 \text{ to } -0.057, p =$	Maturation SES PA score Muscular fitness

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						0.006 Girls $\beta = -0.151$, 95 % CI -0.284 to -0.018, $p = 0.026$ 20mSRT performance was a negative predictor of DBP in 15 year old girls but not boys Boys $\beta = -0.090$, 95 % CI -0.228 to 0.047, $p = 0.196$ Girls $\beta = -0.220$, 95 % CI -0.357 to -0.084, $p = 0.002$ 20mSRT performance was a negative predictor of TC in 12 year olds Boys $\beta = -0.012$, 95 % CI -0.024 to -0.001, $p = 0.039$ Girls $\beta = -0.012$, 95 % CI -0.024 to 0.000, $p = 0.042$ 20mSRT performance was a negative predictor of TC in 15 year old boys but not girls Boys $\beta = -0.024$, 95 % CI -0.034 to -0.013, $p = 0.039$ Girls $\beta = -0.004$, 95 % CI	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						-0.015 to 0.007, $p = 0.457$ 20mSRT performance was a positive predictor of HDL cholesterol in 12 year olds Boys $\beta = 0.008$, 95 % CI 0.003-0.013, $p = 0.003$ Girls $\beta = 0.006$, 95 % CI 0.001-0.010, $p = 0.018$ 20mSRT performance was a positive predictor of HDL cholesterol in 15 year olds Boys $\beta = 0.005$, 95 % CI 0.00-0.01, $p = 0.043$ Girls $\beta = 0.008$, 95 % CI 0.003 0.013, $p = 0.002$ 20mSRT performance was negative predictor of TC:HDL ratio in 12 year olds Boys $\beta = -0.033$, 95 % CI -0.046 to -0.019, $p < 0.001$ Girls $\beta = -0.025$, 95 % CI -0.038 to -0.011, $p < 0.001$ 20mSRT performance was a negative predictor of TC:HDL ratio in 15 year olds	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Boys $\beta = -0.036$, 95 % CI -0.051 to -0.021, $p < 0.001$ Girls $\beta = -0.021$, 95 % CI -0.035 to -0.007, $p = 0.003$	
1885 Garcia-Artero (2007); Spain	Cross- sectional	n = 460 range = 13-18.5 yr	VO _{2max} Léger	CMR was calculated from 4 health markers: TG, HDL, LDL, and glucose	Cardiometabolic	Lipid CMR was different between 20mSRT performance tertiles for boys but not girls Boys 1 st vs 3 rd , $p = 0.003$; 2 nd vs 3 rd , $p = 0.027$ Girls $p = ns$	Age Maturation PA Muscular strength
1910 Ruiz (2007); Spain	Cross- sectional	n = 156 range = 13-18.5 yr age = 14.8 yr	VO _{2max} Léger	Homocysteine (mg/l) was measured from fasting blood samples	Cardiometabolic	20mSRT performance was negatively related with homocysteine in girls but not boys Boys $r = 0.05$, $p = ns$ Girls $r = -0.38$, $p < 0.02$ 20mSRT performance was a negative predictor of homocysteine in girls but not boys Boys $\beta = 0.12$, 95 % CI - 0.03 to 0.06, $p = 0.51$ Girls $\beta = -0.40$, 95 % CI - 0.11 to -0.02, $p = 0.007$	Age Pubertal stage Birth weight Smoking SES Sum of SF Serum folic acid Serum vitamin B12 Genotype marker
1952 Mesa (2006); Spain	Cross- sectional	n = 2090 range = 13-18.5 yr	Stages	TC (mg/dl) was measured from fasting blood samples TG (mg/dl) was	Cardiometabolic	20mSRT performance was not associated with TC Boys p for trend = 0.915 Girls p for trend = 0.903	Age Maturation SES

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				measured from fasting blood samples LDL cholesterol (mg/dl) was measured from fasting blood samples HDL cholesterol (mg/dl) was measured from fasting blood samples Apolipoprotein A (mg/dl) was measured from fasting blood samples Apolipoprotein B (mg/dl) was measured from fasting blood samples Lipoprotein A (mg/dl) was measured from fasting blood samples CMR was calculated from 4 health markers: TG, HDL, LDL, and glucose		20mSRT performance was associated with TG in boys but not in girls Boys p for trend = 0.004 Girls p for trend = 0.191 20mSRT performance was not associated with LDL cholesterol Boys p for trend = 0.202 Girls p for trend = 0.415 20mSRT performance was associated with HDL cholesterol Boys p for trend = 0.013 Girls p for trend = 0.045 20mSRT performance was associated with Apo A in boys but not girls Boys p for trend = 0.028 Girls p for trend = 0.173 20mSRT performance was not associated with Apo B Boys p for trend = 0.840 Girls p for trend = 0.261 20mSRT performance was not associated with	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						lipoprotein A Boys p for trend = 0.773 Girls p for trend = 0.573 20mSRT performance was inversely associated with CMR score in children Normal weight p for trend<0.05 Overweight p for trend<0.05	
2070 Nielsen (2003); International	Cross-sectional	n = 13557 range = 15-20 yr age = 17.2 yr	VO _{2max} Léger	SBP and DBP (mmHg) were measured using both manual and automated techniques	Cardiometabolic	Those in the lowest 20mSRT performance quintile had increased odds of having hypertension compared with those in the higher 20mSRT performance group Boys OR 1.3, 95 % CI 1.1-1.7, p = 0.01 Girls OR 1.5, 95 % CI 1.3-1.8, p<0.00	Age BMI Daily smoking Hash smoking Alcohol Sweets Coffee
2091 Guerra (2002); Portugal	Cross-sectional	n = 529 range = 8-15 yr	VO _{2max} Léger	SBP and DBP (mmHg) were measured using automated blood pressure monitors	Cardiometabolic	20mSRT performance was not associated with SBP Boys $F = 0.313$, p = ns Girls $F = 0.426$, p = ns 20mSRT performance was not associated with	Age Weight Height Fat free mass Fat percentage

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						DBP Boys $F = 0.020$, $p = ns$ Girls $F = 0.125$, $p = ns$	
2118 Boreham (2001); Ireland	Cross- sectional	n = 1015 range = 12 and 15 yr	Laps	SBP and DBP (mmHg) were measured using manual techniques TC (mmol/l) was measured from non- fasting blood samples HDL cholesterol (mmol/l) was measured from non-fasting blood samples TC:HDL ratio was calculated from non- fasting blood samples	Cardiometabolic	20mSRT performance was not a predictor of SBP in 12 year olds Boys $\beta = 0.01$, $p = ns$ Girls $\beta = 0.11$, $p = ns$ 20mSRT performance was not a predictor of SBP in 15 year olds Boys $\beta = 0.03$, $p = ns$ Girls $\beta = -0.09$, $p = ns$ 20mSRT performance was not a predictor of DBP in 12 year olds Boys $\beta = 0.11$, $p = ns$ Girls $\beta = 0.02$, $p = ns$ 20mSRT performance was not a predictor of DBP in 15 year olds Boys $\beta = 0.09$, $p = ns$ Girls $\beta = -0.11$, $p = ns$ 20mSRT performance was not a predictor of TC in 12 year olds Boys $\beta = -0.11$, $p = ns$ Girls $\beta = 0.12$, $p = ns$	Social class Maturation Fatness

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was not a predictor of TC in 15 year olds Boys $\beta = -0.10$, $p = ns$ Girls $\beta = 0.04$, $p = ns$ 20mSRT performance was not a predictor of HDL cholesterol in 12 year olds Boys $\beta = -0.05$, $p = ns$ Girls $\beta = 0.08$, $p = ns$ 20mSRT performance was not a predictor of HDL cholesterol in 15 year olds Boys $\beta = 0.09$, $p = ns$ Girls $\beta = -0.04$, $p = ns$ 20mSRT performance was not a predictor of TC:HDL cholesterol ratio in 12 year olds Boys $\beta = -0.04$, $p = ns$ Girls $\beta = 0.03$, $p = ns$ 20mSRT performance was not a predictor of TC:HDL cholesterol ratio in 15 year olds Boys $\beta = -0.13$, $p = ns$	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Girls $\beta = 0.07$, $p = ns$	

*Note 20mSRT, 20 meter shuttle run test; BP, blood pressure; CMR, cardiometabolic risk; CRP, c-reactive protein; DBP, diastolic blood pressure; HDL, high density lipoprotein; HOMA-IR, homeostatic model assessment – insulin resistance; LDL, low density lipoprotein; MAP, mean arterial pressure; SBP, systolic blood pressure; TC, total cholesterol; TG, triglycerides

Supplement Table 3. 20mSRT performance and cognition/academic achievement in children and youth

Note: β indicates standardized Beta; B indicates unstandardized Beta; ***Indicates questionnaire has good psychometric properties

