



UNIVERSIDADE TÉCNICA DE LISBOA
FACULDADE DE MOTRICIDADE HUMANA



Validity of the 20-m MSRT as a Predictor of VO_{2peak} in Lisbon Elementary School Children

Dissertação elaborada com vista à obtenção do Grau de Mestre na especialidade de
Exercício e Saúde

Orientador: Professora Doutora Maria Helena Santa-Clara Pombo Rodrigues

Júri:

Presidente:

Professor Doutor Luís Fernando Cordeiro Bettencourt Sardinha

Vogais:

Professor Doutor João Manuel Pardal Barreiros

Professor Doutor João Paulo Reis Gonçalves Moreira de Brito

Professora Doutora Paula Marta Pereira Bruno

Pedro Xavier Melo Fernandes Castanheira

2009



Apesar de, porventura erradamente, pensar-se que uma tese de mestrado é no seu todo um investimento pessoal, ela é na verdade um projecto colectivo. Desta forma, gostaria de agradecer:

Aos meus pais e à Débora por terem sempre acreditado no engrandecimento pessoal que advêm deste projecto e por terem criado o suporte familiar essencial para o seu bom término.

Ao Prof. Elvis Carnero pelos seus ensinamentos e opiniões, à Prof. Marta Bruno pela incansável e indescritível dedicação no tratamento estatístico dos dados, e à Prof. Helena Santa-Clara pela impressionante demonstração de confiança neste projecto e pelos incontáveis reforços positivos que me impulsionaram para mais além.

Ao ATL e funcionárias da escola E.B 1 J.I Agualva 1 que foram incansáveis e extremamente pragmáticas na criação das condições ideais de teste.

Ao Conselho Executivo da escola E.B 1 J.I Agualva 1 que sempre proporcionou e valorizou o investimento na formação pessoal dos seus Professores.

A todos os alunos da escola E.B 1 J.I Agualva 1 cuja participação e esforços foram a matriz desta tese.

A todos os meus amigos que, cada um à sua maneira, criaram o suporte social que transpareceu durante todo o período deste mestrado.

À Faculdade de Motricidade Humana pela cedência dos equipamentos.

Contents

ABSTRACT	7
THEME'S PRESENTATION	8
INTRODUCTION	8
ISSUE'S DEFINITION	10
STUDY'S EXTENT	10
ASSUMPTIONS	10
LIMITATIONS.....	10
STUDY'S SIGNIFICANCE.....	11
OPERATIONAL DEFINITIONS AND ABBREVIATIONS.....	11
LITERATURE REVISION	14
INTRODUCTION	14
PHYSIOLOGICAL ASPECTS	14
METABOLIC RESPONSE TO EXERCISE IN CHILDREN.....	14
CARDIOVASCULAR RESPONSE TO EXERCISE IN CHILDREN	17
PULMONARY RESPONSE TO EXERCISE IN CHILDREN.....	20
CONDITIONING AND TRAINING	21
HEALTH-RELATED ASPECTS.....	23
VO _{2MAX} – DEFINITION, CRITERION AND EXPRESSION FORMATS	24
DEFINITION	24
CRITERION.....	24
ABSOLUTE AND RELATIVE EXPRESSION OF VO _{2MAX}	26
EXERCISE TESTING TO OBTAIN A VO _{2MAX}	26
LABORATORY VERSUS FIELD EXERCISE TESTING	27
20-M MULTISTAGE SHUTTLE RUN TEST (20-M MSRT)	27
20-M MSRT ADVANTAGES	29
20-M MSRT DISADVANTAGES.....	29
VALIDATION STUDIES OF THE 20-M MSRT AS AN AEROBIC FITNESS TEST IN CHILDREN AND YOUTH	30
VALIDATION STUDIES OF THE 20-M MSRT AS AN AEROBIC FITNESS TEST IN ADULTS.....	33
VALIDATION STUDIES OF THE 20-M MSRT AS A PERFORMANCE OUTPUT PREDICTOR.....	37
RELIABILITY OF THE 20-M MSRT IN CHILDREN AND YOUTH.....	39
THE 20-M MSRT IN THE WORLD ACROSS THE YEARS.....	40
METHODOLOGY.....	43
INTRODUCTION	43

EXPERIMENTAL CONCEPT	43
PARTICIPANTS	43
INSTRUMENTS AND PROTOCOLS	43
ANTHROPOMETRIC MEASURES	43
20-M MSRT TEST PROTOCOL	44
OXYGEN UPTAKE TRANSIENTS MEASURES	44
TIME AND SPACE RESOURCES	46
DATA ASSORTMENT	47
DATA PREPARATION	47
PROCEDURES	47
BEFORE THE TEST	47
DURING THE TEST	48
AFTER THE TEST	48
STATISTICAL ANALYSIS	48
RESULTS	50
INTRODUCTION	50
EXPLORATORY ANALYSIS	50
VALIDITY OF THE 20-M MSRT AS A PREDICTOR OF VO_{2PEAK} IN LISBON ELEMENTARY SCHOOL CHILDREN	52
VALIDITY OF THE FITNESSGRAM'S PACER TEST SOFTWARE REPORTS AS A PREDICTOR OF VO_{2PEAK} IN LISBON ELEMENTARY SCHOOL CHILDREN	52
COMPARING SEVERAL REGRESSION EQUATIONS THAT PREDICT VO_{2PEAK} USING THE 20-M MSRT IN LISBON ELEMENTARY SCHOOL CHILDREN	56
RELIABILITY OF THE 20-M MSRT IN THE PRESENT STUDY	64
DISCUSSION	67
ANTHROPOMETRIC, SPIROMETRIC AND PERFORMANCE VARIABLES	67
VALIDITY OF THE 20-M MSRT AS A PREDICTOR OF VO_{2PEAK} IN LISBON ELEMENTARY SCHOOL CHILDREN	68
VALIDITY OF THE FITNESSGRAM'S PACER TEST SOFTWARE REPORTS AS A PREDICTOR OF VO_{2PEAK} IN LISBON ELEMENTARY SCHOOL CHILDREN	69
COMPARING SEVERAL REGRESSION EQUATIONS THAT PREDICT VO_{2PEAK} USING THE 20-M MSRT IN LISBON ELEMENTARY SCHOOL CHILDREN	70
RELIABILITY OF THE 20-M MSRT IN THE PRESENT STUDY	72
CONCLUSIONS	74
FUTURE RECOMENDATIONS	74
REFERENCES	75

FIGURE 1: TYPICAL (VO_2 AND VCO_2 ; $ML \cdot MIN^{-1}$)/T LINE GRAPHIC REPRESENTATION AFTER AVERAGING AND SMOOTHING THE DATA MEASURED THROUGH $K4B^2$.	45
FIGURE 2: THE $K4B^2$ METABOLIC MEASURING SYSTEM: 1) BATTERY + ANTENNA; 2) $K4B^2$ UNIT; 3) HR PROBE; 4) MASK; 5) FLOWMETER.	45
FIGURE 3: FRONT IMAGE OF A PARTICIPANT WEARING STILL THE $K4B^2$, COSMED	46
FIGURE 4: BACK IMAGE OF A PARTICIPANT WEARING STILL THE $K4B^2$, COSMED.	46
FIGURE 5: FRONT IMAGE OF A PARTICIPANT WEARING THE $K4B^2$, COSMED WHILE RUNNING	46
FIGURE 6: BACK IMAGE OF A PARTICIPANT WEARING THE $K4B^2$, COSMED WHILE RUNNING	46
FIGURE 7: DISPOSITION OF THE DIFFERENT ELEMENTS OF THE TEST.	46
FIGURE 8: GENDER BASED HISTOGRAM OF MEASURED DIRECTLY VO_{2PEAK} , WITH NORMAL CURVE DISPLAYED.	51
FIGURE 9: MEAN MEASURED DIRECTLY VO_{2PEAK} IN BOYS (FULL LINE) AND GIRLS (DASHED LINE) ACROSS AGES.	51
FIGURE 10: MEAN MEASURED DIRECTLY VO_{2PEAK} (FULL LINE) AND MEAN ESTIMATED VO_{2MAX} BY FITNESSGRAM (DASHED LINE) IN: A) BOYS; B) GIRLS; C) TOTAL, ACROSS AGES.	53
FIGURE 11: BOXPLOTS OF THE MEASURED DIRECTLY VO_{2PEAK} (GREY) AND THE ESTIMATED VO_{2MAX} BY FITNESSGRAM (WHITE) IN: A) WHOLE SAMPLE; B) SPLIT BY GENDER.	53
FIGURE 12: SCATTERPLOT OF ESTIMATED VO_{2MAX} BY FITNESSGRAM AGAINST MEASURED DIRECTLY VO_{2PEAK} VALUES IN: A) BOYS; B) GIRLS; C) BOTH.	55
FIGURE 13: BOXPLOTS OF MEASURED DIRECTLY VO_{2PEAK} AND ESTIMATED VO_{2MAX} FROM PREVIOUSLY REPORTED EQUATIONS.	57
FIGURE 14: GENDER BASED BAR CHART OF THE INDIVIDUAL DIFFERENCE (N=90) BETWEEN MEASURED DIRECTLY VO_{2PEAK} AND ESTIMATED VO_{2MAX} FROM: A) FITNESSGRAM; B) LEGER ET AL (1988); C) D) E) BARNETT ET AL (1993); F) FERNHALL ET AL (1998); G) OLIVEIRA (1998) AND H) I) MATSUZAKA ET AL (2004). THE TWO HORIZONTAL REFERENCE LINES IN EACH GRAPHIC REPRESENT THE \pm SD OF MEAN MEASURED DIRECTLY VO_{2PEAK} ($49.6 \pm 9.0 ML \cdot KG^{-1} \cdot MIN^{-1}$).	59
FIGURE 15: SCATTERPLOT OF MEAN DIFFERENCE (MEASURED VO_{2PEAK} - FITNESSGRAM; Y-AXIS) AGAINST THE AVERAGE VALUE (TRIANGLES; MEAN: MEASURED VO_{2PEAK} , FITNESSGRAM; X-AXIS) AND MEAN DIFFERENCE (MEASURED VO_{2PEAK} - FITNESSGRAM; Y-AXIS) AGAINST MEASURED DIRECTLY VO_{2PEAK} (CIRCLES) (X-AXIS).	61
FIGURE 16: BLAND-ALTMAN PLOT OF MEASURED DIRECTLY VO_{2PEAK} AND THE ESTIMATED VO_{2MAX} FROM: A) FITNESSGRAM; B) LEGER ET AL (1988); C) D) E) BARNETT ET AL (1993); F) FERNHALL ET AL (1998); G) OLIVEIRA (1998) AND H) I) MATSUZAKA ET AL (2004).	62
FIGURE 17: BOXPLOTS REPRESENTATIONS OF MEASURED DIRECTLY VO_{2PEAK} VALUES (A) AND TOTAL NUMBER OF LAPS (B) IN TEST-RETEST TRIALS.	64
FIGURE 18: SCATTERPLOTS OF TEST AGAINST RETEST OF: A) MEASURED DIRECTLY VO_{2PEAK} ; B) RER; C) TOTAL NUMBER OF LAPS; D) HR_{MAX} .	65

TABLE 1: UNIQUE METABOLIC RESPONSES OF CHILDREN RELATIVE TO ADULTS. ADAPTED FROM ACSM (2005).	15
TABLE 2: SOME FACTORS KNOWN TO AFFECT HR RESPONSE TO EXERCISE AMONG CHILDREN AND ADOLESCENTS (BAR-OR 1983).....	18
TABLE 3: PEAK HEART RATE IN CHILDREN AND ADOLESCENTS IN PREVIOUS REPORTED STUDIES. ADAPTED FROM WASHINGTON ET AL (1994).....	19
TABLE 4: UNIQUE CARDIOVASCULAR RESPONSES OF CHILDREN RELATIVE TO ADULTS. ADAPTED FROM ACSM (2005).	20
TABLE 5: UNIQUE PULMONARY RESPONSES OF CHILDREN RELATIVE TO ADULTS. ADAPTED FROM ACSM (2005).....	21
TABLE 6: CARDIOVASCULAR CHANGES THAT OCCUR IN CHILDREN, WITH CONDITIONING. ADAPTED FROM ACSM (2005).....	23
TABLE 7: PULMONARY CHANGES THAT OCCUR IN CHILDREN, WITH CONDITIONING. ADAPTED FROM O. BAR-OR (1983).....	23
TABLE 8: MAXIMAL OXYGEN UPTAKE VALUES (VO_{2MAX} ; $ML.KG^{-1}.MIN^{-1}$) FOUND IN PREVIOUS RESEARCHES INVOLVING CVF TESTING IN CHILDREN AND YOUTH.	25
TABLE 9: MAXIMAL OXYGEN UPTAKE VALUES (VO_{2MAX} ; $ML.KG^{-1}.MIN^{-1}$) ACROSS DIFFERENT STANDARD TESTING DEMANDS IN CHILDREN. ADAPTED FROM WASHINGTON ET AL (1994).	27
TABLE 10: RUNNING SPEED ($KM.H^{-1}$) AND TOTAL NUMBER OF LAPS FOR EACH STAGE OF THE 20-M MSRT ACCORDING TO FOUR MAJOR PROTOCOL VARIANTS. BASED ON TOMKINSON ET AL (2003).	29
TABLE 11: VALIDATION STUDIES OF THE 20-M MSRT IN CHILDREN AND YOUTH.....	30
TABLE 12: VALIDATION STUDIES OF THE 20-M MSRT IN ADULTS.	33
TABLE 13: SEVERAL OTHER 20-M MSRT STUDIES IN CHILDREN AND YOUTH.	36
TABLE 14: PREVIOUSLY REPORTED VO_{2MAX} PREDICTIVE EQUATIONS IN CHILDREN AND YOUTH BASED ON 20-M MSRT PERFORMANCE.	37
TABLE 15: 20-M MSRT VALIDATION STUDIES AS A PERFORMANCE OUTPUT PREDICTOR.....	38
TABLE 16: RELIABILITY OF THE 20-M MSRT IN CHILDREN AND YOUTH.	39
TABLE 17: RELIABILITY OF THE 20-M MSRT IN ADULTS.....	40
TABLE 18: NORMATIVE STUDIES OF THE 20-M MSRT AROUND THE WORLD. ADAPTED FROM T. OLDS ET AL (2006).	41
TABLE 19: SAMPLE'S SELECTION.	43
TABLE 20: RUNNING SPEED ($KM.H^{-1}$) AND NUMBER OF LAPS FOR EACH STAGE OF THE PACER TEST ACCORDING TO THE COOPER INSTITUTE FOR AEROBICS RESEARCH (2002).	44
TABLE 21: AGE (YEARS) AND MORPHOLOGIC CHARACTERISTICS (MEAN \pm SD): HEIGHT (M), WEIGHT (KG), BODY MASS INDEX (BMI; KG/M^2) AND %FAT.	50
TABLE 22: SPIROMETRIC AND CARDIOVASCULAR VARIABLES (MEAN \pm SD) ASSESSED DURING THE PACER TEST.....	50
TABLE 23: GENDER/AGE BASED DISTRIBUTION OF MEASURED DIRECTLY VO_{2PEAK} ($ML.KG^{-1}.MIN^{-1}$; MEAN \pm SD).....	51
TABLE 24: PERFORMANCE VARIABLES ASSESSED (MEAN \pm SD) DURING THE PACER TEST: TOTAL LAPS (N°), MAXIMAL SPEED ($KM.H^{-1}$) TEST TIME (MM:SS) AND DISTANCE (M).	52
TABLE 25: PEARSON PRODUCT-MOMENT CORRELATIONS BETWEEN MEASURED DIRECTLY VO_{2PEAK} ($ML.KG^{-1}.MIN^{-1}$) AND TOTAL NUMBER OF LAPS (N°) AND BETWEEN MEASURED DIRECTLY VO_{2PEAK} ($ML.KG^{-1}.MIN^{-1}$) AND MAXIMAL SPEED ($KM.H^{-1}$).	52
TABLE 26: PEARSON PRODUCT-MOMENT CORRELATIONS BETWEEN MEASURED DIRECTLY VO_{2PEAK} ($ML.KG^{-1}.MIN^{-1}$) AND TOTAL LAPS (N°), WITH WEIGHT, HEIGHT, AND BMI; AND SPEARMAN'S RANK CORRELATION BETWEEN MEASURED DIRECTLY VO_{2PEAK} ($ML.KG^{-1}.MIN^{-1}$) AND TOTAL LAPS (N°) WITH AGE.....	52
TABLE 27: GENDER/AGE BASED MEASURED DIRECTLY VO_{2PEAK} AND ESTIMATED VO_{2MAX} BY FITNESSGRAM (MEAN \pm SD; $ML.KG^{-1}.MIN^{-1}$); MEAN DIFFERENCES ($D \pm$ SD; $ML.KG^{-1}.MIN^{-1}$), PEARSON PRODUCT-MOMENT CORRELATION (R) AND CONCORDANCE CORRELATION (R_c) BETWEEN MEASURED DIRECTLY AND ESTIMATED VO_{2PEAK}	54
TABLE 28: GENDER/AGE BASED MEASURED VO_{2PEAK} AND ESTIMATED VO_{2MAX} VALUES (MEAN \pm SD; $ML.KG^{-1}.MIN^{-1}$) DERIVED FROM SEVERAL 20-M MSRT ESTIMATING EQUATIONS FOR CHILDREN AND YOUTH.	56
TABLE 29: MINIMUM, MAXIMUM AND MEAN DIFFERENCES (MEAN \pm SD; $ML.KG^{-1}.MIN^{-1}$) BETWEEN MEASURED DIRECTLY VO_{2PEAK} AND ESTIMATED VO_{2MAX} FROM PREVIOUSLY REPORTED EQUATIONS.	58
TABLE 30: PEARSON PRODUCT-MOMENT CORRELATIONS AND CONCORDANCE CORRELATION BETWEEN MEASURED DIRECTLY VO_{2PEAK} AND ESTIMATED VO_{2MAX} FROM PREVIOUSLY REPORTED EQUATIONS.	60

TABLE 31: BLAND-ALTMAN APPROACH TO COMPARE MEASURED DIRECTLY VO_{2PEAK} ($ML.KG^{-1}.MIN^{-1}$) AND ESTIMATED VO_{2MAX} ($ML.KG^{-1}.MIN^{-1}$) FROM PREVIOUSLY REPORTED EQUATIONS.....	63
TABLE 32: GENDER BASED BLAND-ALTMAN APPROACH TO COMPARE MEASURED DIRECTLY VO_{2PEAK} ($ML.KG^{-1}.MIN^{-1}$) AND ESTIMATED VO_{2MAX} ($ML.KG^{-1}.MIN^{-1}$) FROM PREVIOUSLY REPORTED EQUATIONS.....	64
TABLE 33: GENDER BASED SPIROMETRIC (VO_{2PEAK} ; $ML.KG^{-1}.MIN^{-1}$ AND RER_{PEAK}), PERFORMANCE (TOTAL LAPS; N°) AND CARDIOVASCULAR (HR_{MAX} ; BPM) OUTPUTS IN TEST AND RETEST TRIALS (MEAN \pm SD); PEARSON PRODUCT-MOMENT CORRELATION (R) AND CONCORDANCE CORRELATION (R_C) BETWEEN TEST AND RETEST TRIALS.....	66

ABSTRACT

Several regression equations have been developed to predict cardiovascular fitness (CVF) based on the 20-m multistage shuttle-run test (MSRT) for children and youth. **Purpose:** The aim of this study was to compare the validity of the FITNESSGRAM's MSRT software reports and previously reported equations as predictors of CVF in Lisbon healthy elementary school children, by measuring the VO_{2peak} during the actual 20-m MSRT performance. **Methods:** Ninety subjects, 45 boys and 45 girls aged 8-10 years, performed the MSRT and their VO_{2peak} was measured in field using a portable gas analyzer (K4b², Cosmed, Italy). Height and weight were measured and %Fat was estimated from skinfold thickness. The total laps (TL) completed were recorded and used to calculate maximal speed (MS), and used as predictors in the equations. **Results:** Using Bland-Altman analytical approach, there were differences between measured and predicted VO_{2peak} . The FITNESSGRAM's reports had the smallest mean difference (d) ($ml \cdot kg^{-1} \cdot min^{-1}$) between measured and the estimated values (1.84), but also had the highest range ($ml \cdot kg^{-1} \cdot min^{-1}$) between upper and lower limits of agreement (LA) (28.6) and the highest slope (1.20) meaning that the equation overpredicted VO_{2peak} in unfit participants and underpredicted VO_{2peak} in fit participants. The lowest slope (0.44) and the lowest range of the LA (22.23) were observed using the equation of Fernhall et al (1998). The comparison using the equation of Leger et al (1988) provided a reasonably large d (4.69), one of the highest ranges (27.08) and the second uppermost slope (1.04). The equation of Matsuzaka et al (2004) including MS showed higher d (5.38) than the equation using TL ($d=4.23$). From the 3 equations of Barnett et al (1993), the equation including skinfold and MS showed the highest d (6.07) but the smallest range (24.09). **Conclusion:** Previously reported equations might be unsuitable for Lisbon healthy elementary school children. We suggest that the FITNESSGRAM's software reports or the equation of Fernhall et al (1998) may be the most appropriate. However, they still yield unsatisfactory accuracy. Our findings advocate that published equations predicting VO_{2peak} from MSRT may provide reasonable group estimates, but poor individual predictions of VO_{2peak} .

THEME'S PRESENTATION

INTRODUCTION

Cardiorespiratory fitness (CRF) is strongly associated with the clustering of cardiovascular disease (CVD) risk factors in children and youth and that association is independent of country, age and sex (Anderssen, Cooper et al. 2007). It is known that higher levels of physical activity and fitness in adults may protect against acute cardiac events by a variety of mechanisms including a reduced myocardial vulnerability to ventricular arrhythmias, reduced platelet stickiness and aggregability, and increased fibrinolysis (Boreham, Twisk et al. 2001). Aerobic fitness appears to be important in reducing the risk of cerebral and cardiovascular morbidity and mortality (Blair, Kohl et al. 1989), even in groups where the risk of heart disease and stroke is high (Ardern, Katzmarzyk et al. 2005). It is also related to better cognitive function in preadolescent children using a task which require variable amounts of executive control (Buck, Hillman et al. 2008).

Even though the state of knowledge, we live in a world where, i.e., in the United States the percentage of children and adolescents who are overweight has more than doubled in the past 30 years. Approximately 11% of children and adolescents are classified as overweight or obese. In Australia, obesity and overweight levels are reported to be between 19% and 23%. In Sweden, adolescents ages 12, 15, and 18 years old are reported to be 12.3%, 11.6%, and 11.4% overweight for boys, respectively, and 6.8%, 5.5%, and 4.8% overweight for girls, respectively (Vincent, Pangrazi et al. 2003). Cumulated results from 28 reports of changes in body mass index (BMI) in 5-16-year-old children and adolescents since 1980 from the US and Australia show that BMI has increased at a median rate of 0.6% per year, comparable to the rate of decline in aerobic performance (Tomkinson, Leger et al. 2003).

This evidence comes even from far east continents where in a sample of 23.5 million 6-to-19-year-olds from seven Asian countries tested between 1977 and 2003, there has been very little change in power and speed test performances of Asian children and adolescents in recent decades, yet alarmingly, there have been consistent declines in cardiovascular endurance fitness performance across all studied Asian nations over the past 10-15 years (Macfarlane and Tomkinson 2007).

Given that the American genotype has not changed substantially over the past two to three decades, we must look to the environment as the primary cause of the obesity epidemic. There has been either an increase or a very modest decline in total energy intake over the past two to three decades (Hill and Melanson 1999). Rowland (2002) concludes that if changes in aerobic fitness of children are declining, and if this fall is not just due to increases in body fat, it might be assumed that sedentary living plays a direct etiologic role. There are several lines of evidence suggesting that in children and adolescents the major factor is sloth, i.e. decreased physical activity. This is because: 1) energy intake appears to be relatively stable; 2) longitudinal studies suggest that activity levels of children and adolescents are declining; 3) there appears to have been an increase in inactivity (e.g. television watching) (Tomkinson, Leger et al. 2003). Several authors found that skinfold thickness increased in both boys and girls as the amount of television watched increased. They also found a relationship between television watching, physical activity, and body composition (Andersen, Crespo et al. 1998).

Overweight children have higher odds of becoming overweight adults than their normoponderal counterparts (Whitaker, Wright et al. 1997). Because risk factor levels in childhood predict risk factor levels in young adulthood, decreasing risk factors in children is an important health consideration on itself (Sallis and McKenzie 1991; Cooley and McNaughton 1999). However, there is a light of hope in this cheerless reality screening. Although modest, research has shown that physical activity patterns established in childhood often track into adulthood (Malina 2001). In addition, getting lower fit people to do something has more health consequences for our nation than getting already fit people to do more (Blair, Kohl et al. 1989).

The American Academy of Physical Education, the American Academy of Pediatric Committees on Sports Medicine, and the American College of Sports Medicine issued strongly worded statements highlighting the need for school physical education programs to adopt health related physical activity goals. In a society in which adult sedentary behaviour contributes substantially to the epidemic of cardiovascular and other chronic disease, there is a rationale for shifting the orientation of physical education to a health focus (Sallis and McKenzie 1991). Because physical activity confers significant protection from chronic diseases such as cardiovascular diseases and non-insulin-dependent diabetes mellitus and because it appears to reduce the risk of osteoporosis and some cancers, there is substantial interest in beginning the prevention of these adult

diseases during the first two decades of life through regular physical activity. In addition to these disease-prevention benefits, physical activity contributes to quality of life, psychological health, and the ability to meet physical work demands and engage in leisure activities (Sallis and Patrick 1994). For adult men, over 2000 calories per week in physical activity leads to important reductions in risk of death (Paffenbarger, Hyde et al. 1986), but how this caloric expenditure standard would apply to children is unclear (Sallis and McKenzie 1991). Some have suggested that physical activity of 1.75 (daily energy expenditure of 1.75 times de basal metabolic rate) should be our target (Hill and Melanson 1999).

It is now known that health behaviour change programs targeted at adults are expensive and typically show high rates of recidivism. Accordingly, many experts have recommended that primary prevention of chronic disease should be pursued through interventions that are directed to children (Pate, Baranowski et al. 1996). Because children receive a majority of their physical activity outside of physical education, a comprehensive approach to promoting children's physical activity must go beyond the confines of the school campus (Sallis and McKenzie 1991; Coleman, Tiller et al. 2005).

The Healthy People 2010 objectives clearly indicate that schools should provide an opportunity for preventing and treating obesity by promoting physical activity and healthy eating habits (Coleman, Tiller et al. 2005). In addition to skill development, elementary exercise behaviours rather than outcomes while intense efforts to improve physical conditioning are withheld until after puberty. Perhaps the chief training benefit should be to initiate activity behaviours and lifestyles for children rather than seeking or expecting significant improvements in VO_{2max} (Payne and Morrow 1993). To children, learning to be physically active might be a more powerful motivator for physical activity than the knowledge about the importance of being physically active (Sallis and McKenzie 1991).

In this way, it is expected that all health care providers regularly assess and counsel their children "patients" about physical activity. This assessment should enable health care providers to 1) categorize their patients as either already engaging in appropriate activity or in need of increased physical activity, 2) determine if a patient is at increased risk for health outcomes that can be beneficially affected by physical activity, and 3) identify barriers to either continued or increased physical activity the children might experience (Sallis and Patrick 1994).

One of the most relevant evolutions was the development of the FITNESSGRAM program (The Cooper Institute for Aerobics Research 1994). Due to its health-related evaluation system (aerobic fitness, body composition, muscular strength and endurance, and flexibility), this program suits on the principles of physical education. In fact, the information given to the students of its performance in the different tests and the respective interpretation based on health indicators might be an important extrinsic motivation factor towards a voluntary and joyful participation in the different activities. The informatization of the process is on the edge of the construction of a national database which may be fundamental for the adequacy of the programs to the student's needs.

The FITNESSGRAM includes two test items designed to test aerobic fitness: the 1 Mile Run/Walk (MRW) and the 20-m Multistage Shuttle Run Test (20-m MSRT) also called PACER test. It urges the necessity to validate every aspect of this program in each age, sex and regional groups in Portugal.

Each VO_{2max} estimation test must be validated directly or indirectly by the standard criterion, the gas analyser. This way, it is possible to assess the predictive power of the test. The direct validation assumes that the subject's gas exchanges are measured in simultaneous with the test performance. The indirect validation or cross validation implies that the subject performs two tests: the test itself and a standard test with known validity. The correlation between the two will establish the validity of the test (Safrit 1990). Furthermore, not only the test has to predict what is suppose to predict, as it has to be reliable, in other words, the results must be similar in repeated trials on the same individual, under similar circumstances (Hopkins 2000).

The 20-m MSRT represents an acceptable field assessment tool for CRF and has been repeatedly employed in different health (Bouziotas, Koutedakis et al. 2001) and fitness (Grant, Corbett et al. 1995) settings. The popularity of the 20-m MSRT is mainly attributed to its practical use for simultaneous measurement of large groups of individuals. The use of 20-m MSRT to predict VO_{2max} was first proposed by Leger and Lambert (1982). This was a development of their earlier track running test (Leger and Boucher 1980), which was re-designed to allow the test to be performed inside gymnasia. The original test, using

2-min incremental speed stages, was later modified to 1-min stages to reduce the overall test time and aid motivation of the participants (Leger and Gadoury 1989). The test has a number of variants, and small differences in the conduct of the test and the analysis of results make it difficult to compare data from different sources.

ISSUE'S DEFINITION

The aim of this study was to compare the validity of the FITNESSGRAM's MSRT software reports and previously reported equations as predictors of cardiovascular fitness (CVF) in Lisbon healthy elementary school children, by measuring the VO_{2peak} during the actual 20-m MSRT performance. Ninety subjects, 45 boys and 45 girls aged 8-10 years (in a balanced experimental design), performed the MSRT and their VO_{2peak} was measured directly in field using a portable gas analyzer (K4b², Cosmed, Rome, Italy).

We additionally assessed the reliability of the 20-m MSRT in a sub-group of subjects that performed the test twice, allowing a discussion around the learning effect.

This study emerges from the necessity to check the validity of the PACER test in Portuguese elementary school children. In agreement with the references gathered, this is the first study to settle on this issue using a portable gas analyzer in children of such a young age.

STUDY'S EXTENT

Ninety subjects (45 boys and 45 girls, ages 8 to 10 years old) took part in this study. They were recruited from a single elementary school "E.B 1 J.I Aqualva 1" where they all participated in Physical Education classes twice a week. The information gathered with each teacher reduced the chances of selecting an individual with a hazard clinical history.

The geographical concentration of the sample was not avoided for several reasons such as the availability of the portable gas analyzer (K4b², Cosmed, Rome, Italy), the reduced time and human resources induced by the holidays period of the subjects.

All the PACER tests were performed in the playing ground of the school E.B 1 J.I Aqualva 1, in a proper space exclusively delimited for the matter. The surface of approximately 375 m² was made of concrete.

Initially there was an intention to create a subgroup in order to determinate %Fat through DEXA (*Dual* energy X-ray absorptiometry). However, we ended up not doing it as logistical factors and the issue's specification itself minimized the relevance of such procedure.

ASSUMPTIONS

The following assumptions were considered:

- 1) All the subjects involved were healthy, with a low CVD risk and absence of chronic obstructive pulmonary disease;
- 2) Assuming that none of the subjects had a particular interest in the matter, and all being recruited by the author, there is no possible way to assure each individual's maximal effort during the test. Therefore, we presuppose that the effort and commitment of the subjects during the test was maximal.
- 3) In a similar way, it is assumed that none of the subjects has any current formal and specific athletic activity, certifying the inexistence of a pronounced discrepancy in the movement economy.

LIMITATIONS

The following limitations were considered:

- 1) The geographical concentration of the sample;
- 2) The loss of some HR data due to technical hitches in the transmission of the electric impulses originated by a non-desired alteration in the positioning of the HR band;

- 3) The convenience sample restrains the generalization of the results to the Portuguese children. We might speculate that the children who did not participate in this study were the ones who did not like physical activity or were not keen on sports. As a consequence, the results may overestimate the general fitness level of this population.