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1591 Niederer (2011); Switzerland	Longitudinal (9 month follow-up)	n = 245 range = 4-6 yr age = 5.2 yr	Stages	Spatial working memory was measured using a validated intelligence scale from the Intelligence and Development Scales *** Attention was measured using the validated KHV-VK test for preschoolers ***	Cognition	20mSRT performance was not associated with spatial working memory at follow-up $\beta = 0.08$, 95 % CI -0.09 to 0.24, p = 0.37 20mSRT performance was positively associated with attention at follow-up $\beta = 0.25$, 95 % CI 0.03-0.48, p = 0.03	Sex Age BMI Migration status Parental education Native language Linguistic region
211 Williams (2013); USA	Cross- sectional	n = 171 range = 17-18 yr age = 17.3 yr	VO2max Fitnessgram	Health knowledge was measured using FitSmart	Cognition	20mSRT performance was associated with Health knowledge (r = 0.26, p = 0.001) 20mSRT performance was a positive predictor of Health knowledge $\beta = 0.56$, p<0.001	Age Sex
280 Du Toi (2011); South Africa	Cross- sectional	n = 212 range = 9-12 yr age = 10.7 yr Boys = 10.8 yr	Laps	Academic achievement was determined by end- year averages from 8 courses	Cognition	20mSRT performance was positively related with academic achievement in girls but not boys Boys r = 0.12, p = ns Girls r = 0.21, p<0.05	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
		Girls = 10.6 yr					
310 Joshi (2011); USA	Cross- sectional	n = 19695 range = Gr. K- 12	VO2max Fitnessgram	Standardized state-wide tests for math and English-Language Arts	Cognition	20mSRT performance was positively related to English Language Arts score $X^2 = 139.5$, $p < 0.001$ 20mSRT performance was positively related to math score $X^2 = 98.6$, $p < 0.001$	n/a
606 Wittberg (2010); USA	Cross- sectional	n = 1941 range = 9-13 yr	Laps	Academic achievement was assessed by state- wide standardized reading, math, science and social studies tests (WESTEST)	Cognition	20mSRT performance was positively associated with reading score in girls but not boys Boys $r = 0.06$, $p = ns$ Girls $r = 0.16$, $p < 0.001$ 20mSRT performance was positively associated with math score in girls but not boys Boys $r = 0.04$, $p = ns$ Girls $r = 0.16$, $p < 0.001$ 20mSRT performance was positively associated with science score in girls but not boys Boys $r = -0.06$, $p = ns$ Girls $r = 0.10$, $p < 0.01$ 20mSRT performance was positively associated with social studies score in girls but not boys Boys $r = -0.03$, $p = ns$	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Girls $r = 0.12$, $p < 0.01$	
1179 Hansen (2014); USA	Cross- sectional	n = 687 age = 7.8	Laps	Academic achievement was assessed using the Weschler Individual Achievement Test (3 rd ed) to measure reading, spelling and mathematics	Cognition	20mSRT performance was not a predictor of reading scores $B = 0.06$, 95 % CI -0.10 to 0.22, $p = 0.463$ 20mSRT performance was a positive predictor of spelling scores $B = 0.32$, 95 % CI 0.14 to 0.50, $p = 0.001$ 20mSRT performance was a positive predictor of mathematics score $B = 0.55$, 95 % CI 0.37 to 0.72, $p < 0.001$	Grade Sex Race Ethnicity Mother's education level Household income BMI Test site
1216 Sardinha (2014); Portugal	Cross- sectional	n = 1531 range = 12-14 yr age = 12.3 yr	VO2max Fitnessgram	Academic achievement was assessed by using end of year grades	Cognition	Fit students from 20mSRT performance had increased odds of having higher academic achievement compared to unfit students High AA OR 2.27, 95 % CI 1.46-3.52 $p < 0.001$	Sex Weight status Cohort
1237 Scudder (2013); USA	Cross- sectional	n = 397 range = 6-9 yr age = 7.6 yr	Laps	Inhibitory control was measured using the Flanker task *** Memory was assessed using the spatial n-back	Cognition	20mSRT performance was negatively associated with compatible reaction time, both congruent and incongruent tests Cong. $\beta = -0.20$, $p < 0.01$ InCong. $\beta = -0.18$, $p < 0.01$	Grade Sex Household income BMI

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				task ***		20mSRT performance was positively associated with compatible accuracy, both congruent and incongruent tests Cong. $\beta = 0.12$, $p < 0.05$ InCong. $\beta = 0.11$, $p < 0.05$ 20mSRT performance was negatively associated with incompatible reaction time, both congruent and incongruent tests Cong. $\beta = -0.12$, $p < 0.05$ InCong. $\beta = -0.11$, $p < 0.05$ 20mSRT performance was positively associated with incompatible congruent accuracy, but not incongruent accuracy Cong. $\beta = 0.13$, $p < 0.05$ InCong. $\beta = 0.08$, $p = ns$ 20mSRT performance was positively associated with working memory during the 1- and 2-back tasks 1-back $\beta = 0.17$, $p < 0.01$ 2-back $\beta = 0.19$, $p < 0.01$	
1316 Rauner (2013);	Cross- sectional	n = 11743 range = Gr. 4- 11	VO2max Fitnessgram	Academic achievement was assessed using state- wide math and reading	Cognition	Unfit students from 20mSRT performance had lower odds of passing the math test	SES Sex Grade level

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
USA				tests		OR 0.70, $p < 0.05$ Unfit students from 20mSRT performance had lower odds of passing the reading test OR 0.70, $p < 0.05$	Ethnicity School
1334 Bass (2013) USA	Cross- sectional	n = 838 range = Gr. 6- 8 age = 13.1 yr	Laps	Academic achievement was assessed using state- wide math and reading tests	Cognition	20mSRT performance was positively related to reading test scores Boys $r = 0.12$, $p < 0.05$ Girls $r = 0.20$, $p < 0.05$ 20mSRT performance was positively related to math test score Boys $r = 0.16$, $p < 0.05$ Girls $r = 0.27$, $p < 0.05$ Children with high 20mSRT performance had increased odds of meeting or exceeding reading test score cut-offs Boys OR 2.81, 95 % CI 1.56-5.07 Girls OR 2.58, 95 % CI 1.39-4.78 Children with high 20mSRT performance had increased odds of meeting or exceeding math test score cut-offs Boys OR 3.15, 95 % CI 1.90-5.21 Girls OR 2.08, 95 % CI 1.10-3.93	SES Age
1335	Cross-	n = 1701	n/a	Academic achievement	Cognition	20mSRT performance was not	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Coe (2013); USA	sectional	range = Gr. 3,6 and 9		was assessed by state- wide math, English and social studies tests		related to math score in Grade 6 or Grade 9 Gr. 6 $r = 0.07$, $p = ns$ Gr. 9 $r = 0.02$, $p = ns$ 20mSRT performance was not related to English score in Grade 6 or Grade 9 Gr. 6 $r = 0.03$, $p = ns$ Gr. 9 $r = 0.04$, $p = ns$ 20mSRT performance was not related to social studies score in Grade 6 or Grade 9 Gr. 6 $r = 0.06$, $p = ns$ Gr. 9 $r = 0.07$, $p = ns$	
1407 Coe (2012); USA	Cross- sectional	n = 312 range = Gr. 6- 8 age = 12.1 yr	Laps	Academic achievement was assessed using end of year grades from 4 classes	Cognition	20mSRT performance was positively related with student grades $r = 0.22$, $p < 0.01$	n/a
1482 Telford (2012); Australia	Cross- sectional	n = 757 range = 8-10 yr	Stages	Academic achievement was determined by government mandated reading, numeracy and writing tests	Cognition	20mSRT performance was not associated with child reading scores $\beta = 2.0$, $SE = 8.7$, $p = 0.82$ 20mSRT performance was positively associated with child numeracy scores $\beta = 17.3$, $SE = 7.7$, $p = 0.03$ 20mSRT performance was	PA level BF % SES

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						positively associated with child writing scores $\beta = 14.9$, $SE = 7.0$, $p = 0.03$	
1502 Padilla- Moledo (2011) Spain	Cross- sectional	n = 684 range = 6-17.9 yr Children = 6- 11.9 yr Adolescents = 12-17.9 yr	VO2max Leger	Academic achievement was self-reported on a Likert scale using the HSBC questionnaire	Cognition	20mSRT performance was not associated with Academic Achievement in children or adolescents Children $\beta = 0.09$, $p = 0.09$ Adolescents $\beta = -0.08$, $p = 0.14$	Sex
1850 Buck (2007); USA	Cross- sectional	n = 74 range = 7-12 yr age = 9.3 yr	Laps	Concentration was measured by the Stroop Colour and Word test Executive function was measured using the Kaufman Brief Intelligence Test	Cognition	20mSRT performance was positively related with the Stroop word, colour, and colour/word tests but not the interference test. Word $r = 0.39$, $p < 0.05$ Colour $r = 0.37$, $p < 0.05$ Colour/Word $r = 0.37$, $p < 0.05$ Interference $r = 0.04$, $p = ns$ 20mSRT performance was not correlated with IQ $r = 0.09$, $p = ns$ Higher 20mSRT performance was associated with better performance on the word, colour and colour/word tests but not the interference test Word $B = 0.38$, $SE = 0.11$, $\beta =$ 0.34	Age IQ BMI

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						$p < 0.01$ Colour B = 0.20, SE = 0.09, β = 0.25 $p < 0.05$ Colour/Word B = 0.24, SE = 0.09, β = 0.30, $p < 0.01$ Interference = no data, p = ns	

Note* 20mSRT, 20 m shuttle run test

Supplement Table 4. 20mSRT performance and mental health in children and youth

Note: β indicates standardized Beta; ***Indicates questionnaire has good psychometric properties

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1344 Rieck (2013); USA	Cross-sectional	n = 986 range = 11-15 yr age = 12.5 yr	VO2max Fitnessgram® HFZ	Depression was self-reported via the CES-DC***	Mental Health	Low 20mSRT performance was not associated with higher odds of depression Boys OR 1.63, 95 % CI 0.08-3.17 Girls OR 2.09, 95 % CI 0.96-4.57	Age SES Ethnicity Other fitness tests
1640 Greenleaf (2010); USA	Cross-sectional	n = 1022 range = 11-15 yr Boy age = 12.5 yr Girl age = 12.4 yr	Fitnessgram® HFZ	Depression was self-reported via the CES-DC***	Mental Health	Girls in the HFZ of 20mSRT performance reported lower levels of depression but no relationship was observed in boys Boys $F = 5.14$, $r = -0.25$, $p = ns$ Girls $F = 8.64$, $r = -0.54$, $p = 0.01$	SES