STUDY'S SIGNIFICANCE

A test is considered valid when it measures effectively what it is suppose to measure. The indirect validation of the 20-m MSRT in children and youth is well established (van Mechelen, Hlobil et al. 1986; Liu, Plowman et al. 1992; Mahoney 1992; Falgairette, Bedu et al. 1994; McVeigh, Payne et al. 1995; Armstrong and Welsman 2000; Boreham, Twisk et al. 2001; Matsuzaka, Takahashi et al. 2004). On the other hand, the direct validation has only been done sporadically using the retroextrapolation method (extrapolation to time zero of recovery of the exponential least squares regression of the first four 20-s recovery VO_2 values) (Mercier, L.A et al. 1983), or ingenious but nonetheless improvised portable devices (Oliveira 1998).

Additionally, of the mentioned studies hardly any report the existence or not of differences between mean determinate and mean predicted VO_2 values (Liu, Plowman et al. 1992; Oliveira 1998).

In the present study, far beyond the determination of the subject's VO_{2peak} , we intended to examine the validity of the FITNESSGRAM's 20-m Multistage Shuttle-Run Test (PACER test) software reports as a predictor of VO_{2peak} , confirm the results obtained by several authors in previous studies (Mercier, L.A et al. 1983; van Mechelen, Hlobil et al. 1986; Liu, Plowman et al. 1992; Mahoney 1992; Falgairette, Bedu et al. 1994; McVeigh, Payne et al. 1995; Oliveira 1998; Armstrong and Welsman 2000; Boreham, Twisk et al. 2001; Matsuzaka, Takahashi et al. 2004) and at last, scrutinize the most appropriate regression equation to estimate the participant's VO_{2peak} among the literature revised (Leger, Mercier et al. 1988; Barnett, C. et al. 1993; McVeigh, Payne et al. 1995; Fernhall, Pitetti et al. 1998; Oliveira 1998; Matsuzaka, Takahashi et al. 2004).

OPERATIONAL DEFINITIONS AND ABBREVIATIONS

PHYSICAL ACTIVITY

It is a broad term that describes any bodily movement produced by skeletal muscles that result in energy expenditure (Caspersen, Powell et al. 1985).

EXERCISE

It is a subset of physical activity that is planned, structured, and repetitive and has as a final or an intermediate objective the improvement or maintenance of physical fitness (Caspersen, Powell et al. 1985).

PHYSICAL FITNESS

It is a set of attributes that are either health or skill related. The degree to which people have these attributes can be measured with specific tests (Caspersen, Powell et al. 1985). Constitutes an essential part of the assessment of health and involves different dimensions such as maximum aerobic capacity (usually expressed as VO_{2max}), neuromuscular capacities such as strength, agility, speed, flexibility, balance, and coordination, and often body composition (Cauderay, Narring et al. 2000).

PHYSICAL CONDITIONING

It is the process by which exercise, repeated during weeks and months, induces morphologic and functional changes in body tissues and systems (Bar-Or 1983). "Conditioning" and "Training" are often used interchangeably and rather loosely in reference to a long term exercise regimens and their effects.

MAXIMAL OXYGEN UPTAKE

The maximal oxygen uptake or maximal aerobic power (VO_{2max}) is defined as the highest oxygen uptake the individual can attain during exercise engaging large muscle groups while breathing air at sea level, exercise time 2 to 6 min, depending on the type of exercise. It represents the maximal ability of the individual to take up, transports, and utilize oxygen by the working muscle (Astrand, Stromme et al. 2003).

PEAK OXYGEN UPTAKE

In adults VO_2 reaches a plateau at VO_{2max} and no further increase is observed with additional increases in the rate of work. This plateau is infrequently found when testing children and VO_{2max} cannot be precisely determined in many pediatric studies. Thus, the peak observed oxygen uptake (VO_{2peak}) is used instead (Washington, Bricker et al. 1994).

RESPIRATORY EXCHANGE RATIO

It is the ratio of the net output of carbon dioxide to the simultaneous net uptake of oxygen at a given site, calculated from the quotient: $RER = CO_{2\text{ produced}} / O_{2\text{ consumed}}$

CARDIAC OUTPUT

It is the volume of blood being pumped by the heart, in particular by a ventricle, in a minute. This is measured in $dm^3 \cdot min^{-1}$ and it is calculated from the expression: $Q = SV \times HR$, where SV means Stroke Volume and HR means Heart rate.

STROKE VOLUME

It is the volume of blood pumped by the right/left ventricle of the heart in one contraction. Specifically, it is the volume of blood ejected from ventricles during systole. The stroke volume is not all of the blood contained in the left ventricle. Normally, only about two-thirds of the blood in the ventricle is put out with each beat. This measure is calculated from the expression: $SV = EDV - ESV$, where EDV means End-Diastolic Volume and ESV means End-Systolic Volume.

ABBREVIATIONS

ATP	Adenosine Triphosphate
(a-v)O ₂	Arteriovenous O ₂ difference
BMI	Body Mass Index
BP	Blood Pressure
bpm	beats per minute
CHD	Coronary Heart Disease
CR	Cardiorespiratory
CRF	Cardiorespiratory Fitness
CVD	Cardiovascular Disease
CVF	Cardiovascular Fitness
<i>d</i>	Mean Difference (ml.kg ⁻¹ .min ⁻¹)
DBP	Diastolic Blood Pressure
FECO ₂	Fractional Concentrations of Carbon Dioxide
FE _{O₂}	Fractional Concentrations of Oxygen
F _R	Respiratory Frequency
HR	Heart Rate
HR _{max}	Maximal Heart Rate
HR _{max1krw}	Maximal Heart Rate attained during 1 Km run/walk test
HR _{maxMSRT}	Maximal Heart Rate attained during the 20-m Multistage Shuttle Run Test
HR _{peak}	Peak Heart Rate
K4b ²	Cosmed Portable gas analyser
LA	Limits of Agreement
LSD	Least Significant Difference
MAS	Maximal Aerobic Speed
MR	Mental Retardation
MRS	Maximal Running Speed
MS	Maximal Speed
MET	Metabolic Equivalent
MRW	1 Mile Run/Walk test
20-m MSRT	20-m Multistage Shuttle Run Test
%Fat	Percentage of Fat
PA	Physical Activity
PACER test	FITNESSGRAM's 20-m Multistage Shuttle-Run Test version
PETCO ₂	End-Tidal Partial Pressure of Carbon Dioxide
PHV	Peak Height Velocity
<i>Q</i>	Cardiac output
Q ₁	Boxplot's lower Quartile
Q ₂	Boxplot's Median
Q ₃	Boxplot's upper Quartile
<i>r</i>	Pearson product-moment correlation coefficient or Spearman's rho
<i>r_c</i>	Concordance correlation coefficient
RER	Respiratory Exchange Ratio
RER _{peak}	Peak Respiratory Exchange Ratio
SBP	Systolic Blood Pressure
SEE	Standard Error of the Estimate
SV	Stroke Volume
TL	Total Number of Laps achieved in the 20-m MSRT
TM	Treadmill
TMT	Treadmill Test
TS	Total Number of Stages achieved in the 20-m MSRT
Tric _{skinf}	Triceps Skinfold Thickness
Tric&Sub _{skinf}	Sum of the Triceps and Subscapular Skinfold Thickness
V _E	Ventilation
V _T	Tidal Volume
VCO ₂	Carbon Dioxide Production
VO ₂	Oxygen uptake
VO _{2max}	Maximal Oxygen Uptake
VO _{2peak}	Peak Oxygen Uptake
VO _{2plateau}	Plateau in Oxygen Uptake
YIET	Yo-Yo Intermittent Endurance Test
Yr	Year

LITERATURE REVISION

INTRODUCTION

This revision starts with an overview of the major children's physiological responses to exercise, followed by the importance of the determination of VO_{2max} attending to health-related aspects. From there, we will proceed with an analysis of the definition of VO_{2max} and we will refer the criterion which may define a maximal test. In addition, relative expression formats of VO_{2max} are mentioned.

At long last, we will revise the articles related to the characteristics, validity and reliability of the 20-m MSRT, as well as, aspects related to the construction of normative tables to VO_{2max} from 20-m MSRT results.

Any study that used the 20-m MSRT to assess the fitness of healthy young people aged 6-19 years was considered as a candidate for this work. Data on elite young sportspeople or groups with specific disabilities or disease conditions (e.g. obesity) were not considered. Studies were located by searching online libraries, PubMed, Web of Knowledge, B-On, using the following keywords: shuttle, shuttle run, MSFT, 20MST, PACER, 20mSRT, beep test, multi-stage, in combination with modifiers child, children, pre-adolescent, adolescent, adolescence, pubescent, pubescence boy, girl, young, youth and infant. No language or date restrictions were applied.

PHYSIOLOGICAL ASPECTS

Knowledge of the normal course of developmental exercise physiology is important to the understanding of aerobic fitness in the pediatric population. To this end, a series of investigations have provided a composite picture of the changes in metabolic, cardiovascular and pulmonary responses to exercise in growing children.

METABOLIC RESPONSE TO EXERCISE IN CHILDREN

For mechanical energy to be released at the myofibrillar level and effect muscle contraction, splitting of adenosine triphosphate (ATP) must take place. This high energy compound is available in small quantities (about 4-5 mmol/kg wet weight) in the resting muscle. However, once contractions start, there is an immediate need for reinforcement of ATP. This can be supplied from: 1) limited stores of creatine phosphate, 2) glycolysis, or 3) the tricarboxylic acid (Krebs) cycle. The former two sources do not require addition of O_2 and are therefore called anaerobic. The latter requires O_2 and is termed aerobic. Muscle contractions that result from anaerobic reactions cannot be sustained longer than 40-50 sec. In contrast, muscle contractions utilizing aerobic energy turnover can last many minutes or even hours. Even though most activities utilize both aerobic and anaerobic pathways, in the jargon of sports scientists physical tasks are subdivided into aerobic type and anaerobic type activities. The former include long-distance running, swimming, cycling, cross-country skiing, and other endurance-requiring tasks. The latter include sprinting, jumping, throwing, and other sports where the required power intensity is high and the duration short (Bar-Or 1983).

Peak oxygen uptake depends upon the pulmonary, cardiovascular, and hematological components of O_2 delivery and the oxidative mechanisms of exercising muscle. It is a function of cardiac output and arteriovenous O_2 difference ($(a-v)O_2$) (Armstrong and Welsman 2000).

MAXIMAL AEROBIC POWER

The maximal oxygen uptake or maximal aerobic power (VO_{2max}) is defined as the highest oxygen uptake the individual can attain during exercise engaging large muscle groups while breathing air at sea level, exercise time 2 to 6 min, depending on the type of exercise (Astrand, Stromme et al. 2003).

Relatively to the limiting factors of VO_{2max} , there are two main theories: the peripheral limitation and the central limitation. Eriksson et al (1971) affirmed that the results obtained favoured the hypothesis that the heart or central circulation did not limit the oxygen transportation, but this is rather limited by the ability of the muscles to consume oxygen. Green and Patla

(1992) for instance, stated that the utilization of oxygen by the working muscle is not the reason limiting the VO_{2max} that can be attained. In their opinion, processes associated with the delivery of oxygen to the working muscle, collectively referred to as central processes, are limiting. Research involving determination of VO_{2peak} is vast. Yet, the interpretation of aerobic fitness during childhood and adolescence is still shrouded with controversy (Armstrong and Welsman 2000).

Maximal VO_2 is considered one of the most valid parameters to evaluate cardiorespiratory (CR) capacity (Massicotte, Gauthier et al. 1985). It reproduces the functional capacity of the lungs, of the cardiovascular (CV) system and of the muscle's mitochondria (Saltin and Strange 1992; Sutton 1992).

Children and adults can be considered to be geometrically similar, because a uniform change in the length dimension would make them identical (Cooke, McDonagh et al. 1991). Perhaps that is the reason why trained preadolescent children have maximal O_2 intakes per kilogram body mass that are comparable to those of adults. These finding might be in the origin of the hypothesis that there is no difference in efficiency between children and adults in the performance of external work (Cooke, McDonagh et al. 1991). However, at a given running speed they utilize more O_2 per kilogram body mass than adults (Davies 1980) (see Table 1).

There are no normative data on VO_2 values for a national sample of children. A summary of small-scale studies reported mean values ranging from 45-60 $ml.Kg^{-1}.min^{-1}$ (Krahenbuhl and Pangrazi 1983), and noted that these values remain stable through adolescence for boys but decline with age for girls (Safrit 1990). Until the age of 12, values grow at the same rate in both genders, even though boys have higher values as early as the age of 5. While VO_{2max} of boys keep increasing until the age of about 18, it hardly develops beyond age 14 in girls. Maximal VO_2 increases by 80% between the ages of 8 and 16 years in girls and by 150% in boys over the same period (Armstrong and Welsman 2000). Maximal VO_2 is strongly related to lean body mass which explains to a great extent, the above sex-related differences (Bar-Or 1983). The values are much higher for boys than girls and the VO_{2max} remains constant for boys between 10-17 years old, while a gradual decrease is seen for girls at 13 years old. This decrease in VO_{2max} per Kg of body weight in the latter group is partially due to the increase of adipose tissue that accompanies the onset of puberty (Massicotte, Gauthier et al. 1985). If we compare boys' and girls' VO_{2peak} ($l.min^{-1}$) at various chronological ages, data suggest that boys' values are about 12% higher than girls' at 10 years and that the gender difference becomes more pronounced during the teen years, typically reaching 37% at age 16 years. Girls' lower levels of physical activity (Anderssen, Cooper et al. 2007) have been proposed as a reason for sex differences in VO_{2peak} (Krahenbuhl, Skinner et al. 1989) but the evidence relating habitual physical activity to VO_{2peak} is unconvincing (Armstrong and Welsman 2000).

Table 1: Unique Metabolic Responses of Children Relative to Adults. Adapted from ACSM (2005).

Variable	Submaximal Exercise	Maximal exercise
VO_2 ($l.min^{-1}$)	Lower	Lower
VO_2 ($ml.Kg^{-1}.min^{-1}$)	Higher	Higher or equal

When the VO_{2max} of children and adolescents of different ages but of the same body weight or body height is compared, it is positively related to age. Thus, the VO_{2max} depends also on the maturity of the individual and not only on his or her body dimensions (Bar-Or 1983; Mota, Guerra et al. 2002). There is some evidence from longitudinal studies that the greatest increase in VO_{2peak} is associated with the attainment of peak height velocity (PHV) in both boys and girls (Kobayashi, Kitamura et al. 1978). After adjusting for age within each stage of maturity, VO_{2max} increases with stage of maturity in males, but decrease with stage of maturity in females. In addition, the CRF expressed either as VO_{2max} per unit body mass or 20-m MSRT-time is inversely and significantly associated with %Fat in both genders. It is a factor accounting 10% in girls and 22% in boys, for the variation in VO_{2max} (Mota, Guerra et al. 2002).

Trowbridge et al (1997) observed that VO_{2max} was significantly lower in African-American compared with Caucasian prepubertal boys and girls. Accordingly to the author, this difference could not be explained by differences in body composition, including fat mass, total soft lean tissue mass or leg soft lean tissue mass. The difference in VO_2 was not apparent during rest or submaximal exercise and was only observed at maximal effort of exercise.

MECHANICAL EFFICIENCY AND ECONOMY OF MOVEMENT

When muscle contraction results in movement, the calorie equivalent of this work is only some 20-25% of the chemical energy utilized during the contraction. The other 75-80% is converted into heat. Interindividual differences in O_2 uptake during walking or running do not necessarily denote differences in mechanical efficiency at the cellular level but rather a difference in the economy of locomotion (Bar-Or 1983).

Daniels et al (1978) showed that young boys are less economical than adults, therefore, they may be unable to sustain the same relative (percent VO_{2max}) performance as adults for runs of similar duration. A few years later, Bar-Or (1983) defined that the mechanical efficiency of cycling is similar in children, adolescents, and adults, ranging between 18 and 30%. In contrast, O_2 cost during walking and running is higher among children, when expressed per kilogram body weight. Cooke et al (1991) reported that the mean difference between children (10-12 years old) and adults in VO_2 of unloaded running is 7 $ml.kg^{-1}.min^{-1}$. These last findings apparently support the hypothesis that there is no difference in efficiency between children and adults in the performance of external work.