Note* 20mSRT, 20 m shuttle run test; CES-DEC, Center for Epidemiology Studies–Depression Scale for Children

Supplement Table 5. 20mSRT performance and psychosocial health in children and youth

Note: β indicates standardized Beta; ***Indicates questionnaire has good psychometric properties

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
116 Zhu (2015); USA	Cross- sectional	n = 141 range = 11-15 yr age = 12.8 yr	Laps	Self-efficacy was self-reported using an adapted questionnaire***	Psychosocial health	20mSRT performance was positively related to self-efficacy $r = 0.71, p < 0.01$ Self-efficacy was positively associated with 20mSRT performance $\beta = 0.45, p < 0.05$	n/a
298 Colquitt (2012); USA	Cross- sectional	n = 122 range = 11-15 yr age = 12.8 yr	VO2max Fitnessgram®	Enjoyment was self-reported via questionnaire***	Psychosocial health	20mSRT performance was positively related to enjoyment $r = 0.41, p < 0.01$ 20mSRT performance was a positive predictor of enjoyment $\beta = 0.41, p < 0.001$	n/a
438 Gao (2009); USA	Cross- sectional	n = 252 range = 12-15 yr age = 13.2 yr	VO2max Fitnessgram®	Self-efficacy was self-reported via an adapted questionnaire*** Outcome expectancy was self-reported using an adapted questionnaire***	Psychosocial health	20mSRT performance was positively related to self-efficacy $r = 0.37, p < 0.01$ 20mSRT performance was positively related with outcome expectancy $r = 0.23, p < 0.01$	n/a
1278 Gao (2013); USA	Cross- sectional	n = 276 range = 12-15 yr age = 13.3 yr	Laps	Self-efficacy was self-reported using an adapted questionnaire***	Psychosocial health	20mSRT performance was positively related to self-efficacy $r = 0.27, p < 0.01$	n/a

1304 Morales (2013); Spain	Cross- sectional	n = 1158 range = 8-11 yr Boy age = 9.5 yr Girl age = 9.5 yr	Stages	Moods and emotions were self-reported using the KIDSCREEN-52 questionnaire***	Psychosocial health	20mSRT performance was positively associated with moods and emotions in girls but not boys Boys $\beta = 0.05$, $p = 0.43$ Girls $\beta = 0.14$, $p = 0.009$	Age
-------------------------------------	---------------------	--	--------	--	------------------------	---	-----

Note* 20mSRT, 20 meter shuttle run test

Supplement Table 6. 20mSRT performance and self-esteem and physical self-perception in children and youth

Note: β indicates standardized Beta; B indicates unstandardized Beta; ***Indicates questionnaire has good psychometric properties

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
3075 Olive (2012); Australia	Longitudinal (4 yr. follow- up)	n = 821 range = 7- 9 yr baseline age = 8.2 yr follow-up boy age = 12.09 yr follow-up girl age = 12.05 yr	Square Root of CRF	Body dissatisfaction was measured from an adapted version of the Body Self- Esteem Scale for children***	Self- esteem/Physical self-perception	20mSRT performance at baseline was not associated with body dissatisfaction at 4-yr follow up Boys β = -4.67, SE = 3.33, p = 0.16 Girls β = -4.96, SE = 3.95, p = 0.21	BF%
298 Colquitt (2012); USA	Cross- sectional	n = 122 range = 11-15 yr age = 12.8 yr	PACER VO _{2max} Fitnessgram	Perceived usefulness was measured using a published survey***	Self-esteem/ Physical self- perception	20mSRT performance was positively related with perceived usefulness r = 0.35, p<0.01 20mSRT performance was a positive predictor of perceived usefulness β = 0.27, p = 0.005	n/a
650 Gao (2008); USA	Cross- sectional	n = 307 range = 12-15 yr age = 13.4 yr	PACER VO _{2max} Fitnessgram	Perceived sport competence was measured using the Self-Perceived Competence in Physical Education Scale***	Self-esteem/ Physical self- perception	20mSRT performance was a positive predictor of perceived sport competence β = 0.44, p<0.01	n/a
1175 Riiser (2014);	Cross- sectional	n = 120 age =	Speed at the last completed	Self-determined motivation was measure by a translated	Self-esteem/ Physical self-	20mSRT performance was a positive predictor	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
Norway		13.7 yr	stage	version of the Behavioural Regulation in Exercise-2 questionnaire***	perception	of self-determined motivation B = 2.08, p<0.01	
1304 Morales (2013); Spain	Cross- sectional	n = 1158 range = 8- 11 yr	Half stages	Self-perception and autonomy were measured using the KIDSCREEN-52 questionnaire***	Self-esteem/ Physical self- perception	20mSRT performance was a positive predictor of self-perception in girls but not boys Boys β = 0.036, p = 0.528 Girls β = 0.131, p = 0.014 20mSRT performance was not a predictor of autonomy Boys β = 0.0045, p = 0.418 Girls β = 0.045, p = 0.418	Age BMI Muscular fitness
1640 Greenleaf (2012); USA	Cross- sectional	n = 1022 range = 11-15 yr	PACER Fitnessgram Healthy Fitness Zone	Self-esteem and self-concept were measured using the Self-Description questionnaire II*** Body satisfaction was measured using the BPSS- R questionnaire***	Self-esteem/ Physical self- perception	Those in the healthy fitness zone for 20mSRT performance reported higher self-esteem for girls but not boys Boys F = 5.24, r = 0.25, p = ns Girls F = 8.46, r = 0.53, p = 0.01 Those in the healthy fitness zone for 20mSRT performance reported higher self-perceived	SES

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						strength Boys $F = 10.21$, $r = 0.35$, $p = 0.001$ Girls $F = 6.87$, $r = 0.48$, $p = 0.01$ There were no differences in perceived flexibility between healthy fitness and needs improvement zones for 20mSRT performance Boys $F = 5.04$, $r = 0.24$, $p = \text{ns}$ Girls $F = 5.70$, $r = 0.44$, $p = \text{ns}$ Those in the healthy fitness zone for 20mSRT performance reported higher self-perceived endurance Boys $F = 80.18$, $r = 0.97$, $p = 0.0001$ Girls $F = 26.14$, $r = 0.93$, $p = 0.0001$ Those in the healthy fitness zone for 20mSRT performance reported higher body satisfaction Boys $F = 8.29$, $r = 0.31$, $p = 0.01$ Girls $F = 8.41$, $r = 0.53$, p	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						= 0.01	
1846 Cairney (2008); Canada	Cross- sectional	n = 2245 range = 9- 10 yr age = 9.4 yr	Stages	Perceived adequacy, predilection, and enjoyment in physical education was measured using the CSAPPA questionnaire***	Self-esteem/ Physical self- perception	20mSRT performance was a positive predictor of perceived adequacy $\beta = 0.080$, SE = 0.187, p<0.001 20mSRT performance was a positive predictor of predilection $\beta = 0.033$, SE = 0.080, p<0.05 20mSRT performance was not a predictor of enjoyment in physical education $\beta = 0.001$, SE = 0.001, p = ns	n/a
1865 Shen (2007); USA	Cross- sectional	n = 653 range = 11-15 yr age = 12.4 yr	Laps	Perceived competence was measured using the Intrinsic Motivation Inventory questionnaire***	Self-esteem/ Physical self- perception	20mSRT performance was positively related with perceived competence in physical education r = 0.21, p<0.01	n/a
2043 Jurimae (2003); Estonia	Cross- sectional	n = 525 range = 10-17 yr	Minutes	Perceived endurance was measured using a self- report questionnaire	Self-esteem/ Physical self- perception	20mSRT performance was positively related with perceived endurance range of r = 0.31–0.51, p<0.05	n/a
3075 Olive (2012); Australia	Cross- sectional	n = 821 range = 7- 9 yr	Square Root of CRF	Body dissatisfaction was measured from an adapted version of the Body Self-	Self- esteem/Physical self-perception	Low 20mSRT performance was associated with greater	BF%

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
		age = 8.2 yr		Esteem Scale for children***		body dissatisfaction Boys $\beta = -9.87$, SE = 2.60, p<0.001 Girls $\beta = -11.03$, SE = 3.47, p = 0.002	

Note* 20mSRT, 20 meter shuttle run test; BPSS-R, Body Parts Satisfaction Scale–Revised; CSAPPA, Children's Self-Perceptions of Adequacy in and Predilection for Physical Activity

Supplement Table 7. 20mSRT performance and quality of life and wellbeing in children and youth

***Indicates questionnaire has good psychometric properties (i.e., valid, reliable, appropriate for age group tested).

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
220 Borras (2011); Spain	Cross- sectional	n = 302 range = 11-12 yr Boy age = 11.6 yr Girl age = 10.7 yr	Decimal Stages	Satisfaction with health, Physical Comfort, Emotional Comfort and Restricted Activity were self-reported using the Spanish version of the CHIPCE/PRF***	Quality of Life & Wellbeing	20mSRT performance was not related to satisfaction with health $r = 0.053$, $p = 0.37$ 20mSRT performance was negatively related to physical comfort $r = -0.13$, $p = 0.03$ 20mSRT performance was not related to emotional comfort $r = -0.08$, $p = 0.19$ 20mSRT performance was not related to restricted activity $r = -0.01$, $p = 0.86$	n/a
1175 Riiser (2014); Norway	Cross- sectional	n = 120 age = 13.7 yr	Speed	Health related quality of life was self-reported using the Norwegian version of the KIDSCREEN***	Quality of Life & Wellbeing	20mSRT performance was a positive predictor of HRQoL $\beta = 4.16$, 95 % CI 0.3-8.02, $p < 0.05$	n/a
1304 Morales (2013); Spain	Cross- sectional	n = 1158 range = 8-11 yr Boy age = 9.5 yr Girl age = 9.5 yr	Stages	Physical wellbeing and psychological wellbeing were self-reported using the KIDSCREEN-52***	Quality of Life & Wellbeing	20mSRT performance was positively associated with physical wellbeing in girls but not boys Boys $\beta = 0.09$, $p = 0.12$ Girls $\beta = 0.16$, $p = 0.001$ 20mSRT performance was	Age BMI Muscular fitness