Whatever the hypothesis, VO_{2max} appears to account for less variation in children's running performance than has been reported for adults (Krahenbuhl and Pangrazi 1983). Certain studies have also indicated that factors such as biological maturity, height, weight, % body fat and mechanical efficiency could influence the physical performances of children, and might be partially responsible for the important variations between sex and age groups (Massicotte, Gauthier et al. 1985). In fact, maximal running speed, anaerobic capacity, running economy, fractional utilization of VO_{2max} (percent VO_{2max}), and the velocity at which onset of blood lactate accumulation occurs (V_{OBLA}) have been found to relate to distance-running performance in children (Cureton, Boileau et al. 1977; Krahenbuhl and Pangrazi 1983; Olds and Dollman 2004).

In addition, the higher O_2 cost in young children cannot be explained merely by difference in resting metabolism, which is only some 1-2 $ml.kg^{-1}.min^{-1}$, but rather by their relatively "wasteful" gait while walking or running. With training, running becomes more economical in children and adolescents, as reflected by a lower O_2 cost. No matter the underlying mechanism, the high metabolic cost of walking or running makes the small child less of an "aerobic machine" than might be expected from his high maximal O_2 uptake. If one takes the difference between VO_{2max} and the O_2 uptake needed for a given task to represent metabolic reserve, one can see that children are at a disadvantage (Bar-Or 1983).

ANAEROBIC CHARACTERISTICS

Whether or not children have inferior maximal aerobic power is still debatable. This is not the case with anaerobic performance which is lower in the child in absolute and relative terms alike, whether scaled to body weight, height squared, or lean body mass. Bar-Or (1983) using the Windgate anaerobic test, verified that even when normalized for body weight, however, the power produced by an 8 year old boy is still only 70% of that generated by an 11 year old boy. As we see, while maximal aerobic power does not change (in boys) or even decrease (in girls) with age, there is a progressive growth-related in anaerobic performance.

Differences in maximal aerobic power can be accounted for by the mass of active muscle tissue. This is not so with anaerobic capacity. The markedly lower anaerobic capacity of young child reflects, to a great extent, a qualitative deficiency in his muscle or in the recruitment of motor units within the muscle, which reduces its ability to perform anaerobically (Eriksson, Grimby et al. 1971; Bar-Or 1983).

O_2 UPTAKE TRANSIENTS

In line with the notion that children are less capable than adults of exerting anaerobically is the pattern of their O_2 uptake transients when they increase metabolic rate. At first, the aerobic supply of additional energy lags behind the actual demands for energy and an " O_2 deficit" is contracted. The balance of chemical energy at these initial stages is facilitated by anaerobic pathways. During submaximal exercise, the aerobic energy supply will gradually catch up with the demand and a new metabolic "steady-state" will be established within 2-5 min.

Children have shorter O_2 uptake transients than adults (Freedson, Gilliam et al. 1981). The O_2 deficit of 10 to 11 year old boys is smaller than among young adults. Even more marked is the high rate at which boys increase their O_2 uptake compared with adults. Children in this age-group reached 55% of their final O_2 uptake within 30 sec and a steady state within 2 min. The adults (20-22 years old) reached only 33% of their final O_2 uptake within the first 30 sec and required some 3-4 min to reach a steady-state (Máček and J. 1980).

An interesting question is whether due to their shorter O_2 transients, children do not need to resort much to anaerobic pathways (hence the smaller O_2 deficit and lactate production), or whether these shorter transients are compensatory for their low glycolytic capacity. Another possibility is that the shorter transients in children are a reflection of a smaller body and the resulting shorter circulation time (Cumming 1978).

CARDIOVASCULAR RESPONSE TO EXERCISE IN CHILDREN

The main role of the cardiovascular system during exercise is in the transport of additional O_2 to the exercising muscles and of CO_2 from the muscles. Other roles include transport of nutrients, metabolites, and hormones, retention of osmotic and acid-base balance, and convection of heat from body core to periphery.

CARDIAC OUTPUT AND STROKE VOLUME

As in adults, the cardiac output (Q) of children rises at the beginning of exercise, or upon transition to a higher level of exercise. A new steady-state is established within a few minutes. Typically, children have a markedly lower stroke volume (SV) and a higher heart rate (HR) than adults. The end result is a somewhat lower Q at each metabolic level (Bar-Or, Shephard et al. 1971).

Stroke volume responses to submaximal exercise are more secure and demonstrate that in upright exercise progressive increases in intensity promote rises in SV to about 40 to 60% of VO_{2peak} when SV levels off (Eriksson, Grimby et al. 1971). The same authors identified an increase in the Q when subjects went from rest to exercise, but absolute values for Q at a given O_2 uptake in these 13 to 14 year old boys were 1-2 $liters \cdot min^{-1}$ lower than those observed in material comprising adult men.

The few data available are consistent and indicate that Q at VO_{2peak} increases with age, with values rising from 12.5 to 21.1 $l \cdot min^{-1}$ in males and 10.5 to 15.5 $l \cdot min^{-1}$ in females, over the 9 to 20 years age range (Armstrong and Welsman 2000). The rise in VO_{2peak} with age during childhood and adolescence is, it appears, primarily due to an increase in SV and therefore Q . Arteriovenous oxygen difference ($(a-v)O_2$) may increase with age, but additional insights into oxygen delivery and subsequent oxidative metabolism at VO_{2peak} may be dependent on technological advances in non-invasive methodology (Armstrong and Welsman 2000).

Explanations are speculative but whereas there is no evidence to indicate sex differences in either HR or $(a-v)O_2$ at VO_{2peak} , boys have been observed to have greater SV than girls (Armstrong and Welsman 2000). Boys' greater muscle mass during puberty is highly likely to contribute to the progressive increase in sex differences in VO_{2peak} , as it will not only facilitate the use of O_2 during exercise but also enhance the venous return to the heart, through the peripheral muscle pump, and therefore augment SV (Rowland, Goff et al. 2000). Bar-Or et al (1971), in a sample of subjects with 29 girls and 27 boys, 10-13 years old, verified that for a given VO_2 level the girls had a higher Q , higher HR, and a lower $(a-v)O_2$, than boys.

A potential handicap due to low Q may exist, however, during maximal exercise when peripheral O_2 extraction can no longer rise or when the child is exposed to combined stress of exercise and extreme heat (Bar-Or 1983).

HEART RATE AND EXERCISE

The immediate response of the cardiovascular system to exercise is an increase in HR due to a decrease in vagal tone. This response is followed by an increase in sympathetic outflow to the heart and systemic blood vessels. During dynamic exercise, the HR increases linearly with the rate of work. During low to moderate levels of exercise at a constant work rate, the HR reaches a steady state within 1 min and increases proportional to the rate of work in subjects with a normal sinus node.

Young children compensate for a smaller heart and lower SV by an increased HR at a given rate of work. Thus, they attain higher maximal HR than adults (Washington, Bricker et al. 1994).

Peak Heart Rate (HR_{peak}) and Respiratory Exchange Rate (RER) have been utilized as markers of maximal exercise effort, since both rise with increasing work intensity during a progressive test (Issekutz, Birkhead et al. 1962).

The HR_{max} of children and adolescents ranges between 195 and 215 bpm (see Table 3) and starts declining with age only after maturity has been reached. Such decline is independent of sex, level of conditioning, climate, or other environmental conditions. It is equivalent to 0.7-0.8 bpm per year (Bar-Or, Shephard et al. 1971; Bar-Or 1983; Rowland and Cunningham 1997; Armstrong and Welsman 2000). It is believed that the circulatory responses to submaximal exercise of 10 to 13 year old children are similar to those of adults, but the child operates at lower Q and higher $(a-v)O_2$ than adults at any given metabolic level (Bar-Or, Shephard et al. 1971) (see Table 4).

As VO_{2max} is approached, increases in the VO_2 of children are achieved more by increases in HR as $(a-v)O_2$ difference reaches a maximum value. Conversely, adults perform submaximal work at a relatively lower $(a-v)O_2$ difference than children and increases in VO_2 as VO_{2max} is approached are brought about more by an increase in $(a-v)O_2$ difference and relatively less by an increase in HR (Bar-Or, Shephard et al. 1971) (see Table 2).

Table 2: Some Factors known to Affect HR Response to Exercise Among Children and Adolescents (Bar-Or 1983).

	Submaximal HR (at a given exercise)	Maximal HR
Age	Young > Old	No effect
Sex	Females > Males	No effect
Adiposity	Obese > Lean	No effect
Climatic stress	(+)	No effect
Emotional stress	(+)	No effect
Active muscle mass	Small > Large	Large > Small
Body position	Upright > Supine	Upright > Supine
Conditioning	(-)	No effect or slight (-)
Deconditioning	(+)	No effect
Heat acclimatization	(-)	No effect
Habituation	(-)	No effect

As a rule, women have a higher HR than men at a given exercise level. Traditionally this difference has been attributed to the low blood haemoglobin concentration of the mature female. However, a higher HR is also found among preadolescent girls whose haemoglobin level is not different from that of boys and even among children as young as 6 years old (Bar-Or, Shephard et al. 1971). The degree of relative tachycardia in females ranges between 10 and 20 bpm. It could result from a lower SV or from lower levels of habitual activity (Gilliam, Freedson et al. 1981).

Submaximal HR in children declines with age. Heart rate can be as much as 30-40 bpm higher in an 8 year old child than it is in an 18 year old performing the same absolute task. Such a difference is partially due to the greater relative exercise intensity performed by the younger children, but is also found at equal relative metabolic loads (Bar-Or 1983).

Table 3: Peak Heart Rate in Children and Adolescents in Previous Reported Studies. Adapted from Washington et al (1994).

Author, year	Age	Population	Sex	HR _{peak}	Protocol
Astrand (1952)	7-9	Scandinavia	Male	208	Cycle ergometer
	7-9		Female	211	
	10-11		Male	211	
	10-11		Female	209	
	12-13		Male	205	
	12-13		Female	207	
Goldberg (1966)	6-16	USA	Male	194 ± 9	Continuous, cycle ergometer
			Female	193 ± 8	
Wilmore (1967)	7-9	USA	Female	195	Intermittent, multistage, cycle ergometer
	10-11		196		
	12-13		194		
Shepard (1969)	11-13	Canada	Both	193	Continuous, progressive, treadmill
Eriksson et al (1971)	6-15	Scandinavia	Male	200.5 ± 2.9	Continuous, cycle ergometer
Hermansen (1971)	10-12	Scandinavia	Male	205.9-206.6	
Skinner (1971)	6-15	USA	Male	197.7-200.9	Continuous, progressive, treadmill
			Female	203.0-203.5	
			Male	199.1	Intermittent, multistage, treadmill
			Female	202.7	
Wilmore (1974)	8-12	USA	Male	196.8 ± 7.7	
Cumming (1977)	5-16	Canada	Male	170 ± 17	Intermittent, graded, supine cycle
			Female	174 ± 11	
Cumming (1978)	4-18	Canada	Both	193-206	Bruce, treadmill
Lock (1978)	5-16	USA	Both	142 ± 4.9	Fixed rate of work, submaximal, supine cycle
Riopel (1979)	4-21	USA, black	Male	179-186	Modified Balke, treadmill
			Female	181-194	
		USA, white	Male	183-192	
			Female	187-191	
James (1980)	5-33	USA 95% white + 5% black	Both	187-199	Continuous, progressive, constant workload, electrically braked, cycle ergometer
Alpert (1982)	6-15	USA, black	Male	188-194	Continuous, graded Mechanically braked, cycle ergometer
			Female	185-195	
		USA, white	Male	191-194	
			Female	194-195	
Wilmore (1982)	13-15		Male	202.1 ± 8.5	
Massicotte et al (1985)	10-17	Canada	Male	194.5 ± 6.9-200.4 ± 6.1	Progressive, cycle ergometer
			Female	193.4 ± 6.3-200.9 ± 5.1	
Mahon and Vaccaro (1989)	10-14	USA	Male	188.0 ± 4.2	Progressive, treadmill
T. W. Rowland and Cunningham (1992)	7-10		Both	No plateau: 199 ± 5.0 Plateau: 201 ± 5.0	Graded, treadmill
Liu et al (1992)	12-15	USA	Male	194.8 ± 14.6	Progressive, treadmill
			Female	195.5 ± 10.3	
Barnett et al (1993)	12-17	Hong Kong	Male	200.0 ± 9.0	
			Female	203.0 ± 7.0	
T. W. Rowland (1993)	10-13	USA	Both	204 ± 7.0	Treadmill

Falgairrette et al (1994)	10-11	France	Male	203 ± 8.0	Cycle Ergometer
McVeigh et al (1995)	M: 13,68±0,3 F: 13,68±0,3	Scotland	Male Female	203.32 ± 8.8 201.32 ± 0.2	Shuttle run test
Rogers et al (1995)	Pre-pubertal 6,6-14,2	USA	Both	Prepuberty M: 201 ± 10 F: 207 ± 9.0 Puberty M: 200 ± 10 F: 204 ± 6.0	Progressive, treadmill
Baquet et al (2002)	8-11	France	Both	C: 205 ± 8.0 E: 202 ± 8.0	Shuttle run test
Pitetti et al (2002)	8-15	USA	Male Female	189 ± 9.0 199 ± 8.0	Progressive, treadmill
Matsuzaka et al (2004)	8-17	Japan	Males Females:	8-10: 196.5 ± 7.0 11-13: 200.9 ± 5.5 8-10: 194.7 ± 9.5 11-13: 196.3 ± 8.4	Graded, treadmill

Age in years and HR_{max} in bpm.

ARTERIAL BLOOD PRESSURE

Children's arterial blood pressure (BP) response to exercise follows a pattern similar to adults. The latter, normally have their systolic blood pressure (SBP) increased between 6-8 mmHg for each 1 Met augment in work, whereas in children the increases are 2-3 mmHg and in adolescents 4-6 mmHg (ACSM 2005).

In rhythmic exercise, the higher contractile forces of the myocardium cause an increase in intraventricular systolic pressure. This is one mechanism by which more blood is made available to the periphery. Such an increase is manifested, among other functions, by a rise in arterial SBP. In contrast, DBP, which depends primarily on the peripheral vascular resistance, changes little with exercise and its direction cannot be predicted. When rhythmic exercise is performed, the rise in systolic SBP is proportional to exercise intensity and to the overall metabolic level. Such pattern is operative in all healthy individuals, irrespective of age. A lower exercise blood pressure in the young child is in line with the lower Q and SV (Bar-Or 1983).

However, the blood pressure responses to static contraction are different from those that occur during rhythmic contractions. Most striking is the dissociation between metabolic demands and blood pressure: both SBP and DBP increase above and beyond those expected merely from an increase in metabolism. Whenever the static effort exceeds some 20% of the maximal voluntary contraction of the respective muscle group, a steady-state cannot be reached. Fatigue keeps increasing and exhaustion takes place within minutes, even though HR may remain as low as 110-120 bpm (Lind 1970).

Table 4: Unique Cardiovascular Responses of Children Relative to Adults. Adapted from ACSM (2005).

	Submaximal Exercise	Maximal exercise
Heart Rate	Higher	Higher
Cardiac Output	Lower	Lower
Stroke Volume	Lower	Lower
Lactate Concentration	Lower	Lower

PULMONARY RESPONSE TO EXERCISE IN CHILDREN

PULMONARY VENTILATION

It is well documented that ventilation (V_E) at rest is higher in children than in adults (Gaultier, Perret et al. 1981) (see Table 5). The difference in ventilation between children and adults, both at rest and during exercise, even though blood gas tensions are similar, may be due to morphological, metabolic, and central nervous system factors (Gratas-Delamarche, Mercier et al. 1993).

In absolute terms, ventilation grows with age. Whereas a 6 year old child may reach a maximal pulmonary ventilation of 30-40 l.min⁻¹, a young adult can reach 100-120 l.min⁻¹ and more. When expressed per kilogram body weight, maximal ventilation is about the same in children, adolescents and adults. Submaximal ventilation per kilogram, on the other hand, is higher in children, diminishing regularly with age.

During progressive exercise, pulmonary ventilation increases linearly with the metabolic rate until about 60% of VO_{2max} when a ventilatory “breaking point” is reached, beyond which ventilation rises in a more accelerated manner. Such accelerated ventilation parallels, and is presumably triggered by, an increasing concentration of blood lactate and hydrogen ions. The ventilator breaking point is related to, and often coincides with, the “anaerobic threshold” (Bar-Or 1983).