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						positively associated with psychological wellbeing in girls but not boys Boys $\beta = 0.01$, $p = 0.82$ Girls $\beta = 0.11$, $p = 0.03$	
1456 Catro-Pinero (2012); Spain	Cross- sectional	n = 691 range = 12-17 yr Child = 9.3 yr Adol = 14.8 yr	VO2max Leger	Health complaints were self- reported using the HBSC questionnaire***	Quality of Life & Wellbeing	20mSRT performance was inversely associated with health complaints Child $\beta = -0.15$, $p = 0.005$ Adolescents $\beta = -0.18$, $p = 0.002$ Low 20mSRT performance was associated with higher odds of having health complaints Child OR 2.56, 95 % CI 1.30-5.03 Adolescent OR 1.99, 95 % CI 1.16-3.43	Sex
1478 Mota (2012); Portugal	Cross- sectional	n = 553 range = 12-18 yr age = 15.4 yr	Laps	Self-rated health was self- reported answering one Likert scale question	Quality of Life & Wellbeing	Girls with positive self-rated health completed more laps on the 20mSRT than negative raters $p < 0.001$ Unfit girls had higher odds of reporting negative self-rated health OR 2.93, 95 % CI 1.82-4.72, $p < 0.001$	BMI
1502 Padilla- Moledo (2011); Spain	Cross- sectional	n = 684 Child = 6-11.9 yr Adol = 12-17.9 yr	VO2max Leger HFZ	Perceived health status and life satisfaction were self-reported using the HBSC questionnaire [only valid for adolescents]	Quality of Life & Wellbeing	20mSRT performance was positively associated with perceived health in adolescents but not children Child $\beta = 0.08$, $p = 0.11$	Sex

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Adolescents $\beta = 0.15$, $p = 0.008$ 20mSRT performance was positively associated with life satisfaction Child $\beta = 0.14$, $p = 0.006$ Adolescents $\beta = 0.17$, $p = 0.002$ Adolescents with poor 20mSRT performance were more likely to report their health as 'fair' as opposed to 'excellent'. No increased odds were observed in children Child OR 3.77, 95 % CI 0.82-17.36 Adol OR 2.76, 95 % CI 1.06-7.20 Adolescents with poor 20mSRT performance had higher odds of reporting 'low' life satisfaction Child OR 1.77, 95 % CI 0.93-3.37 Adol OR 1.66, 95 % CI 1.03-2.68	

Note* 20mSRT, 20 meter shuttle run test; CHIPCE/PRF, Children Health and Illness Profile-Children Edition/Parent Report Form; HRQoL, health related quality of life; HBSC, Health Behaviour in School-aged Children

Supplement Table 8. 20mSRT performance and bone health in children and youth

Note: β indicates standardized Beta; B indicates unstandardized Beta

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1836 Vincente- Rodriguez (2008); Spain	Cross- sectional	n = 278 range = 13-18.5 yr	VO _{2max} Léger	Bone mineral content (g) was measured using a DXA scan	Bone Health	20mSRT performance was a positive predictor of bone mineral content Boys B = 0.050*, SE = 0.238, sr = 0.426, p<0.0001 Girls B = 0.041*, SE = 0.181, sr = 0.359, p<0.0001	Height Maturity

Note* 20mSRT, 20 m shuttle run test; DXA, dual energy x-ray absorptiometry.

* indicates unstandardized Beta

Supplement Table 9. 20mSRT performance and musculoskeletal fitness in children and youth

Note: β indicates standardized Beta

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
292 Klincarov (2011); Macedonia	Cross-sectional	n = 156 range = 14 yr age = 14.0 yr	Stages	Standing broad jump (cm) was used to measure strength Hand grip dynamometry (kg) was used to measure strength 30 sec sit-up (#) was used to measure muscular endurance Bent arm hang (sec) was used to measure muscular endurance 10x5m shuttle run (sec) was used to measure agility	Musculoskeletal fitness	Children in the highest 20mSRT performance category scored highest on the SBJ $F = 20.99, p < 0.0001$ No differences in 20mSRT performance categories were observed for HG strength $F = 1.43, p = 0.24$ Children in the highest 20mSRT performance category scored highest on the 30 sec sit-ups $F = 20.48, p < 0.0001$ Children in the highest 20mSRT performance category scored highest on the Bent Arm Hang $F = 32.51, p < 0.0001$ Children in the highest 20mSRT performance category scored highest on	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						the 10x5m test $F = 8.32, p < 0.001$	
298 Colquitt (2012); USA	Cross-sectional	n = 122 range = 11-15 yr age = 12.8 yr	VO2max Fitnessgram	Push up test (#) was used to measure muscular endurance Trunk lift test (inches) was used to measure lower back strength Curl-up (#) test was used to measure muscular endurance Sit and Reach test (inches) was used to measure flexibility	Musculoskeletal fitness	20mSRT performance was positively related to the push-up test $r = 0.67, p < 0.01$ 20mSRT performance was positively related to the trunk lift test $r = 0.27, p < 0.01$ 20mSRT performance was positively related to the curl-up test $r = 0.70, p < 0.01$ 20mSRT performance was not associated with the sit and reach test $r = 0.01, p = ns$	n/a
1335 Coe (2013); USA	Cross-sectional	n = 1701 range = Gr. 3, 6, 9	n/a	Sit and Reach test (inches) was used to measure flexibility Curl-up test (#) was used to measure muscular endurance Push up test (#) was used to measure	Musculoskeletal fitness	20mSRT performance was not related to the sit and reach test Gr. 6 $r = -0.01, p = ns$ Gr. 9 $r = 0.03, p = ns$ 20mSRT performance was positively related to the curl-up test Gr. 6 $r = 0.13, p < 0.01$	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				muscular endurance		Gr. 9 $r = 0.11$, $p < 0.01$ 20mSRT performance was positively related to the push up test Gr. 6 $r = 0.11$, $p < 0.01$ Gr. 9 $r = 0.08$, $p < 0.05$	
1407 Coe (2012); USA	Cross-sectional	n = 312 range = Gr. 6-8 age = 12.1 yr	Laps	Curl-up test (#) was used to measure muscular endurance Push up test (#) was used to measure muscular endurance Sit and Reach test (inches) was used to measure flexibility	Musculoskeletal fitness	20mSRT performance was positively related to the curl-up test $r = 0.41$, $p < 0.01$ 20mSRT performance was positively related to the push-up test $r = 0.53$, $p < 0.01$ 20mSRT performance was not associated with the sit and reach test $r = 0.08$, $p = ns$	n/a
1433 Llorente-Cantarero (2012); Spain	Cross-sectional	n = 141 range = 7-12 yr	Stages	No. of abdominals performed in 30 sec was used to measure muscular endurance Standing broad jump (metres) test was used to measure lower body strength	Musculoskeletal fitness	Children in the high 20mSRT performance group completed more abdominals $p = 0.005$ Children in the high 20mSRT performance group performed better on the standing broad jump	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
				Handgrip dynamometry (kg) was used to measure upper body strength		p<0.0001 No differences were observed in 20mSRT performance in handgrip dynamometry p = ns	
1653 Houwen (2010); Netherlands	Cross-sectional	n = 60 range = 6-12 yr age = 8.9 yr	Stages	Sit and Reach test (cm) was used to measure flexibility Standing broad jump test (cm) was used to measure lower body strength Handgrip dynamometry (kg) was used to measure upper body strength No. of abdominals performed in 30 sec was used to measure muscular endurance 10x5m (sec) shuttle run was used to measure agility	Musculoskeletal fitness	20mSRT performance was not related to the sit and reach test r = 0.20, 95 % CI -0.08 to 0.45, p = ns 20mSRT performance was not related to the standing broad jump test r = 0.32, 95 % CI 0.05-0.55, p = ns 20mSRT performance was not related to the handgrip strength test r = 0.16, 95 % CI -0.12 to 0.42, p = ns 20mSRT performance was positively related to performance on the sit-up test r = 0.36, 95 % CI 0.09-0.58, p<0.05	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						20mSRT performance was not related to 5x10m test $r = -0.10$, 95 % CI -0.37 to 0.18, $p = ns$	
1885 Garcia-Artero (2007); Spain	Cross-sectional	$n = 460$ range = 13-18.5 yr Boy age = 15.2 yr Girl age = 15.3 yr	VO2max Leger	Muscular strength score was generated from the standing broad jump, handgrip strength and bent arm hang tests	Musculoskeletal fitness	20mSRT performance was positively related to muscular strength score Boys $r = 0.26$, $p < 0.001$ Girls $r = 0.29$, $p < 0.001$	n/a
2155 Reeves (1999); USA	Cross-sectional	$n = 51$ range = 5-6 yr age = 5.4 yr	Laps	Strength was measured following Bruininks-Oseretsky Test of Motor Proficiency	Musculoskeletal fitness	20mSRT performance was positively related to strength $r = 0.41$, $p = 0.002$	n/a
2458 Coledam (2015); Brazil	Cross-sectional	$n = 781$ range = 10-18 yr	Criterion Reference	Strength was measured using the elbow-flexion test *no units provided*	Musculoskeletal fitness	Meeting the HFZ of 20mSRT performance was associated with meeting criterion reference for strength PR 1.81, 95 % CI 1.47-2.24, $p < 0.001$	Sex Age SES

20mSRT, 20 meter shuttle run test; HFZ, healthy fitness zone; HGS, hand-grip strength; NIZ, needs improvement zone; SAR, sit and reach; SBJ, standing broad jump

Supplement Table 10. 20mSRT performance and motor skill development in children and youth

Note: β indicates standardized Beta

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
534 Marshall (1996); Canada	Cross- sectional	n = 200 range = 5-6 & 9-10 yr	Laps	Motor skill was assessed by the TGMD-2	Motor Skill	20mSRT performance was positively associated with motor skill $r = 0.48, p < 0.05$	n/a
628 Okely (2001); Australia	Cross- sectional	n = 2026 range = Gr. 8- 10	Laps	Motor skill was assessed using an observational behavioural instrument	Motor Skill	20mSRT performance was positively related with all motor skills r range = 0.10-0.34, all $p < 0.01$ 20mSRT performance was positively associated with sprint run Gr. 8 Boys $\beta = 0.19$, $p < 0.01$ Gr. 10 Boys $\beta = 0.18$, $p < 0.01$ Gr. 8 Girls $\beta = 0.17$, $p < 0.01$ Gr. 10 Girls $\beta = 0.29$, $p < 0.01$ 20mSRT performance was positively associated with vertical jump in Gr. 8 girls only	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Gr. 8 Boys $\beta = 0.03$, p = ns Gr. 10 Boys $\beta = 0.05$, p = ns Gr. 8 Girls $\beta = 0.15$, p<0.05 Gr. 10 Girls $\beta = 0.09$, p = ns 20mSRT performance was positively associated with catch in girls only Gr. 8 Boys $\beta = 0.07$, p = ns Gr. 10 Boys $\beta = -0.03$, p = ns Gr. 8 Girls $\beta = 0.11$, p<0.05 Gr. 10 Girls $\beta = 0.14$, p<0.05 20mSRT performance was positively associated with throw in all children except Gr. 8 boys Gr. 8 Boys $\beta = 0.05$, p = ns Gr. 10 Boys $\beta = -0.09$, p<0.05 Gr. 8 Girls $\beta = 0.11$, p<0.05	

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
						Gr. 10 Girls $\beta = 0.14$, $p < 0.05$ 20mSRT performance was positively associated with kick in all children except Gr. 10 girls Gr. 8 Boys $\beta = 0.19$, $p < 0.01$ Gr. 10 Boys $\beta = 0.20$, $p < 0.01$ Gr. 8 Girls $\beta = 0.13$, $p < 0.05$ Gr. 10 Girls $\beta = 0.06$, $p = ns$ 20mSRT performance was positively associated with strike in Gr. 10 boys only Gr. 8 Boys $\beta = 0.02$, $p = ns$ Gr. 10 Boys $\beta = 0.16$, $p < 0.01$ Gr. 8 Girls $\beta = 0.18$, $p = ns$ Gr. 10 Girls $\beta = 0.06$, $p = ns$	
1140 Burns (2015); USA	Cross- sectional	n = 384 range = Gr. K to 6	Laps	Motor skill was assessed by the TGMD-2	Motor skill	Children scoring 1 point higher on the TGMD-2 had increased odds of	n/a

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
		age = 8.4 yr				meeting the HFZ for 20mSRT performance OR 70.3, 95 % CI 2.17- 269.3, $p < 0.05$	
1653 Houwen (2010); Netherlands	Cross- sectional	n = 60 range = 6.12 yr age = 8.9 yr	Stages	Motor skill was assessed by the TGMD-2 Plate tapping test was used to assess coordination	Motor skill	20mSRT performance was positively related to the locomotor TGMD-2 score $r = 0.33$, 95 % CI 0.06- 0.56, $p < 0.05$ 20mSRT performance was positively related to the object control TGMD-2 score $r = 0.31$, 95 % CI 0.04- 0.54, $p < 0.05$ 20mSRT performance was not related to plate tapping test score $r = -0.09$, 95 % CI -0.36 to 0.19, $p = ns$	Age Sex

Note* 20mSRT, 20 m shuttle run test; TGMD, Test of Gross-Motor Skill Development

Supplement Table 11. 20mSRT performance and risk of injury/harm in children and youth

Reference # Author (year); country	Study Design	Sample	Metric and/or equation used from 20mSRT	Outcome	Outcome Category	Main Findings	Covariates controlled for
1094 Martin- Diener (2013); Switzerland	Cross- sectional	n = 249 range = 7-9yr age = 7.9 yr	Laps	Any type of injury that occurred in the past 12 month was reported by the parents/guardians in a retrospective questionnaire	Risk of injury	Those in the highest 20mSRT performance tertile had lower odds of sustaining an injury than those in the lowest 20mSRT performance tertile High 20mSRT OR 0.29, 95 % CI 0.16-0.63 Middle 20mSRT OR 0.37, 95 % CI 0.16-0.85	Sex Age Parental education Migration background Overweight MVPA Motor quotient Training in sports club
1592 Clark (2011); Ireland	Cross- sectional	n = 1590 range = 12-16yr age = 13.9 yr	Laps	Injuries were categorized as previous fractures reported retrospectively by the parent/guardian	Risk of injury	Those in the highest 20mSRT performance tertile had an increased odds of fractures compared with those in the lowest 20mSRT performance tertile High 20mSRT OR 1.50, 95 % CI 1.08-1.60 Middle 20mSRT OR 1.19, 95 % CI 0.89-1.60 p for trend = 0.012	Age Sex Pubertal status

Note* 20mSRT, 20 meter shuttle run test

Appendix C: Ethics

Description: Ethics approval for RBC-CAPL data collection and analysis described in Chapter 5.

Page removed to protect personal information

Appendix D: In the News

Description: A summary of news articles for the publication found in Chapter 6.

Research on the 50-country comparison of young people's aerobic fitness (September 2016). I provided eight original media interviews, and one live interview with CTV News (<https://youtu.be/FRzZMW7kWM0>). This research was communicated by at least 100 broadcast, print and digital media outlets in over a dozen countries: ABC News (<http://www.abc.net.au/news/2016-09-21/fittest-kids-from-tanzania/7865978>), BBC World Service, CBC News (<https://au.news.yahoo.com/sa/a/32689167/australian-kids-fail-the-fitness-test/#page1>), Channel 7 News (<https://au.news.yahoo.com/sa/a/32689167/australian-kids-fail-the-fitness-test/#page1>), Fox News (<http://www.foxnews.com/health/2016/09/30/child-fitness-gap-linked-to-income-inequality.print.html>), Globe and Mail (<http://www.theglobeandmail.com/life/health-and-fitness/fitness/canadianchildren-finish-19th-in-study-of-the-beep-fitness-test/article31985138/>), Reuters Health (<http://www.reuters.com/article/us-health-children-fitness-international-idUSKCN11Z2ZS?feedType=RSS&feedName=healthNews>), Science Daily (<https://www.sciencedaily.com/releases/2016/09/160921104324.htm>), South China Morning Post (<http://www.scmp.com/print/lifestyle/health-beauty/article/2022001/hong-kong-kids-would-be-left-gasping-global-childrens-race>), United Press International (http://www.upi.com/Health_News/2016/09/22/US-ranks-near-bottom-among-countries-for-youth-fitness-study-says/3581474552993/), Valley News Live (<http://www.valleynewslive.com/content/news/UND-Study-American-kids-rank-at-bottom-of-worldwide-fitness-ranking-394848781.html>) , etc. (Altmetric Score = 434[Top 5%]).

Appendix E: Accepted ESSR Proposal Form

Description: Accepted ESSR ‘perspective for progress’ proposal form for the manuscript found in Chapter 7.

EXERCISE AND SPORT SCIENCES REVIEWS

Proposal Submission Form

Handling editor: Ryan Rhodes

Succinct title: Making a case for cardiorespiratory fitness surveillance among children and youth

Corresponding author and institution: Mark S. Tremblay; Children's Hospital of Eastern Ontario Research Institute

Synopsis

Novel hypothesis: We hypothesize that the development of cardiorespiratory fitness guidelines would help support regional, national, and international surveillance efforts.

Supporting author publications:

References supporting our abstract (proposed authors in bold):

Ortega FB, Ruiz JR, Castillo MJ, Sjöström M. Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity*. 2008;32:1–11.

Ruiz JR, Castro-Piñero, Artero EG, Ortega FB, Sjostrom M, Suni J, Castillo MJ. Predictive validity of health-related fitness in youth: a systematic review. *Br J Sports Med* 2009;43:909-23.

Lang JJ, Belanger K, Poitras V, Janssen I, Tomkinson GR, Tremblay MS. Systematic review of the relationship between 20 m shuttle run test performance and health markers in children and youth. *J Sci Med Sport*. 2017a; Submitted

Tomkinson GR, Lang JJ, Tremblay MS, Dale M, Leblanc A, Belanger K, Ortega FB, Leger L. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. *Br J Sports Med*. Published online first: 23 April 2016.

Lang JJ, Tremblay MS, Léger L, Olds TS, Tomkinson GR. International variability in 20 m shuttle run performance in children and youth: who are the fittest from a 50-country comparison? A systematic literature review with pooling of aggregate results. *Br J Sports Med*. Published online first: 20 Sept 2016

Ruiz JR, Cavero-Redondo I, Ortega FB, Welk GJ, Andersen LB, Martinez-Vizcaino V. 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*. Published online first: 26 September 2016.