RESPIRATORY FREQUENCY AND TIDAL VOLUME

Exercising children have a markedly higher respiratory frequency (F_R) than older individuals who perform the same task, maximal or submaximal. Compared with adolescents and adults, children respond to exercise with relative tachypnea and with shallow respirations (Bar-Or 1983). Gratas-Delamarche et al (1993) results consistently show a higher respiratory centre response to CO₂ in children than in adults. First, the relatively rapid and shallow breathing does not allow children to wash out alveolar air as efficiently as adults. Second, a high F_R leads to an increase in the total mechanical work of breathing due to an increase in viscous and turbulent work. These two effects oblige children to breathe more than adults for a similar metabolic level expressed per kilogram body mass.

In a 5 year longitudinal descriptive study, T. W. Rowland and Cunningham (1997) provide a picture of ventilatory responses to exercise in children that generally supports that created by previous cross-sectional studies. Tidal Volume (V_T) increases in proportion to body mass as the child grows, but the F_R at a given submaximal work load progressively declines. As a result, submaximal V_E rises slower with age than would be expected for body mass. This is also reflected in a steady fall in the V_E response to a given submaximal metabolic rate (ie, decline in (V_E/VO₂) with age. When the influence of sex was examined, only the magnitude of the V_E/VO₂ was observed to be gender related, with the girls demonstrating a greater value at all years. The explanation for this effect of gender on ventilator efficiency is unknown. Virtually the same patterns were observed at maximal exercise. Tidal Volume per kilogram remained stable while F_R declined. The scaling exponent of 0.92 for V_E relative to mass indicates that body mass rose more rapidly than V_E during the 5 year study. At maximal exercise, however, the fall the maximal V_E per kilogram was not statistically significant. V_E/VO₂ fell with advancing age only in the boys, with greater values again evident in the girls.

VENTILATORY EQUIVALENT

During exercise the ventilatory equivalent to oxygen is also higher in children (Gratas-Delamarche, Mercier et al. 1993) (see Table 5). This ratio, which represents the number of litres of air needed to ventilate the lungs in order to supply 1 litre of O₂, is a numerical expression of ventilatory efficiency. A high ratio reflects low efficiency. Ventilatory equivalent has been shown to decrease with age, whether during submaximal or maximal exercise. The major implication to children of less efficient ventilation is the greater cost O₂ of respiration (Bar-Or 1983).

Table 5: Unique Pulmonary Responses of Children Relative to Adults. Adapted from ACSM (2005).

Variable	Submaximal Exercise	Maximal exercise
Tidal Volume	Lower	Lower
Ventilation	Lower	Lower
V _E /VO ₂	Lower	Lower
RER	Higher	Higher

CONDITIONING AND TRAINING

CONDITIONING AND METABOLISM

Baquet et al (2002) findings from their study were that a 7 week training program significantly improved children’s absolute VO_{2peak} and relative to body mass VO_{2peak}. Payne and Morrow (1993) cited Bar-Or 1989 to suggest that children’s aerobic

trainability prior to puberty was less than expected and to reinforce their results in which the aerobic benefit of training revealed small-to-moderate for the under 13 years old age group, even though their athletic performance may improve.

In a review, Borms (1986) referred that the positive effect of endurance-type training programs on VO_{2max} is not consistently shown in studies on children. The author further stated that although improvements in running performance may be demonstrated, this often occurs without concurrent improvements in VO_{2max} . How can one reconcile an improvement in long-distance running with no concomitant increase in maximal O_2 uptake? One possibility is that the child improves his running style such that his movements are mechanically more economical. Another explanation is an improvement in anaerobic capacity, even though the conditioning program is considered aerobic. It is also possible that maximal aerobic power of the prepubescent does improve but maximal O_2 uptake is not a valid or sensitive indicator to show it. A final reason for the apparent lack of conditioning effects in prepubescent could be the high habitual activity of the "controls". It is quite possible that even though they do not participate in the regimented program, these children are so active in their free time that intergroup differences are quite negligible (Bar-Or 1983).

CONDITIONING AND CARDIOVASCULAR SYSTEM

Stroke volume at rest and during all levels of exercise increase with conditioning in pubertal children and in adults. In contrast, in prepubertal children conditioning does not seem to be accompanied by changes in SV. An increased SV may reflect higher blood volume and improved venous return, but it may also reflect improved myocardial contractility. Although SV is related to heart volume, the myocardial mass of children does not seem to affect their athletic capability. Total blood volume and total hemoglobin are higher in trained boys than in untrained ones and they increase as a result of conditioning. A higher volume facilitates a higher venous return as well as more effective heat convection from body core to periphery. Hemoglobin concentration, on the other hand, does not increase with conditioning (Bar-Or 1983).

A reduction of resting submaximal HR is a most sensitive response to conditioning and often precedes any changes in VO_{2max} (Rosenthal and Bush 1998). The reduction in HR_{max} is milder, seldom exceeding 5-7 bpm. The mechanism of conditioning-induced bradycardia is not clear. It may be secondary to an increase in SV, but it also reflects a stronger parasympathetic and a weaker sympathetic drive (Bar-Or 1983) (see Table 6).

Maximal Q increases with training in children and in adults, in proportion to the increase in VO_{2max} . In adults, the maximal (a-v) O_2 increases with conditioning, reflecting an increased peripheral O_2 extraction. In children, however, neither the submaximal or maximal (a-v) O_2 changed with conditioning (Bar-Or 1983).

Systolic arterial blood pressure at rest decreased in adults following conditioning, by about 7 mmHg. Blood Pressure during submaximal exercise does not change with conditioning in prepubertal boys nor are there conditioning-related changes in total peripheral resistance (Bar-Or 1983).

Table 6: Cardiovascular Changes that Occur in Children, with Conditioning. Adapted from ACSM (2005).

<i>Variable</i>	<i>Change</i>
<i>Morphologic</i>	
Heart Volume	Increase
Blood Volume	Increase
Total haemoglobin	Slight Increase
<i>Functional</i>	
Stroke Volume – submax, max	Increase
Heart Rate – submax	Decrease
Heart Rate – max	No change or Decrease
Cardiac Output – submax	No change or Decrease
Cardiac Output – max	Increase
Myocardial O ₂ requirements	Decrease
Arterio-mixed-venous O ₂ difference – submax, max	No change
Muscle blood flow – submax, max	No change
Systolic Blood Pressure – submax	No change
Systolic Blood Pressure –max	Increase
Diastolic Blood Pressure – submax, max	No change or Increase
Total peripheral resistance – submax, max	No change

CONDITIONING AND PULMONARY SYSTEM

Pulmonary functions, both static (volumes, capacities) and dynamic (flow, diffusion), are strongly related to body size. It is therefore hard to distinguish between growth and conditioning-related changes in the pulmonary response to exercise (Bar-Or 1983).

With conditioning, V_E and F_R during a standard task are reduced and O₂ extraction from the inspired air is greater. This important effect of conditioning may reflect a lower reliance on anaerobic metabolic pathways, a lower O₂ cost of exercise, or a reduction in chemoreceptor sensitivity. In contrast, the maximal V_E and F_R become higher with conditioning in proportion to the increase in maximal O₂ uptake (Bar-Or 1983) (see Table 7).

Table 7: Pulmonary Changes that Occur in Children, with Conditioning. Adapted from O. Bar-Or (1983).

<i>Function</i>	<i>Change</i>
Vital Capacity	No change
Ventilation - submax	Decrease
Ventilation - max	Increase
Respiratory Frequency - submax	Decrease
Tidal Volume - max	Increase
Ventilatory Equivalent – submax, max	Decrease or No change
Respiratory Muscle Endurance	Increase
Pulmonary Diffusing Capacity	No change

HEALTH-RELATED ASPECTS

There is now strong evidence of widespread and quite rapid secular changes in children's aerobic test performance. This is true in most age-sex groups, with a sample-weighted mean decline of 0.43% of mean values per year. This decline is most marked in older age groups and the rate of decline is similar for boys and girls (Tomkinson, Leger et al. 2003). Declines in fitness performance in this population have not been entirely due to increases in fatness and it is possible that climate factors or the cultural involvement can both play a relevant role in this decrease (Olds, Tomkinson et al. 2006).

Cardiorespiratory fitness is strongly associated with the clustering of cardiovascular disease (CVD) risk factors in children and youth, and that association is independent of country, age and sex (Anderssen, Cooper et al. 2007). It is known that higher

levels of physical activity and fitness in adults may protect against acute cardiac events by a variety of mechanisms including a reduced myocardial vulnerability to ventricular arrhythmias, reduced platelet stickiness and aggregability, and increased fibrinolysis (Boreham, Twisk et al. 2001). Aerobic fitness appears to be important in reducing the risk of cerebral and cardiovascular morbidity and mortality (Blair, Kohl et al. 1989), even in groups where the risk of heart disease and stroke is high (Ardern, Katzmarzyk et al. 2005). It was also related to better cognitive function in preadolescent children using a task that required variable amounts of executive control (Buck, Hillman et al. 2008).

Maximal oxygen uptake (VO_{2max}) is accepted as the criterion measure of cardiorespiratory fitness (CRF) (ACSM 2005). Boreham et al (2001) results indicated that the observed relationship between CRF and CHD risk status in adolescents are mediated by fatness whereas the observed relationships with fatness are independent of fitness.

It may be possible to express a lower limit of aerobic power that even when other health-related factors are absent may represent a risk. In girls this has been stated as $30 \text{ ml.kg}^{-1}.\text{min}^{-1}$ while for boys this is claimed to be $35 \text{ ml.kg}^{-1}.\text{min}^{-1}$ (Cooley and McNaughton 1999). For instance, Astrand et al (2003) stated that a VO_{2max} of $44 \text{ ml.kg}^{-1}.\text{min}^{-1}$ for young male adults (18-29 years of age) has been generally accepted as criterion below which both health and fitness may be compromised.

For this amount of reasons, a comprehensive aerobic fitness survey could also result in the construction of a set of world fitness norms, which might be used in the same way as the world standard body mass index (BMI) values for children that have been widely adopted, based on large international data sets (Cole, Bellizzi et al. 2000).

VO_{2MAX} – DEFINITION, CRITERION AND EXPRESSION FORMATS

DEFINITION

Maximal oxygen uptake (VO_{2max}) is accepted as the criterion measure of CRF (Massicotte, Gauthier et al. 1985). It limits the capacity to perform aerobic exercise and it is widely recognized as the best single measure of aerobic fitness (Armstrong and Welsman 2000).

Maximal oxygen uptake is the product of the maximal cardiac output (Q) (L blood. min^{-1}) and arterial-venous oxygen difference ($a-vO_2$) ($\text{ml } O_2 \text{ per l blood}$) (ACSM 2005). It is defined as the highest oxygen uptake the individual can attain during exercise engaging large muscle groups while breathing air at sea level, exercise time 2 to 6 min, depending on the type of exercise (Bassett and Howley 2000; Astrand, Stromme et al. 2003).

CRITERION

There is an almost unanimous consensus in respect to the magnitude of the maximum ability of the individual to take up, transport, and utilize oxygen by the working muscles, as a good indicator of the well functioning of the respiratory, cardiovascular and muscle-skeletal systems (Astrand, Stromme et al. 2003). Any disorder in one of the elements of these "aerobic team" will certainly reflect in the final output.

The traditional laboratory exercise model assumes that VO_2 rises with increasing exercise intensity up to a point beyond which no additional increase in VO_2 takes place, despite a well-motivated subject being able to increase further the intensity of his/her exercise. Exercise above the intensity where VO_2 levels off or plateaus is assumed to be supported exclusively by anaerobic resynthesis of adenosine triphosphate (ATP) resulting in an intracellular accumulation of lactate, acidosis, and inevitably termination of exercise (Armstrong and Welsman 2000).

Maximal VO_2 at the end of an exhaustive treadmill or leg cycle ergometer test must reach a plateau in order to reach a true VO_{2max} . If this criterion is not met, the appropriate term to describe CVF is VO_{2peak} rather than VO_{2max} . It is now generally accepted in CVF test that a progressively increasing work rate to a point of fatigue produces a VO_{2peak} that closely approximates to a VO_{2max} , even when a plateau in VO_2 is not evident (Bar-Or 2000; Pitetti, Fernhall et al. 2002).

Prediction of VO_{2max} from the submaximal HR response has been repeatedly criticized for its lack of accuracy, particularly on an individual basis (Davies 1968; Leger and Lambert 1982). It is limited by one or more assumptions: 1) linearity between

HR/VO₂; 2) accuracy of the prediction of HR_{max}; 3) O₂ constant cost of external work; and, at last, 4) diurnal variation of HR (McArdle, 1986) cited by Buono, Roby et al (1991).

T. W. Rowland (1993) and T. W. Rowland and Cunningham (1992) indicate that a HR_{peak} of 200 bpm and a RER of 0.99 would serve as effective criteria for achievement of VO_{2max} during treadmill running testing on children. Barnett et al (1993) goes further in the achievement of a VO_{2max} by means of gas analyser where is accepted the highest VO₂ averaged over 4 consecutive 15-s period as VO_{2peak}, providing the subject at least two of the following: 1) voluntary exhaustion; 2) a HR greater than 90% of the HR_{max} predicted (220-age); 3) a RER greater than 1.0; or 4) subjective opinion of experienced testers.

The problem of a true VO₂ measurement with children seems not only to rely on the type of ergometer used, but also on the amount of time that is required to reach a maximal effort (Massicotte, Gauthier et al. 1985). In addition, differences in physical activity are influenced by the beliefs/perceptions that children hold about their physical abilities, which likely also affect their motivation to persist in tests of endurance such as shuttle run (Cairney, Hay et al. 2008). In a study of Cairney et al (2008), children with high-perceived adequacy and low BMI scores completed the highest number of MSRT stages. The children who performed the poorest had both low perceived adequacy and high BMI. This illustrates the importance of considering psychological factors when interpreting physiologic assessments in children.

Data are consistent and demonstrate that over the range of 8 to 18 years old, boys' mass related VO_{2peak} remains unchanged at approximately 48-50 ml.kg.min⁻¹ whereas girls experience a progressive decline from about 45-35 ml.kg.min⁻¹ (Armstrong and Welsman 2000; Mota, Guerra et al. 2002).

Table 8: Maximal Oxygen Uptake Values (VO_{2max}; ml.kg⁻¹.min⁻¹) Found in Previous Researches Involving CVF Testing in Children and Youth.

Author	Participants	VO ₂ test criterion	VO _{2max}
Eriksson et al (1971)	8M 13-14 years	Cycle Ergometer	49
G. S. Krahenbuhl and Pangrazi (1983)	10M&F + 11 M&F 10.0 ± 0.4 e 9.9 ± 0.6	Treadmill	55.1 ± 3.6 and 48.6 ± 4.5
Massicotte et al (1985)	159M + 162F 10-17 years	Cycle Ergometer	M: 52.7-57.4 F: 41.9-48.3
Mahon and Vaccaro (1989)	16M 10-14 years	Treadmill	E: 45.9 ± 4.5 C: 45.4 ± 5.5
Armstrong (1991)	199M + 164F 11-16 years	Treadmill	M: 49 ± 6.0 F: 41 ± 6.0
Buono et al (1991)	45M + 45F 5 th 10.6 ± 0.5 8 th 12.8 ± 0.6 11 th 16.1 ± 1.1	Treadmill	5 th 47.8 ± 7.1 8 th 47.6 ± 8.2 11 th 48.0 ± 8.1
Cooke et al (1991)	16M 10-12 years	Treadmill	64.3 ± 3.2 (trained)
T. W. Rowland and Cunningham (1992)	15M&F 7-10 years	Treadmill	No plateau: 51.9 ± 4.6 Plateau: 47.5 ± 5.7
Gratas-Delamarche et al (1993)	9M 9.8-10.8 years	Cycle Ergometer	49.7 ± 1.7
T. W. Rowland (1993)	6M + 3F 10-13 years	Treadmill	53.9 ± 4.3
Cureton et al (1995)	376M + 183F 8-17	Treadmill	8-11 yrs Male: 49.6 Female: 45.4
Rogers et al (1995)	Prepuberty 15M + 15F 6.6-9.5 years Puberty 15M 12.6-14.2 years and 15F 11.1-12.9 years.	Treadmill	Prepuberty (6.6-9.5) M: 50.7 ± 9.6 F: 49.1 ± 6.5 Puberty M: 53.4 ± 5.9 F: 45.4 ± 7.3

ABSOLUTE AND RELATIVE EXPRESSION OF VO_{2MAX}

For tasks that involve moving the whole body from one place to another, the child whose body mass is smaller may not need as high an absolute VO_{2max} as that of the heavier adolescent or adult. To compare the maximal aerobic power of individuals who differ in body mass, one should therefore express VO_{2max} in relative rather than absolute terms, e.g., per body height, surface area, lean body mass, or body weight (Bar-Or 1983).