Lang JJ, Tremblay MS, Ortega FB, Ruiz JR, Tomkinson GR. Review of criterion-referenced standards for cardiorespiratory fitness: What proportion of 1 142 026 international children and youth are apparently healthy? *Br J Sports Med*. 2017; In press.

ESSR Editorial Office
American College of Sports Medicine
401 West Michigan Street
Indianapolis, IN 46202-3233 USA

Tel: 317-634-8932
Email: essr@acsm.org
Online: www.acsm-essr.org

Revised 6/23/2017

EXERCISE AND SPORT SCIENCES REVIEWS

One-page abstract: The purpose of this paper is to propose the development of cardiorespiratory fitness (CRF) guidelines for children and youth. We hypothesize that the development of CRF guidelines would help support regional, national, and international surveillance initiatives. To achieve our purpose we review recent research conducted by the author group that investigates the importance of CRF for current and future health among children and youth. Ortega *et al* (2008) published a narrative review that highlighted the importance of CRF for health among children and youth. Ruiz *et al* (2009) assessed the predictive validity of CRF in kids and concluded that there was strong evidence supporting CRF as a predictor of future health. Recently, Lang *et al* (2017a) conducted a systematic review of 142 papers and identified that the 20 m shuttle run test (20mSRT), a field-based measure of CRF, was a strong indicator of physiological, physical, psychosocial and cognitive health among children and youth. As such, we discuss the 20mSRT as an effective field-based measurement of CRF for population-based research and surveillance by reviewing new international 20mSRT studies. A recent paper by the authors calculated international normative-referenced data using 20mSRT scores from over 1.1 million children and youth representing 50 countries (Tomkinson *et al.*, 2016). A follow-up paper using the same dataset compared 20mSRT performance across 50 countries and identified similar patterns as those reported for physical activity (Lang *et al.*, 2016). The large sample size and country representation of these studies support the international utility of the 20mSRT. As a result, we propose that the 20mSRT would be an effective surveillance indicator to better understand the health of children and youth across cities, regions, and countries. We further highlight potential novel CRF surveillance strategies in the school and clinical setting. To support CRF surveillance we describe the importance of guidelines as a knowledge translation tool, and we compare our guidelines proposal to that of the physical activity guidelines. We conclude by highlighting recent research developing international CRF criterion-referenced standards, which could be used as an initial benchmark to identify the minimum amount of CRF needed for health among children and youth aged 8–18 years (Ruiz *et al.*, 2016). A recent review supported these standards as the best option for international research (Lang *et al.*, 2017b). These standards could be used as the basis to develop future CRF guidelines.

Coauthors names and institutions: Lang JJ, Children's Hospital of Eastern Ontario; Tomkinson GR, University of North Dakota; Janssen I, Queen's University; Ruiz JR, University of Granada; Ortega FB, University of Granada

ESSR Editorial Office
American College of Sports Medicine
401 West Michigan Street
Indianapolis, IN 46202-3233 USA

Tel: 317-634-8932
Email: essr@acsm.org
Online: www.acsm-essr.org

Revised 6/23/2017

EXERCISE AND SPORT SCIENCES REVIEWS

Deadline preference:

(Proposal will be scheduled for the next available issue if selected issue is unavailable.)

- ☐ June 1 (January issue)
- ☐ September 1 (April issue)
- ☐ December 1 (July issue)
- ☒ March 1 (October issue)

Topic area (choose only one):

- ☐ Athlete Care and Clinical Medicine
- ☐ Biomechanics and Neural Control of Movement
- ☐ Cardiovascular, Renal, and Respiratory
- ☐ Clinical Exercise Physiology
- ☐ Environmental and Occupational Physiology
- ☐ Epidemiology and Biostatistics
- ☒ Fitness Assessment, Exercise Training, and Performance
- ☐ Immunology, Genetics, and Endocrinology
- ☐ Metabolism and Nutrition
- ☐ Physical Activity/Health Promotion Interventions
- ☐ Skeletal Muscle, Bone, and Connective Tissue

ESSR Editorial Office
American College of Sports Medicine
401 West Michigan Street
Indianapolis, IN 46202-3233 USA

Tel: 317-634-8932
Email: essr@acsm.org
Online: www.acsm-essr.org

Revised 6/23/2017

Appendix F: Summary of Included Studies (Supplemental Manuscript 1)

Description: A summary of the 177 studies originally included in Supplemental Manuscript 1, and subsequently used in Chapters 4, 6, and Supplemental Manuscript 2.

Supplement 2. Summary of the included studies that have assessed 20 m shuttle run performance of 9- to 17-year-olds.

Reference	Year	Country	Year of testing	Sex	Age range (years)	Sample range	Total sample size	Sampling strategy	Sampling base
Secchi <i>et al</i> [39]	2014	Argentina	2012	M, F	11–17	53–137	1259	NP	S
Hardy <i>et al</i> [40]	2010	Australia	2010	M, F	9–15	114–720	5021	P	S
McIntyre[41]	2009	Australia	1999–00, 2009	M, F	10	41–48	178	NP	O
Booth <i>et al</i> [42]	2006	Australia	2004	M, F	9–15	58–375	3143	P	S
Lloyd & Antonas[43]	2000	Australia	1998?	M, F	11, 12	69–71	280	NP	S
Mulkearns <i>et al</i> [44]	2000	Australia	1996–98	F	9, 10	66–79	145	NP	O
Cooley & McNaughton[45]	1999	Australia	1998	M, F	11–16	339–636	6161	P	S
Booth <i>et al</i> [46]	1997	Australia	1997	M, F	9–15	122–491	3934	P	S
Okely <i>et al</i> [47]	1997	Australia	1996	M	14 ^c	51	51	NP	O
ACHPER[48]	1996	Australia	1994	M, F	9–17	46–104	1282	P	S
Vandongen <i>et al</i> [49]	1995	Australia	1990	M, F	10–12 ^b	162	972	NP	O
ASC[50]	1994	Australia	1993	M, F	12–17	301–542	4458	NP	N
Jenner <i>et al</i> [51]	1992	Australia	1988	M, F	11, 12 ^c	264–283	1094	NP	S
Ortega <i>et al</i> [52]	2008 ^a	Austria	2006–08	F	13	70	70	P	O
Vandendriessche <i>et al</i> [53]	2012	Belgium	2010?	M, F	9–11 ^b	132–235	1018	P	S
Seghers & Rutten[54]	2010	Belgium	2007	M, F	11, 12	69–90	318	P	O
Ortega <i>et al</i> [52]	2008 ^a	Belgium	2006–08	M, F	13–15	40–44	170	P	O
Verstraete <i>et al</i> [55]	2007	Belgium	2002, 2004	M, F	9–11 ^c	77–221	1000	P	S
Cardon <i>et al</i> [56]	2004	Belgium	2002?	M, F	9, 10 ^c	111–186	591	P	S
Heyters & Marique[57]	2004	Belgium	1994, 2004	M, F	10–17	90–306	6418	P	S
Telama <i>et al</i> [58]	2002 ^a	Belgium	1994–95	M, F	12, 15	423–700	2225	NP	S

Baquet <i>et al</i> [59]	2000	Belgium	1997	M, F	12–16	51–81	455	NP	S
Lefèvre <i>et al</i> [60]	1998	Belgium	1993, 1997	M, F	12–17	166–288	5535	P	S
Beunen <i>et al</i> [61]	1991	Belgium	1990	M, F	9–17	211–423	5281	P	S
Poortmans <i>et al</i> [62]	1986	Belgium	1984–85	M, F	12–17	41–82	347	P	S
Cazorla <i>et al</i> [63]	1988	Benin	1987	M, F	9–17	139–425	4141	P	N
Ribeiro <i>et al</i> [64]	2013	Brazil	2006	F	9–17	54–236	1290	P	O
Hobold[65]	2003	Brazil	2001?	M, F	9–17	91–153	1951	P	O
Pieta[66]	2000	Brazil	1998?	M, F	9–14	83–142	1287	NP	O
Veldhuizen <i>et al</i> [67]	2014	Canada	2005–09	M, F	9–15 ^c	40–1012	15217	P	O
Voss <i>et al</i> [68]	2014	Canada	2005–07, 2011	M, F	10–15 ^c	49–189	951	NP	S
Leone <i>et al</i> [69]	2011	Canada	2009–11	M, F	9–12	128–277	1828	P	S
Reed <i>et al</i> [70]	2006	Canada	2004	M, F	10, 11	51–59	219	P	O
Massicotte[71]	1990	Canada	1989–90	M, F	9–17	75–402	5210	P	N
Léger <i>et al</i> [30]	1984	Canada	1981	M, F	9–17	155–404	5584	NP	S
Garber <i>et al</i> [72]	2014	Chile	2011	M, F	13, 14	3416– 4886	15863	P	N
Kain <i>et al</i> [73]	2004	Chile	2002	M, F	11–13	43–68	327	NP	S
Wong <i>et al</i> [74]	2001	China (HK)	1999?	M, F	9–11	77–318	1214	P	S
Wang <i>et al</i> [75]	2011	China (ML)	2010	M, F	9–10 ^c	72–243	548	P	O
Cubides <i>et al</i> [76]	2015	Colombia	2014	M, F	13, 14	41–48	350	P	O
Ortega[77]	2013	Colombia	2008	M, F	9–17	63–853	36117	P	O
Tovar Mojica <i>et al</i> [78]	2008	Colombia	2004	M	9–16	44–89	517	NP	O
Cazorla <i>et al</i> [79]	1985	Côte d’Ivoire	1985	M, F	9–17	92–314	3273	P	N
Shahin[80]	2011	Cyprus	2005	M, F	10	343–372	715	P	O
Tinazci & Emiroğlu[81]	2010	Cyprus	2006	M, F	9–11 ^c	574–1537	6187	P	S