Compelling arguments have been presented to question the validity of using simple ratio scaling to remove the influence of body mass from size-dependent performance measures such as VO_{2peak} . Studies in which appropriate allometric techniques have been used to control for body size have provided new insights into VO_{2peak} during growth. A different picture emerged from the allometric analysis, with VO_{2peak} significantly increasing into puberty with no decline into adulthood evident. These data challenge the conventional interpretation of VO_{2peak} during growth in both boys and girls. They demonstrate that there is a progressive increase in aerobic fitness in both sexes over the age range 11 to 13 years independent of the influence of body size. Peak VO_2 remains lower in girls than boys, with sex differences increasing over the age range observed (Armstrong and Welsman 2000). Maximal VO_2 relative to body surface area ($ml \cdot m^{-2} \cdot min^{-1}$ or $ml \cdot Kg^{-0.67} \cdot min^{-1}$) or body mass to the 0.75 power ($ml \cdot Kg^{-0.75} \cdot min^{-1}$) is a more appropriate method of data expression than a simple ratio standard of VO_2 relative to body mass ($ml \cdot kg^{-1} \cdot min^{-1}$) (Rogers, Olson et al. 1995).

Although this practice has been criticized (Mota, Guerra et al. 2002), VO_{2max} was expressed per unit of body mass ($ml \cdot kg^{-1} \cdot min^{-1}$). The primary concern is that if the physiological variable of interest and the body size dimension do not increase in a constant linear proportion to one another, the data will be inappropriately skewed. The problem with scaling VO_2 relative to body mass is that these two variables do not meet this criterion. In fact, the VO_2 -to-body mass relationship has been repeatedly demonstrated to be a curvilinear exponential function (Bergh, Sjodin et al. 1991). Therefore, the use of a simple ratio standard can lead to misleading conclusions regarding differences or similarities between individuals of varying sizes. Because of the large body size difference between children and adults, the potential magnitude of error from an inappropriate scaling factor is even greater when making comparisons between these populations (Rogers, Olson et al. 1995).

Despite the state of knowledge, the great majority of the studies revised had the VO_{2max} values expressed in $ml \cdot Kg^{-1} \cdot min^{-1}$ as they do not find any practical reason for the use of others measuring units in alternative to the "traditional" (Krahenbuhl and Williams 1992). By following these authors' steps, we were able to properly compare the measured directly VO_{2peak} results with the estimated FITNESSGRAM's software reports, as well as with numerous others branded authors, promoting the discussion.

EXERCISE TESTING TO OBTAIN A VO_{2MAX}

Exercise testing of children differs from adult exercise testing in many ways beyond the technical issues related to test performance. Disease processes that produce myocardial ischemia are relatively rare in children compared with adults. Exercise testing may be useful in these cases, but the use of testing to assess functional capacity or cardiac rhythms will be encountered more often. Although the precise role of exercise testing will vary somewhat from place to place and from case to case, exercise testing is often essential to diagnose and to direct treatment in a wide variety of situations. The pediatric professionals should be available to discuss when a test might be of value in a specific case in addition to providing advice about the specifics of the performance of the test and offering age- and size-appropriate normal data from the laboratory with test interpretation.

Here are stated some of the most common reasons to pediatric exercise testing (Washington, Bricker et al. 1994; Paridon, Alpert et al. 2006):

- Evaluate specific signs or symptoms that are induced or aggravated by exercise;
- Assess or identify abnormal responses to exercise in children with cardiac, pulmonary, or other organ disorders, including the presence of myocardial ischemia and arrhythmias;
- Assess efficacy of specific medical or surgical treatments;

- Evaluate prognosis, including both baseline and serial testing measurements;
- Establish baseline data for institution of cardiac, pulmonary, or musculoskeletal rehabilitation;
- To estimate levels of functional capacity for and to improve the safety of vocational, recreational, and athletic recommendations;
- To evaluate fitness levels.

LABORATORY VERSUS FIELD EXERCISE TESTING

The standard test for assessing aerobic endurance fitness is the direct measurement of the subject's VO_{2max} whilst running to exhaustion on a treadmill in the laboratory environment (Aziz, Tan et al. 2005) (see Table 8).

Although it is a useful criterion of the overall capacity of the individual to perform exercise aerobically, the direct determination of VO_{2max} requires sophisticated instrumentation, laboratory time and trained personnel. Consequently, such measurements are not practicable for a large group of individuals. Therefore, attempts have been made to develop simple non-invasive, valid and reliable field tests to provide an estimate of VO_{2max} requiring limited equipment (Mahoney 1992). Such tests require the subject either to cover as much distance as possible in a set time (Cooper 1970), cover a set distance in the fastest time possible (Kline, Porcari et al. 1987) or perform a 20-m multistage shuttle run test (Leger and Lambert 1982). These field tests are maximal from the onset and require well motivated subjects with some knowledge of pace judgment and understanding of the test requirements (van Mechelen, Hlobil et al. 1986; Ramsbottom, Brewer et al. 1988; Mahoney 1992).

From a metabolic point of view, previous studies suggest increased energy demands during field tests, in specific the 20-m MSRT, compared to treadmill running (Flouris, Metsios et al. 2005) (see Table 9). This may well be attributed to differences in factors such as intensity, exercise mode, technique, and musculature employed between the two conditions. These factors should be considered in the design of physical training programs that incorporate shuttle running elements. This information should also be taken into account when designing the physical training for sports incorporating shuttle running (for example, football, basketball, rugby).

Table 9: Maximal Oxygen Uptake values (VO_{2max} ; $ml.kg^{-1}.min^{-1}$) Across Different Standard Testing Demands in Children. Adapted from Washington et al (1994).

Protocol	VO_{2max}
Stepwise loading, cycle ergometer	35.6-60.6
Supramaximal, cycle ergometer	49-55.4
Discontinuous, cycle ergometer	41.8-56.6
Other, cycle ergometer	32.6-61.4
Increasing-grade stepwise loading, treadmill	47.7-61.0
Increasing-speed stepwise loading, treadmill	45.7-58.2
Increasing-grade and increasing-speed stepwise loading, treadmill	45.9-61.3
Walking, treadmill	43.1-53.5

20-M MULTISTAGE SHUTTLE RUN TEST (20-M MSRT)

Unfavourable comparisons between the fitness of American children relative to Europeans in the 1950s and 1960s inspired efforts to introduce standardized testing and physical education in US schools (Olds, Tomkinson et al. 2006). As test batteries have been developed and revised, test items considered appropriate and valid for students in Grades 5-12 have been recommended for younger children without additional supportive research. The FITNESSGRAM (The Cooper Institute for Aerobics Research 1994) national test batteries are designed to be administered to children beginning in kindergarten (Rikli, Petray et al. 1992). It evaluates the health-related components of fitness (aerobic fitness, body composition, muscular strength and endurance, and flexibility) identified as the most important to assess. The FITNESSGRAM includes two test items designed to test aerobic fitness: the 1 Mile Run/Walk (MRW) and the 20-m Multistage Shuttle Run Test (20-m MSRT) (Vincent, Barker et al. 1999).

The 20-m MSRT was developed by Leger and Lambert (1982) to simulate a graded exercise test with approximately 1 MET level increase between stages. Rather than running around an oval track, subjects run back and forth between parallel lines placed 20-m apart. Because no grade can be involved as on a treadmill, the energy demand is elevated by increasing the speed. This was a development of their earlier track running test (Leger and Boucher 1980), which was re-designed to allow the test to be performed inside gymnasiums. The original test, using 2-min incremental speed stages, was later modified to 1-min stages to reduce the overall test time and aid motivation of the participants (Leger and Gadoury 1989). This is so since it is well known that the utilization of running tests with young population encompasses certain problems such as the time or distance required motivation and anaerobic contribution (Massicotte, Gauthier et al. 1985).

The 20-m MSRT represents an acceptable field assessment tool for CRF and has been repeatedly employed in different health (Bouziotas, Koutedakis et al. 2001) and fitness (Grant, Corbett et al. 1995) settings. However, the popularity of the 20-m MSRT is mainly attributed to its practical use for simultaneous measurement of large groups of individuals. Studies evaluating its accuracy in predicting laboratory VO_{2max} have reported contradictory results (Grant, Corbett et al. 1995; Boreham, Twisk et al. 2001). More importantly, the efficacy (that is, the extent to which a specific procedure produces a valid classification of data in relation to established criteria) of the original 20-m MSRT model in screening for CRF remains unknown (Flouris, Metsios et al. 2005). The test is incremental and hence submaximal for much of its duration, requiring maximal effort only the last minute or so. For this reason it does not require maximal motivation throughout. This should reduce, if not eliminate, the influence that average $\%VO_{2max}$ used during the test has on the results (Liu, Plowman et al. 1992). During the MSRT, and especially in the latter stages, the HR response is curvilinear suggesting that the 1min interval between shuttle levels, do not allow steady state conditions to be attained (Ramsbottom, Brewer et al. 1988).

The 20-m MSRT is probably the most common test to assess the aerobic fitness of children and adolescents. It seems that the MSRT is a better predictor of VO_{2max} than endurance runs, because it is an indoor test (no weather or surface differences) and it has a protocol with increasing speed which is much in accordance with the loading pattern of a VO_{2max} (van Mechelen, Hlobil et al. 1986). Indeed, as for continuous running, the repetition of short bouts of exercise at high percentage of maximal aerobic speed (MAS), alternating with identical passive recovery time, allows subjects to reach high levels of VO_2 and even elicit VO_{2peak} (Baquet, Berthoin et al. 2002). According to van Mechelen et al (1986), it is not necessary to run more than five minutes at a steady pace to significantly improve the correlation between VO_{2max} and running performance. Findings from other investigations suggest that the VO_{2max} during the 20-m MSRT is significantly higher compared to a treadmill test. This may well be attributed to differences in factors such as intensity, exercise mode, technique, and musculature employed between the two conditions (Flouris, Metsios et al. 2005).

The test has a number of variants (see Table 10), and small differences in the conduct of the test and the analysis of results make it difficult to compare data from different sources. Tomkinson et al (2003) analyzed secular trends in 20-m MSRT performance. Their meta-analysis highlighted the difficulties arising from the different metrics used for expressing 20-m MSRT results, the existence of variant forms of the test, artefacts created by averaging over a range of age groups, and the exact time at which stage numbers were called on the tapes used. Some studies called elapsed time on the minute, while others called stages at the end of the stages (which might differ by a second or two from the exact minute). Individual results have been reported as the number of completed stages, the running speed at the last completed stage, the number of completed laps (or stages plus laps), the number of minutes the test lasted, or as an estimated VO_{2max} based on regression equations. Hence, stages are not always precisely 1-min long (Tomkinson, Leger et al. 2003).

Table 10: Running Speed (Km.h⁻¹) and Total Number of Laps for Each Stage of the 20-m MSRT According to Four Major Protocol Variants. Based on Tomkinson et al (2003).

Stage	Leger and Lambert, (1988)		Eurofit (Council of Europe 1988)		QUB (Queens University of Belfast)		FITNESSGRAM (The Cooper Institute for Aerobics Research 2002)	
	Speed	Laps	Speed	Laps	Speed	Laps	Speed	Laps
1	8.5	7	8.0	7	8.0	7	8.0	7
2	9.0	8	9.0	8	8.5	7	9.0	8
3	9.5	8	9.5	8	9.0	8	9.5	8
4	10.0	8	10.0	8	9.5	8	10.0	9
5	10.5	9	10.5	9	10.0	8	10.5	9
6	11.0	9	11.0	9	10.5	9	11.0	10
7	11.5	10	11.5	10	11.0	9	11.5	10

There are consistent and systematically evidences that boys easily outperformed girls, the differences being significant ($p \leq 0.0001$) (Olds, Tomkinson et al. 2006). There was a significant negative relationship ($r = -0.44$, $p = 0.007$) between the performance index and the average annual temperature of the capital city of the country, so that children from colder countries performed better. However, when these countries are excluded from the analysis, there is no longer a significant relationship between temperature and performance ($r = 0.018$, $p = 0.34$) (Olds, Tomkinson et al. 2006).

The best performing nations were Northern European countries, notably Iceland and Baltic states. The top seven nations were all from Northern and Central Europe. The worst performing nations were Southern European countries (Greece, Italy, Portugal, and Spain), Brazil, and some of the developed Rim nations (USA, Singapore, Hong Kong, and Australia). In Europe, there was a clear north-south fitness gradient (Olds, Tomkinson et al. 2006).

20-M MSRT ADVANTAGES

The 20-m MSRT has some advantages such as the constant environmental conditions, a graded physiological response and the absence of individual pace control (Paliczka, Nichols et al. 1987; Ramsbottom, Brewer et al. 1988). It can be administered in a relatively small space either indoors or outdoors and is therefore practical for schools without extensive facilities. In these settings, the test is often conducted in small groups, making it feel less like a test and more like a running game (Cairney, Hay et al. 2008). Students run together, so no student is considerably ahead of or behind the others, as is often the case during the MRW. Students continue to run as long as possible, so the last student finished does the best, as opposed to the MRW, in which the last student finished is the slowest (Vincent, Barker et al. 1999). For those whose sports involve frequent stops, starts, and turns, it may be a more specific test than a continuous directional run (Liu, Plowman et al. 1992). The scoring method is easy to administer and interpret, and has realistic endpoints. Motivation is partially controlled by the predetermined workloads and body weight is an integral part of the test (Mahoney 1992).

The 20-m MSRT is a maximal running test with increasing workload and therefore resembles laboratory exercise test protocols used to assess VO_{2max} . It is an indoor test requiring only a standard school gym (12x21 m) and is not influenced by differences in running surface or weather conditions (van Mechelen, Hlobil et al. 1986).

20-M MSRT DISADVANTAGES

Since the origin of the 20-m MSRT, the focus of research attention has been on its ability to estimate aerobic fitness. There are problems inherent with this process due to sample selection, the nature of the statistical analysis (error of prediction), and the issue of habituation or practice (Lamb, 2007). In recent years, attention has been afforded to addressing the utility of the 20-m MSRT (and another running-based tests) as markers of sport-specific endurance performance (Lemmink, Verheijen et al. 2004; Aziz, Tan et al. 2005), rather than VO_{2max} (Lamb and Rogers 2007).

Lamb and Rogers (2007) in their study affirmed that the 95% Limits of Agreement (LA) utilized indicated that the degree of random error was too large for the test to be considered to have sufficient reliability to monitor changes in aerobic capacity in active male and female undergraduates. The data also indicated that a small learning or habituation effect occurred between the first and second applications of the test, suggesting that, notwithstanding, researchers and practitioners should

me minded to provide at least one practice trial before utilizing the scores from the 20-m MSRT. They further stated that unless researchers explore ways in which improvements can be made, there is a case for the test to be abandoned as a credible predictor of aerobic capacity or exercise performance.

The 20-m MSRT frequently underestimates VO_{2max} (Sproule, Kunalan et al. 1993; Berthoin, Gerbeaux et al. 1994; Grant, Corbett et al. 1995), in particularly when VO_{2max} is beyond the healthy untrained subjects as in the original (Leger and Lambert 1982) study.

VALIDATION STUDIES OF THE 20-M MSRT AS AN AEROBIC FITNESS TEST IN CHILDREN AND YOUTH

In the following section we briefly summarize all the revised validation studies of the 20-m MSRT in children and youth (see Table 11). Each study has its own purpose and method, elevating the importance of knowing what has or has not been done across the years on this subject.

Table 11: Validation Studies of the 20-m MSRT in Children and Youth.

Author	Participants	VO ₂ test criterion	Performance variable	<i>r</i>	<i>R</i> ²	VO _{2max}
van Mechelen et al (1986)	41M, 41F 12-14 years	Treadmill 8 Km. h ⁻¹ 2.5% ↑/2min	Total stages	0.68M 0.69F 0.76 M&F		Treadmill M: 53.2 ± 5.4 F: 44.1 ± 4.8
Armstrong et al (1988)	77M 11-14 years	Treadmill 10 Km. h ⁻¹ 2% ↑/3min	½ Stage	0.54		###
Leger et al (1988)	200M&F 8-19 years	Retroextrapolation at the end of MSRT + Cycle Ergometer	Max Speed		0.71	8-11 years M&F: 40.5 ± 7.1 41.3 ± 9.0 13 years M: 53.9 ± 5.0; 52.9 ± 2.9
Boreham et al (1990)	23M, 18F 14-16 years	Treadmill 6, 8, 10 Km. h ⁻¹ and then 2% ↑/1min	Total laps	0.64M 0.90F 0.87 M&F		###
Liu et al (1992)	22M, 26F 12-15 years	Treadmill progressive	Total laps	0.65M 0.51F 0.69M&F 0.72M&F		M: 54.34 ± 6.95 F: 46.29 ± 6.04 T: 49.97 ± 7.59
Mahoney (1992)	22M, 26F 53M, 50F 12 years	Treadmill M: 8; 9,4 Km. h ⁻¹ 5% ↑/2min F: 4.8 Km. h ⁻¹ 5% ↑/2min	Total laps	0.83M 0.76F		M: 43.8 ± 4.5 F: 38.5 ± 4.4
Barnett et al (1993)	27M + 28F 12-17 years	Treadmill 8, 9, 10 Km. h ⁻¹ 0% ↑/2min, then 2.5-3% ↑/2min, until exhaustion.	Max Speed	0.72 0.82		M: 53.9 ± 5.6 F: 45.1 ± 4.6
Falgairrette et al (1994)	16M + 28M 10-11 years	Cycle Ergometer	Max Speed	0.93		Low alt.: 51.0 ± 7 High alt.: 41.0 ± 4
McVeigh et al (1995)	15M + 18F M: 13.68 ± 0.3 F: 13.68 ± 0.3	Treadmill 7-8 Km. h ⁻¹ 0% /2min, then 9-10 Km. h ⁻¹ 2% /2min, until exhaustion.	Max Speed	0.60M 0.65F	0.68M 0.85F	M: 53.17 ± 5.3 F: 42.94 ± 5.8
Oliveira (1998)	77M + 80F 15.0 ± 0.8	Portable gas analyser Cortex Metamax, CBS	Total laps		0.78	M: 53.1 ± 6.8 F: 40.6 ± 4.5

Pitetti et al (2002)	13M + 38F 8-15 years	Treadmill Modified Bruce Protocol	Total laps	0.57		M: 49.4 ± 7.7 F: 45.0 ± 7.2
Matsuzaka et al (2004)	62M + 70F 8-17	Treadmill	Total laps	0.77	0.80	8-10M: 52.1 ± 6.2 11-13M: 56.0 ± 5.0 8-10F: 49.3 ± 5.5 11-13F: 50.2 ± 5.4

VO_{2max} in ml.kg⁻¹.min⁻¹.

In order to validate the 20-m MSRT as an estimate of VO_{2max} and to compare the results of this test with the results of a 6 min endurance run, van Mechelen et al (1986) performed the 20-MSRT and the 6 min endurance run in 82 subjects (41 boys and 41 girls) aged 12-14, and had their VO_{2max} directly measured during maximal treadmill running. The test result was expressed as "palier" (one palier is approximately one minute). The mean results of the 20-MST were, for boys, 8.0 palier (± 1.7) and for girls, 6.4 palier (± 1.5). The mean results of the 6 min endurance run were for boys, 1264.4 meters (± 160.8), and for girls, 1103.9 meters (± 144.7). The mean VO_{2max} for boys was 53.2 ± 5.4 ml.kg⁻¹.min⁻¹ and for girls, 44.1 ± 4.8 ml.kg⁻¹.min⁻¹. The correlation coefficient between VO_{2max} and the 20-m MSRT was found to be 0.68 (± 3.9) for boys, 0.69 (± 3.4) for girls and 0.76 (± 4.4) for both sexes, and that of VO_{2max} with the 6 min endurance run was 0.51 (± 4.6) for boys, 0.45 (± 4.3) for girls and 0.63 (± 5.3) for both sexes. The conclusion was that the 20-m MSRT is a suitable tool for the evaluation of maximal aerobic power. The authors further declared that the differences in validity between the 20-m MSRT and the 6 minutes endurance run were statistically not significant ($p > 0.05$), and for reasons of practicability the 20-m MSRT should be preferred to the 6 minutes endurance run when used in physical education classes.

Leger et al (1988) determined the maximal aerobic power of school children, healthy adults attending fitness class and athletes performing in sports with frequent stops and starts (e.g. basketball, fencing and so on). The last stage number announced was used to predict VO_{2max} (ml.kg⁻¹.min⁻¹) from the speed (km.h⁻¹) corresponding to that stage (speed = 8 + 0.5 stage n°) and age (year): $VO_{2max} = 31.025 + 3.238(\text{Speed}) - 3.248(\text{Age}) + 0.1536(\text{Age})(\text{Speed})$, ($r=0.71$) with 188 boys and girls aged 8-19 years. Right upon termination VO₂ was measured with four 20 sec samples and VO_{2max} was estimated by retroextrapolating the O₂ recovery curve at time zero of recovery. For adults, similar measurements indicated that the same equation could be used keeping age constant at 18 ($r=0.90$, $n=77$ men and women 18-50 years old). Test-retest reliability coefficients were 0.89 for children (139 boys and girls 6-16 years old) and 0.95 for adults (81 men and women, 20-45 years old)

Boreham et al (1990) compared the validity of two field tests of aerobic fitness for predicting VO_{2max} in a group of adolescent schoolchildren. Twenty four school boys (mean age=15.6 ± 0.6 yr) and twenty four schoolgirls (mean age = 15.4 ± 0.7 yr) underwent three different tests of aerobic capacity. The tests were: (a) VO_{2max}, using a continuous, progressive treadmill protocol, (b) a physical work capacity test at a heart rate of 170 bpm (PWC170), and (c) a 20-m MSRT. Results (means ± SD) for boys ($n=23$) and girls ($n=18$) respectively were: VO_{2max}, 53.3 ± 5.9 and 42.6 ± 5.8 ml.kg⁻¹.min⁻¹; PWC170, 2.84 ± 0.47 and 1.86 ± 0.39 w.kg⁻¹; 20-MST, 81.7 ± 15.9 and 50.4 ± 12.5 laps. Linear regression of PWC170 and 20-m MSRT on VO_{2max} scores ($n=41$), revealed similarly high powers of prediction for both field tests (PWC170 vs. VO_{2max}, $r=0.84$; 20-m MSRT vs. VO_{2max}, $r=0.87$), with an indication that the 20-m MSRT may be preferable for use with girls.

Liu et al (1992) had a threefold purpose: to determine (a) the test-retest reliability of the 20-m MSRT (number of laps), (b) the concurrent validity of the 20-m MSRT (number of laps), and (c) the validity of the prediction equation for VO_{2max} developed by Leger et al (1988) on Canadian children for use with American children 12-15 years old. An intraclass coefficient of 0.93 was obtained on 20 students (12 males; $r=0.91$ and 8 females; $r=0.87$) who completed the test twice, 1 week apart. The number of laps run correlated significantly with VO_{2peak} in males ($n=22$; $r=0.65$; $F[1, 20]=14.30$ $p \leq 0.001$), females ($n=26$; $r=0.51$; $F[1, 24]=8.34$; $p \leq 0.01$), and males and females ($r=0.69$; $F[1, 46]=42.54$, $p \leq 0.001$). When the measured VO_{2peak} (49.97 ± 7.59 ml.kg⁻¹.min⁻¹) was compared with the estimated VO_{2max} (48.72 ± 5.72 ml.kg⁻¹.min⁻¹) predicted from age and maximal speed of the 20-m MSRT (Leger, Mercier et al. 1988) no significant difference was found, $t(47)=-1.631$; $p \geq 0.11$, between the means; the r was 0.72 and SEE was 5.26 ml.kg⁻¹.min⁻¹.

Mahoney (1992) investigated the validity of 20-m MSRT and PWC170 (Physical Working Capacity) field tests with laboratory-measured VO_{2peak} in a multiracial UK population of 12 year olds: 103 subjects completed the 20-m MSRT and 96 of these completed the PWC170. Peak VO_2 was $43.8 \text{ ml.kg}^{-1}.\text{min}^{-1}$ for boys and $38.5 \text{ ml.kg}^{-1}.\text{min}^{-1}$ for girls. Pearson product-moment correlation showed 20-m MSRT to be a reliable measure of cardiorespiratory fitness ($r=0.83$, boys; $r=0.76$, girls, $p<0.03$), while correlations with PWC170 were lower ($r=0.64$, boys; $r=0.54$, girls) and not significant. The 20-m MSRT was consistent in retest ($n=20$) - reliability coefficients $r=0.73$, boys; $r=0.88$, girls; $p<0.01$. The results suggest that the 20-m MSRT is a valid measure of fitness in this population when compared with VO_{2peak} . PWC170 is less valid, possibly due to cultural and social backgrounds.

Barnett et al (1993) determined the validity of the 20-m MSRT for predicting VO_{2peak} in Hong Kong Chinese students, ages 12-17 years (55 subjects, 27 boys and 28 girls). A correlation of 0.72 ($p<0.001$) was found between VO_{2peak} and predicted VO_{2max} using an equation previously developed with Canadian children (Leger, Mercier et al. 1988). However, maximal shuttle run speed alone was a better predictor in this group ($r=0.74$, $SEE=4.6 \text{ ml.kg}^{-1}.\text{min}^{-1}$, $p<0.001$). Multiple-regression analysis (best-subsets) was performed and the best predictor variables were maximal speed and sex with either triceps skinfold or weight. For practical application in the school setting, the equation $peak \text{ } VO_2 = 24.2 - 5.0(\text{sex}) - 0.8(\text{age}) + 3.4(\text{maximal speed})$ ($r=0.82$, $SEE=4.0$), where for sex, male = 0 and female = 1, is suggested.

Falgairrette et al (1994) performed field tests of running and laboratory tests in La Paz [high altitude (HA), 3700 m] and in Clermont-Ferrand [low altitude (LA), 300 m], in order to investigate their validity at HA. Prepubertal boys of mean ages 10.6 years (HA1, $n=16$; LA1, $n=28$) and pubertal boys of 13.7 years (HA2, $n=12$; LA2, $n=41$) took part in the study. All the boys performed a 30-m sprint (V30m), a 30-s shuttle run (V30s) and a 20-m MSRT. In the laboratory, the boys performed a force-velocity test (P_{max}), a Wingate test (PWing) and a graded test to measure maximal oxygen consumption (VO_{2maxB} ; direct method) on a cycle ergometer. At similar ages, the VO_{2max} , MAS in the 20-m MSRT and VO_{2maxB} were lower by 9%, 12% and 20%, respectively, at HA than at LA ($p<0.05$). However, the relationships between VO_{2maxB} and VO_{2max} 20-m MSRT (litres per minute) at HA ($r=0.88$) and at LA ($r=0.93$) were identical.

McVeigh et al (1995) examined the reliability and validity of the 20-m MSRT as a predictor of VO_{2peak} in Edinburgh school children. Thirty-three children (15 boys, 18 girls) performed three shuttle tests and three laboratory treadmill tests of VO_{2peak} . Multiple regression analysis revealed that the prediction of VO_{2peak} ($\text{ml.kg}^{-1}.\text{min}^{-1}$) from 20-m MSRT performance was improved by including skinfold thickness measurements in the prediction models, particularly with the female group. Predictive power was greatest for females when using maximal speed (km.h^{-1}) best of three repeated tests and triceps skinfold thickness (mm) ($R^2=0.85$, $SEE = 2.4$) [$VO_{2max} = 7.26(\text{MS};\text{km.h}^{-1}) - 0.56(\text{Tric}_{\text{skinfold}};\text{mm}) - 26.86$], and for males when MS best of three repeated tests and the sum of the triceps and subscapular skinfolds ($R^2=0.68$, $SEE=3.23$) [$VO_{2max} = 4.94(\text{MS};\text{km.h}^{-1}) - 0.30(\text{Tric}\&\text{Sub}_{\text{skinfold}};\text{mm}) - 1.16$]. When using 20-m MSRT performance to predict VO_{2peak} ($\text{ml.kg}^{-1}.\text{min}^{-1}$) on children of this age group, body composition measures must be included in the equation.

Oliveira (1998) assessed the validity of the Mercier et al (1983) 1 min stages version of the 20-m MSRT in 157 Portuguese students (15 ± 0.8 yr). Participants were randomly distributed into validation and a control group. The first performed the 20-m MSRT with a portable gas analyser and the last were used solely to the indirect validation of the test and the development of VO_{2max} estimation equations. The authors observed moderate correlations ($R^2=0.78$) when the VO_{2max} was expressed relatively to body weight ($\text{ml.kg}^{-1}.\text{min}^{-1}$). As follows: $VO_{2max} = 25.0 + 6.6(\text{sex}) + 0.23(\text{laps})$. However, the allometric representation of VO_{2max} ($\text{ml.kg}^{-0.67}.\text{min}^{-1}$) showed higher predictive power ($R^2=0.83$). Thus: $VO_{2max} = 83.3 + 35.1(\text{sex}) + 0.77(\text{laps})$. The 20-m MSRT was consistent in retest ($r=0.93$).

Fernhall et al (1998) developed two regression equations to predict CVF based on the 20-m MSRT in children with mild and moderate mental retardation (MR) which was posterior validated to nondisabled, healthy youths (ages 8 to 15 years; 38 females and 13 males) (Pitetti, Fernhall et al. 2002). Participants performed two modified Bruce protocol treadmill tests (TMT) and two 20-m MSRT on separate days. Cardiovascular fitness (VO_{2peak} , $\text{ml.kg}^{-1}.\text{min}^{-1}$) was measured during the TMT and computed for the 20-m MSRT using both regression equations. Results indicate that the test-retest correlations for the 20-m MSRT (n° of laps; $r=0.89$) and TMT (VO_{2peak} , $\text{ml.kg}^{-1}.\text{min}^{-1}$; $r=0.86$) were high. Predicted VO_{2peak} values were moderately significant (nondisabled youth: $r=0.55$, $p<0.01$; youth with MR: $r=0.66$, $p<0.01$) when compared with TMT VO_{2peak} .

Matsuzaka et al (2004) examined the validity of a 20-m MSRT as an aerobic fitness test for Japanese children, adolescents, and young adults. Participants were 62 boys and 70 girls aged 8–17 years and 56 men and 99 women aged 18–23 years. The authors observed high multiple correlations for adults ($R^2=0.88$) and for children and adolescents ($R^2=0.80$). Therefore, it was suggested that their multiple regression equations are more appropriate for predicting VO_{2peak} in Japanese children, adolescents, and adults.

VALIDATION STUDIES OF THE 20-M MSRT AS AN AEROBIC FITNESS TEST IN ADULTS

In order to truly understand the vast investigation fields on the 20-m MSRT, one have to look inevitably to what has been done in adults too. Table 12 summarizes the validation studies of the 20-m MSRT in adults.

Table 12: Validation Studies of the 20-m MSRT in Adults.