Telama <i>et al</i> [58]	2002 ^a	Czech Republic	1994–95	M, F	12, 15	95–120	439	NP	S
Gajda[82]	1994	Czech Republic	1992–93	M, F	9–14	83–188	1514	NP	S
Nielsen & Andersen[83]	2003	Denmark/GL	1996–97	M, F	15–17	183–3198	9342	P	N
Cazorla <i>et al</i> [84]	1986	Djibouti	1986	M, F	12–17	43–182	1524	P	N
Jürimäe <i>et al</i> [85]	2007 ^a	Estonia	2002	M, F	11–17	108–234	2559	NP	S
Jürimäe & Saar[86]	2003	Estonia	2001?	M, F	10–16 ^b	44–62	419	P	O
Telama <i>et al</i> [58]	2002 ^a	Estonia	1994–95	M, F	12, 15	312–334	1290	NP	S
Jürimäe & Volbekiene[87]	1998 ^a	Estonia	1992	M, F	11–17	133–294	3188	NP	S
Raudsepp & Jürimäe[88]	1997	Estonia	1994–95	F	9, 10	56 – 57	113	P	O
Raudsepp & Jürimäe[89]	1996	Estonia	1994–95	M	9, 10	50–55	105	P	O
Kull & Jürimäe[90]	1994	Estonia	1992–93	M, F	16, 17 ^c	43–131	339	NP	O
Palomäki <i>et al</i> [91]	2015	Finland	2003, 2010	M, F	15, 16	320–571	3560	P	N
Telama <i>et al</i> [58]	2002 ^a	Finland	1995	M, F	12, 15	254–307	1109	NP	S
Viljanen <i>et al</i> [92]	2000	Finland	1993	M, F	10, 14, 16	63–249	1010	NP	S
Ortega <i>et al</i> [52]	2008 ^a	France	2006–08	F	13–14	44–46	90	P	O
Baquet <i>et al</i> [93]	2006	France	1997, 2001	M, F	11–15 ^c	44–54	197	NP	O
Baquet <i>et al</i> [94]	2001	France	1997	M, F	11–14	45–88	468	P	O
Baquet <i>et al</i> [95]	1999	France	1992	F	10	49	49	NP	O
Carzola <i>et al</i> [96]	1997	France	1996–97	M, F	9–11	47–110	498	P	S
Carzola <i>et al</i> [97]	1999	France	1999	M, F	11–17	85–234	2217	P	S
Cazorla[98]	1987	France	1987	M, F	9–12	119–693	3642	P	N
Brunet & Van Praagh[99]	1985	France	1984–85	M, F	9, 10	40–43	83	NP	O
Ortega <i>et al</i> [52]	2008 ^a	Germany	2006–08	M	14–16	53–61	169	P	O

Telama <i>et al</i> [58]	2002 ^a	Germany	1994–95	M, F	12, 15	216–264	977	NP	S
Tambalis <i>et al</i> [100]	2015	Greece	2014	M, F	9–17	490–20149	196370	P	N
Tambalis <i>et al</i> [101]	2011	Greece	1997, 2007	M, F	9	2253–15219	66573	P	N
Ortega <i>et al</i> [52]	2008 ^a	Greece	2006–08	M, F	14	41–60	212	P	O
Tokmakidis <i>et al</i> [102]	2006	Greece	2004	M, F	9–11	43–68	300	NP	O
Manios <i>et al</i> [103]	2004	Greece	2001–02	M, F	11 ^c	52–61	113	P	O
Georgiadis[104]	1993	Greece	1990–91	M, F	9–17	255–297	4814	P	N
Welk <i>et al</i> [105]	2015	Hungary	2013	M, F	11–17	110–204	2187	P	N
Ortega <i>et al</i> [52]	2008 ^a	Hungary	2006–08	M, F	13–15	52–59	279	P	O
Telama <i>et al</i> [58]	2002 ^a	Hungary	1994–95	M, F	12, 15	100–114	439	NP	S
Barabás[106]	1992	Hungary	1990?	F	14–17	49–130	389	NP	O
Gunnarsson & Sigriksson[107]	1999	Iceland	1998	M, F	9–15	359–528	6130	NP	N
Ortega <i>et al</i> [52]	2008 ^a	Italy	2006–08	F	16	43	43	P	O
Grassi <i>et al</i> [108]	2006	Italy	2001	F	14–17	42–53	95	NP	O
Cilia <i>et al</i> [109]	1998	Italy	1997	M, F	15–17	44–68	273	NP	O
Cilia <i>et al</i> [110]	1997	Italy	1997	M, F	12–17	109–404	3203	NP	S
Cilia <i>et al</i> [111]	1996	Italy	1995	M, F	12–14	85–120	659	NP	S
Cilia & Bellucci[112]	1993	Italy	1992	M, F	12–14	189–259	1363	NP	S
Council of Europe[113]	1986	Italy	1985–86	F	14–17	44–79	264	NP	S
MoECSSaT[114]	2015	Japan	2014	M, F	9–17	920–1125	18147	P	N
MoECSSaT[114]	2014	Japan	2013	M, F	9–17	742–1089	15947	P	N
MoECSSaT[114]	2013	Japan	2012	M, F	9–17	916–1125	18062	P	N
MoECSSaT[114]	2012	Japan	2011	M, F	9–17	823–1113	17024	P	N

MoECSSaT[114]	2011	Japan	2010	M, F	9–17	822–1117	17323	P	N
MoECSSaT[114]	2010	Japan	2009	M, F	9–17	807–1124	16850	P	N
MoECSSaT[114]	2009	Japan	2008	M, F	9–17	788–1119	16989	P	N
MoECSSaT[114]	2008	Japan	2007	M, F	9–17	765–1127	16285	P	N
MoECSSaT[114]	2007	Japan	2006	M, F	9–17	782–1128	16771	P	N
MoECSSaT[114]	2006	Japan	2005	M, F	9–17	765–1121	16510	P	N
MoECSSaT[114]	2005	Japan	2004	M, F	9–17	706–1126	15610	P	N
MoECSSaT[114]	2004	Japan	2003	M, F	9–17	676–1126	15619	P	N
MoECSSaT[114]	2003	Japan	2002	M, F	9–17	636–1126	14796	P	N
MoECSSaT[114]	2002	Japan	2001	M, F	9–17	657–1123	15423	P	N
MoECSSaT[114]	2001	Japan	2000	M, F	9–17	606–1060	14514	P	N
MoECSSaT[114]	2000	Japan	1999	M, F	9–17	701–1059	15354	P	N
MoECSSaT[114]	1999	Japan	1998	M, F	9–17	885–1120	18508	P	N
Sauka <i>et al</i> [115]	2011	Latvia	2004–09	M, F	9–17	62–312	3400	P	N
Jūrimāe <i>et al</i> [85]	2007 ^a	Lithuania	2002	M, F	11–17	214–426	4878	NP	S
Jūrimāe & Volbekiene[87]	1998 ^a	Lithuania	1992	M, F	11–17	66–150	1601	NP	S
Cazorla <i>et al</i> [116]	1986	Mauritius	1986	M, F	9–17	137–160	2695	P	N
Esteve-Lanao <i>personnel comm</i>	2015	Mexico	2010–11	M, F	9–16	42–685	7449	NP	S
Galaviz <i>et al</i> [117]	2012	Mexico	2009	M, F	11	43–49	92	NP	O
Dadouchi <i>et al</i> [118]	2003	Morocco	2002	M, F	12–17	77–302	2522	NP	O
Brouwer <i>et al</i> [119]	2013	Netherlands	2005–07	M, F	15, 16 ^c	126–144	540	P	S
Slinger <i>et al</i> [120]	2009	Netherlands	2002–05	M, F	12–15	53–308	1043	P	O
van Mechelen <i>et al</i> [121]	1991	Netherlands	1987	M, F	12–16	101–285	1874	P	N
Haugen <i>et al</i> [122]	2013	Norway	2004–06	M, F	13–15 ^c	352–406	2305	NP	O

Cossio-Bolaños & Arruda[123]	2009	Peru	2001	M, F	9–12	58–81	560	P	O
Gonzalez-Suarez & Crimmer-Sommers[124]	2011	Philippines	2007–08	M, F	11 ^b	152–188	340	P	O
Bronikowski & Bronikowska[125]	2009	Poland	2007?	M	13, 14 ^c	84–115	314	NP	O
Pilicz <i>et al</i> [126]	2003	Poland	1999	M, F	9–17	2055–3470	53813	P	N
Maciaszek & Osinski[127]	2001	Poland	1999?	M, F	10–14	245–371	3123	P	O
Mleczko <i>et al</i> [128]	2000	Poland	1991–92	M, F	15–17	198–303	1431	P	O
Santos <i>et al</i> [129]	2014	Portugal	2008	M, F	10–17	486–1695	19861	P	N
Coelho-Silva <i>et al</i> [130]	2013	Portugal	2008–09	M, F	11–15	61–143	1755	P	S
Rodrigues <i>et al</i> [131]	2013	Portugal	1997–00, 2006	M, F	9–15 ^c	52–168	835	NP	O
Marta <i>et al</i> [132]	2012	Portugal	2011	M, F	10, 11 ^c	47–111	312	NP	O
Silva <i>et al</i> [133]	2012	Portugal	2008–09	M, F	10–17	179–493	5712	P	O
Mota <i>et al</i> [134]	2002	Portugal	1998–00	M, F	9 ^c	59–61	120	P	O
Oliveira <i>personal comm</i>	2002	Portugal	2001–02	M, F	11–17	80–250	2169	NP	O
Kim <i>et al</i> [135]	2011	Republic of Korea (South Korea)	2009?	F	15,16	41	82	NP	O
Cazorla <i>et al</i> [136]	1988	Senegal	1988	M, F	9–17	86–163	2243	P	N
Bovet <i>et al</i> [137]	2007	Seychelles	2004	M, F	11–16	174–658	5371	P	N
Cazorla <i>et al</i> [138]	1990	Seychelles	1990	M, F	9–17	142–258	3803	P	N
Krska <i>et al</i> [139]	<i>In press</i>	Slovakia	2014–15	F	9–11	76–105	181	NP	O
Krska <i>et al</i> [140]	<i>In press</i>	Slovakia	2014–15	M	9–11	80–108	188	NP	O
Kyselovicová[141]	2000	Slovakia	1994	F	15	54	54	P	O
Kasa & Majherová[142]	1997	Slovakia	1996	M	11, 13	40–45	85	NP	O

Moravec <i>et al</i> [143]	1996	Slovakia	1993	M, F	9–17 ^c	74–554	3630	NP	N
Belej <i>et al</i> [144]	1995	Slovakia	1993?	M, F	14	344–345	689	NP	S
Pienaar <i>et al</i> [145]	2015	South Africa	2012	M, F	16 ^c	45–81	182	P	O
Du Toit <i>et al</i> [146]	2011	South Africa	2002	F	11 ^c	44	44	NP	O
Pienaar & Vilioen[147]	2010	South Africa	2000–01	M	10–15	82–134	585	P	S
Du Preez[148]	2008	South Africa	2005	M	10, 11	44–45	89	NP	O
Stadler[149]	2007	South Africa	2005	F	11, 12	40–51	91	NP	O
Du Toit <i>et al</i> [150]	2005	South Africa	2003?	F	15, 16	46–81	127	NP	O
Van Gent <i>et al</i> [151]	2002	South Africa	2000–01	F	10–15	92–117	622	P	S
du Randt[152]	1996	South Africa	1995–96	M, F	11–16	45–210	1230	P	N
Gulías-González <i>et al</i> [153]	2014	Spain	2010	M, F	9–16	52–142	1494	P	S
Torrijos-Niño <i>et al</i> [154]	2014	Spain	2010	M, F	9–10 ^c	93–347	876	P	S
Chillón <i>et al</i> [155]	2011	Spain	1999–00	M, F	9–16	75–206	2224	P	O
Castro-Pinero <i>et al</i> [156]	2009	Spain	2006–07	M, F	9–17	45–171	1898	P	S
Ortega <i>et al</i> [52]	2008 ^a	Spain	2006–07	M, F	15	44	44	P	O
Ortega <i>et al</i> [157]	2005	Spain	2001–02	M, F	12–17	78–290	2026	P	S
García Baena[158]	1999	Spain	1999	M, F	13–16	46–70	364	P	S
Prat <i>et al</i> [159]	1998	Spain	1984–85	M, F	10–17	199–267	3932	P	S
Tercedor & Delgado-Fernando[160]	1998	Spain	1995	M, F	10	124–160	284	P	O
Sainz[161]	1996	Spain	1990–94	M, F	9–17	54–671	6658	P	S
Ureña[162]	1996	Spain	1995–96	M, F	14–16	99–106	613	NP	O
Brito Ojeda <i>et al</i> [163]	1995	Spain	1993?	M, F	12–17	41–50	388	P	S
Sainz[164]	1992	Spain	1986–89	M, F	10–15	69–180	1300	NP	S
Rivas[165]	1987	Spain	1987	M, F	9–17	88–303	3849	NP	O

Lieveld <i>et al</i> [166]	1993	Suriname	1991?	M	14	70–115	270	NP	O
Ortega <i>et al</i> [52]	2008 ^a	Sweden	2006–08	F	15	45	45	P	O
Schmid <i>et al</i> [167]	2007	Switzerland	2005	M, F	11	127–138	265	P	O
Cauderay <i>et al</i> [168]	2000	Switzerland	1996–97	M, F	9–17	92–214	2959	P	S
Aandstad <i>et al</i> [169]	2006	Tanzania	2003	M, F	9–10	61–76	274	P	O
Çalis <i>et al</i> [170]	1992	Turkey	1991	M	15	45	45	NP	O
Richards[171]	2014	Uganda	2010	M, F	11–14	44–291	1446	P	O
Sandercock <i>et al</i> [172]	2015	UK (England)	2014	M, F	10 ^c	150–157	307	NP	O
Sandercock <i>et al</i> [173]	2012	UK (England)	2006–10	M, F	10–16	60–933	7354	NP	S
Boddy <i>et al</i> [174]	2010	UK (England)	1999–2010	M, F	9–10 ^b	104–1213	27942	P	O
Sandercock <i>et al</i> [175]	2010	UK (England)	1998, 2008	M, F	10 ^c	145–158	618	NP	O
Liverpool City Council[176]	2003	UK (England)	2000–03	M, F	11–12 ^b	610–1123	3466	NP	O
Boreham <i>et al</i> [177]	2001	UK (N Ireland)	1989–90	M, F	12, 15	251–258	1015	P	S
Watkins <i>et al</i> [178]	2001	UK (N Ireland)	1989–90, 1999–01	M, F	12, 15	251–532	3032	P	N
Twisk <i>et al</i> [179]	1999	UK (N Ireland)	1992–93	M, F	15	229–230	459	P	S
Mahoney[180]	1992	UK (N Ireland)	1990–91	M, F	12	50–53	103	NP	O
Mahoney & Boreham[181]	1991	UK (N Ireland)	1989?	M, F	9, 11 ^c	57–87	299	P	S
Riddoch[182]	1990	UK (N Ireland)	1988–89	M, F	11–17	180–237	3012	P	N
Nichols & Riddoch[183]	1986	UK (N Ireland)	1986	M, F	12–15 ^c	44–163	474	P	S
Lewitt <i>et al</i> [184]	2012	UK (Wales)	2007	M, F	13	97–141	238	P	O
Ranson & Taylor[185]	2012	UK (Wales)	2009–10	M, F	9–11 ^c	95–164	821	NP	S
Welk <i>et al</i> [186]	2015	USA	2011–14	M, F	10–17	1591– 12143	114585	NP	S
Carrel <i>et al</i> [187]	2012	USA	2008–10	M, F	9–17	132–2015	18005	NP	S
Liu <i>et al</i> [188]	2012	USA	2010?	M, F	11–13 ^c	103–121	672	NP	O

Welk <i>et al</i> [Error! Reference source not found.]	2011	USA	2009–10	M, F	11–17	62–187	1637	NP	O
Beets <i>et al</i> [190]	2005	USA	2001–02	M, F	10–15	71–2787	34524	P	O
Beets & Pitetti[191]	2004	USA	1999–01	M, F	10–17	41–63	365	NP	S
Welk <i>et al</i> [192]	2004	USA	2002?	M, F	9–16 ^b	43–68	519	NP	O
Lloyd <i>et al</i> [193]	2003	USA	2001?	M, F	10–12	44–67	111	NP	O
Chun <i>et al</i> [194]	2000	USA	1997	M, F	12	116–120	236	NP	O
Wolford[195]	1998	USA	1997	M, F	10, 11 ^c	43–57	195	NP	O
Mahar <i>et al</i> [196]	1997	USA	1995	M, F	10 ^c	98–111	209	NP	O

^aIndicates multinational study. ^b Age reported as an age range. ^c Age reported as a mean and standard deviation. ? indicates year of testing not reported. P = Probability sampling; NP = non-probability sampling; N = national sample; NP = State/Provincial sample; O = other sample (e.g. local, city or school level). Note, some studies reported data for children outside of the target age range (e.g. <9 and >17 years of age), however this table only reports the data for 9- to 17-year-olds.

Appendix G: Summary of Conference Abstracts and Presentations

Description: Citations of published conference abstracts and presentations for the works included in this dissertation.

Citations

1. **Lang JJ**, Dale M, LeBlanc A, et al. Systematic review and analysis of 20 m shuttle run results in children and youth. *Medicine and Science in Sports and Exercise*. 2016; 48(5S Suppl 1):1059–60.
2. Tomkinson GR, **Lang JJ**, Tremblay MS, et al. International normative 20 m shuttle run values from 850 036 children and youth representing 48 countries. *Medicine and Science in Sports and Exercise*. 2016;48(5S Suppl 1):108–9.
3. Tomkinson GR, **Lang JJ**, Tremblay MS, et al. International normative 20 m shuttle run values from 850 036 children and youth representing 48 countries. Poster presented at the 2016 annual ACSM conference, Boston, United States.
4. **Lang JJ**, LeBlanc AG, Dale M, Belanger K, Ortega FB, Léger L, Tremblay MS, Tomkinson GR. Systematic review and analysis of 20 m shuttle run results in children and youth. Poster presented at the 2016 annual ACSM conference, Boston, United States.
5. **Lang JJ**, Tremblay MS, Léger L, Olds T, Tomkinson G. International variability in 20 m shuttle run performance in children and youth: Who are the best from a 50 country comparison? Oral presented at the 2016 International Congress on Physical Activity and Public Health, Bangkok, Thailand.
6. **Lang JJ**, Tremblay MS, Tomkinson G. Comparing cardiorespiratory fitness standards: What proportion of children and youth are healthy? Poster presented at the 2016 International Congress on Physical Activity and Public Health, Bangkok, Thailand.