Author	Participants	VO ₂ test criterion	MSRT performance variable	Correlation Coefficients
Leger and Lambert (1982)	59M + 32 F M : 24.8 ± 5.5 F : 27.3 ± 9.2	Retroextrapolation at the end of MSRT	Max Speed	0.84
Paliczka et al (1987)	9M 26-47 years	Treadmill Mean speed of the 10 km run 2% ↑/1min	Total laps	0.93
Ramsbottom et al (1988)	36M, 38F 19-36 years	Treadmill Exhaustion speed 7-10 min 3.5% initial and then 2.5% ↑/3min	Total stages	0.82M 0.89F 0.92 M&F
Sproule et al (1993)	16M, 4F 21-35 years M 20-21 years F	Treadmill Taylor Protocol	Total stages	0.86 M&F
Grant et al (1995)	22M 19-29 years	Treadmill Based on the recommendations of the <i>British Association of Sports Science</i>	Total stages	0.86
St Clair Gibson et al (1998)	20M&F 22 ± 3	Treadmill 2min: 4.8 km.h ⁻¹ and then 2.5° ↑/2min till 20° and then 0.4 km.h ⁻¹ ↑/2min	Max Speed at the last completed stage	0.71 runners 0.61 squash players 0.67 combined
Wilkinson et al (1999)	25M, 2F 22.6 ± 3.4 years	Treadmill 4 min: Exhaustion speed 8-12 min 1% ↑/1min and then 0.1 km.h ⁻¹ ↑/5 sec	Mean speed	0.91
Flouris et al (2005)	110M 21.6 ± 2.5	Treadmill Bruce Protocol	Max Speed	0.96

Leger and Lambert (1982) pioneer study performed the primary 20-m MSRT in 91 adults (32 females and 59 males, aged 27.3 ± 9.2 and 24.8 ± 5.5 year respectively and with VO_{2max} (mean ± SD) of 39.3 ± 8.3 and 51.6 ± 7.8 ml. kg⁻¹. min⁻¹ respectively) and had VO_{2max} estimated by the retroextrapolation method (extrapolation to time zero of recovery, of the exponential least squares regression of the first four 20 sec recovery VO_2 values). Starting at 8 km.h⁻¹ and increasing by 0.5 km.h⁻¹ every 2 min, the 20-m MSRT enabled prediction of the VO_{2max} (ml.kg⁻¹.min⁻¹) from the maximal speed (MS, km.h⁻¹) by means of the following regression equation: $VO_{2max} = 5.857(MS) - 19.458$; $r=0.84$ and $SEE=5.4$. Later, the multistage protocol was slightly modified to its final version, in which the test started at stage 7 Met and continued with a 1 Met (3.5 ml O₂.kg⁻¹.min⁻¹) increment every 2 min. Twenty-five of the 91 subjects performed the 20-m MSRT twice, once on a hard, low-friction surface (vinyl-asbestos tiles) and another time on a rubber floor, as well as a walking maximal multistage test on an inclined treadmill. There was no difference between the means of these tests or between the slopes of the VO_{2max} - maximal speed regressions for the two types of surfaces. The 20-m MSRT and another maximal multistage field test involving continuous track running gave comparable results ($r=0.92$, $SEE=2.6$, $n=70$). Finally, test and retest of the 20-m MSRT also yielded

comparable results ($r=0.975$, $SEE=2.0$, $n=50$). The authors concluded that the 20-m MSRT is a valid and reliable test for the prediction of the VO_{2max} of male and female adults, individually or in groups, on most gymnasium surfaces.

Paliczka et al (1987) assessed the validity of a 20-m MSRT both as a field test of CR endurance and as a predictor of competitive performance in a 10 km race. Nine male subjects (age 35.4 ± 5.8 years) (mean \pm SD) underwent a laboratory test of maximum oxygen uptake on a treadmill (VO_{2max} 59.0 ± 9.9 ml.kg⁻¹.min⁻¹), completed the 20-m MSRT (score 105 ± 23.7 laps/11.4 \pm 2.7 paliers) and competed in a 10 km race (finishing time 41.8 ± 7.3 minutes). Analysis using Pearson's Product Moment Coefficient revealed high correlations between these variables (20-m MSRT vs. VO_{2max} , $r=0.93$; 20-m MSRT vs. 10 km, $r=-0.93$; VO_{2max} vs. 10 km, $r=-0.95$). These results confirm that the 20-m MSRT is a valid field test of CR endurance and suggest that it can additionally be used to predict relative running performance over 10 km.

Sproule et al (1993) compared the results of direct and indirect measurements of VO_{2max} of an Asian population. Twenty subjects (16 male, 4 female), all of whom were physical education students, were assessed directly using laboratory treadmill running to determine VO_{2max} . The indirect estimation of VO_{2max} was obtained using a modified form of Leger and Lambert (1982) 20-m MSRT. Pearson product-moment correlations confirmed test-retest reliability for both direct and indirect measurements ($r=0.90$ and $r=0.91$ respectively). Differences for test-retest were found to be not significant. No differences were found between the HR_{max} responses of the subjects for the direct and indirect tests. Of the subjects 75% had a lower predicted VO_{2max} value ($p<0.01$) compared with results gained by direct measurements. The reasons pointed out for this difference reclined on the different racial groups used as subjects, the climatic conditions in Singapore, or the small sample size.

Grant et al (1995) compared the results from a Cooper walk run test, a 20-m MSRT, and a submaximal cycle test with the direct measurement of maximum oxygen uptake on a treadmill. Three predictive tests of maximum oxygen uptake-linear extrapolation of heart rate of VO_2 collected from a submaximal cycle ergometer test (predicted_{L/E}), the Cooper 12 min walk, run test, and a 20-m MSRT- were performed by 22 young healthy males (mean \pm SD; age 22.1 ± 2.4 years; body mass 72.4 ± 8.9 kg) and the values compared to those obtained by direct measurement on a maximal treadmill test. All of the subjects were regular exercisers. The (mean \pm SD) from the various tests in ml.kg⁻¹.min⁻¹ were as follows: treadmill 60.1 ± 8.0 , Cooper 60.6 ± 10.3 , 20-m MSRT 55.6 ± 8.0 , and predicted_{L/E} 52.0 ± 8.4 . The Cooper test had a correlation with the treadmill test of 0.92, while the 20-m MSRT and the predicted_{L/E} had correlations of 0.86 and 0.76 respectively. Both the 20-m MSRT and predicted_{L/E} showed systematic underprediction of the treadmill value. On average, the 20-m MSRT was 4.5 ± 0.9 ml.kg⁻¹.min⁻¹ lower than the treadmill VO_{2max} while the predicted_{L/E} was 7.8 ± 1.4 ml.kg⁻¹.min⁻¹ lower than the treadmill VO_{2max} . These findings indicate that, for the population assessed, the Cooper walk run test was the best predictor of VO_{2max} among the three tests.

St Clair Gibson et al (1998) assessed the validity of the 20-m MSRT in predicting VO_{2max} in athletes trained in sports with different physical demands. Over a 14 day period, 10 male long-distance runners (continuous high-intensity exercise) and 10 male squash players (intermittent high-intensity exercise) performed, in random order, the 20-m MSRT and a maximal treadmill protocol of increasing elevation which elicited VO_{2max} . Compared with direct measurement of VO_{2max} , the 20-m MSRT significantly underpredicted the VO_{2max} of the group of runners ($n=10$; $p<0.01$) and of the athletes as a whole ($n=20$; $p<0.01$). The correlation between VO_{2max} determined by the 20-m MSRT and by the treadmill protocol was 0.61 ($p<0.05$) for the squash players, 0.71 ($p<0.05$) for the runners and 0.67 ($p<0.01$) for the athletes as a whole. The authors affirmed that as the subjects represented athletes trained specifically in sports with different physical demands this may explain why the correlations between the 20-m MSRT and direct measurement of VO_{2max} were less robust than those reported in previous studies. Furthermore, the results indicate that there are sport-specific differences when predicting VO_{2max} from the 20-m MSRT.

Wilkinson et al (1999) determined the incidence of subject drop-out on a 20-m MSRT and a modified incremental 20-m MSRT in which speed was increased by 0.014 m.s⁻¹ every 20-m shuttle to avoid the need for verbal speed cues. Analysis of the 20-m MSRT with 208 elite female netball players and 381 elite male lacrosse players found that 13 ± 3 players stopped after the first shuttle of each new level, in comparison with 5 ± 2 players on any other shuttle. No obvious drop-out pattern was observed on the incremental shuttle run test with 273 male and 79 female undergraduate students. The mean difference between a test-retest condition ($n=20$) for peak shuttle running speed (-0.03 ± 0.01 m.s⁻¹) and maximal heart rate (0.4 ± 0.1

bpm) on the incremental test showed no bias ($p > 0.05$). The relationship ($n = 27$) between peak shuttle running speed on the incremental shuttle run test ($4.22 \pm 0.14 \text{ m}\cdot\text{s}^{-1}$) and $\text{VO}_{2\text{max}}$ ($59.0 \pm 1.7 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) was high ($r = 0.91$; $p < 0.01$), with a standard error of prediction of $\pm 2.6 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$. These results suggest verbal cues during the multi-stage shuttle run test may influence subject drop-out. The 20-m MSRT shows no obvious drop-out pattern and provides a valid estimate of $\text{VO}_{2\text{max}}$.

Flouris et al (2005) started from the assumption that a $\text{VO}_{2\text{max}}$ of $44 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ is an accepted criterion ($\text{VO}_{2\text{CR}}$) below which health and fitness for young male adults may be compromised. In this way, the authors developed new algorithms validated for $\text{VO}_{2\text{CR}}$ screening using the 20-m MSRT. Maximal oxygen uptake was assessed in 110 males using a stationary gas analyser in a treadmill test (TMT) and in 40 of these subjects using a portable gas analyser in the 20-m MSRT. Significant energy cost variance (EC_v) was detected between TMT and 20-m MSRT ($p < 0.001$), correlated significantly with subject height, and was a significant predictor of $\text{VO}_{2\text{max}}$ differences between TMT and 20-m MSRT. Two equations predicting $\text{VO}_{2\text{max}}$ during 20-m MSRT (EQ_{MST}) and TMT (EQ_{TMT}) were developed. The R^2 of EQ_{MST} was 0.92 ($p < 0.001$). Predicted $\text{VO}_{2\text{max}}$ values from EQ_{MST} correlated with directly measured 20-m MSRT $\text{VO}_{2\text{max}}$ at $r = 0.96$ ($p < 0.001$). ANOVA detected no mean difference ($p > 0.05$) between predicted and measured values.

HEALTH-RELATED UTILIZATION SCOPE OF THE 20-M MSRT IN CHILDREN AND YOUTH

The 20-m MSRT has been widely used in different settings not necessarily seeking its validation. Here are stated some references of its wide utility (see Table 13).

Cooley and McNaughton (1999) used the 20-m MSRT to determine the aerobic fitness of Tasmanian secondary school children aged 11 to 16 years. Boreham et al (2001) examined the independence and relative strengths of association between CHD risk status and both body fatness and CRF in 12 and 15 years old adolescents. Chun et al (2000) validated criterion-referenced standards for cardiovascular endurance tests using the criterion-groups validation model in adolescents' boys and girls. Mota et al (2002) with a cross-sectional study aimed to 1) estimate changes in body composition and CRF across stages of pubertal maturation, and 2) describe the relationship between maturity status and body fatness, regional fat distribution, and CRF. Buck et al (2008) investigated the relation between aerobic fitness and interference control (one component of executive control) in 74 children between 7 and 12 years of age. At last, Baquet et al (2002) examined the $\text{VO}_{2\text{peak}}$ responses of prepubescent children following a 7-week aerobic training.

Table 13: Several Other 20-m MSRT Studies in Children and Youth.

Author	Participants	Performance variable	Results
Cooley and McNaughton (1999)	2981M 3180F 11-16 years	Total laps	11 years M: 46.9 F: 44.5 12 years M: 47.9 F: 43.8
Chun et al (2000)	240M + 232F 75% M= 12 years 85% F=12 years	Total laps	M MRW: 9.72 ± 2.30 M MSRT: 48.23 ± 19.75 F MRW: 10.99 ± 2.25 F MSRT: 36.14 ± 14.60
Boreham et al (2001)	251M 12 years 258F 12 years 252M 15 years 254F 12 years	Total laps	12 years (shuttles) M: 58.7 F: 44.3 15 years (shuttles) M: 79.2 F: 45.4
Baquet et al (2002)	23M + 30F 8-11 years	Max Speed at the last complete stage	Control 46.2 ± 8.5 Pre-training 43.9 ± 6.2
Mota et al (2002)	254M + 240F 9-12 years	Max Speed	Male: 9.1 ± 0.7 years: 48.2 ± 3.4 10.3 ± 1.0 years: 46.9 ± 4.7 11.9 ± 1.2 years: 48.6 ± 4.2 Female: 9.2 ± 0.7 years: 46.7 ± 2.8 10.4 ± 0.9 years: 45.7 ± 2.8 ml
Buck et al (2008)	74 M&F 7-12 years	Total laps	Shuttles: All: 23.2 ± 12.2 7-8 years: 21.1 ± 10.9 9-10 years: 23.5 ± 11.7 11-12 years: 25.8 ± 15.9

Results in VO_{2max} ($ml \cdot kg^{-1} \cdot min^{-1}$) or Total Laps.

PREVIOUSLY REPORTED VO_{2MAX} PREDICTIVE EQUATIONS IN CHILDREN AND YOUTH BASED ON THE 20-M MSRT PERFORMANCE

In general, there are a few unanimous concerns to consider in the development of VO_{2peak} estimating equations from the 20-m MSRT. In the first issue, and as referred earlier, we have to consider that the methods by which VO_{2peak} is predicted from the 20-m MSRT vary along the researchers. These include maximal shuttle run speed or “paliers”, with each palier lasting 1 min, the total number of laps (one lap being 20 m), or a combination of these methods (Leger and Lambert 1982; van Mechelen, Hlobil et al. 1986; Boreham, Paliczka et al. 1990; Anderson 1992; Barnett, C. et al. 1993; McVeigh, Payne et al. 1995). Therefore, this will obviously reflect on the existing variables in the 20-m MSRT VO_{2peak} predictive equations. In example, Barnett et al (1993) McVeigh et al (1995) recognized shuttle run performance (maximal running speed) as the best single predictor of VO_{2peak} . This is a one of a kind reason to insert this variable in a VO_{2max} equation. However, this is not a consensual position as we may see from Table 14.

The second issue concerns the necessity to include skinfold measurements in the estimation of VO_{2peak} from the 20-m MSRT. Researchers who included skinfold measurements in their regression formulas reported an improvement in the predictive power of the test (Barnett, C. et al. 1993; McVeigh, Payne et al. 1995). At last, the third issue is how gender and age affect the VO_{2peak} predictive power of the 20-m MSRT. Leger and Gadoury (1989) and Leger and Lambert (1982) affirmed that gender had little effect on predicting VO_{2peak} . However, Barnett et al (1993) found that an inclusion of gender factor in the determination of the regression equations used to predict VO_{2peak} improved the accuracy of prediction. Boreham et al (1990) indicated a higher correlation between the 20-m MSRT and VO_{2peak} for girls than boys ($r=0.90$ vs. $r=0.64$, respectively), and

McVeigh et al (1995) found that the plot of VO_{2peak} against 20-m MSRT performance showed distinct gender groups and divergent regression lines for the two genders. Leger et al (1988) and Barnett et al (1993) identified age as a critical variable since it takes into account, at least partially, changes in running economy that result from growth.

Table 14: Previously Reported VO_{2max} Predictive Equations in Children and Youth Based on 20-m MSRT Performance.

Author	Participants	Performance variable	Correlation Coefficient	Equation
Leger et al (1988)	200M&F 8-19 years	Max speed	0.71	$VO_{2max} = 31.025 + 3.238(\text{Speed}) - 3.248(\text{Age}) + 0.1536(\text{Age})(\text{Speed})$
Barnett et al (1993)	27M + 28F 12-17 years	Max speed	0.84	1) $VO_{2max} = 28.3 - 2.1(\text{sex: M=0; F=1}) - 0.7(\text{Tric}_{skinf}) + 2.6(\text{MS})$
			0.85	2) $VO_{2max} = 25.8 - 6.6(\text{sex: M=0; F=1}) - 0.2(\text{weight}) + 3.2(\text{MS})$
			0.82	3) $VO_{2max} = 24.2 - 5.0(\text{sex: M=0; F=1}) - 0.8(\text{age}) + 3.4(\text{MS})$
McVeigh et al (1995)	15M + 18F M: 13.68 ± 0.3 F: 13.68 ± 0.3	Max speed	0.68M	M: $VO_{2max} = 4.94(\text{MS}) - 0.30(\text{Tric}\&\text{Sub}_{skinf}) - 1.16$
			0.85F	F: $VO_{2max} = 7.26(\text{MS}) - 0.56(\text{Tric}_{skinf}) - 26.86$
Oliveira (1998)	77M + 80F 15.0 \pm 0.8	Total laps	0.78	$VO_{2max} = 25 + 6.6(\text{sex: M=2; F=1}) + 0.23(\text{TL})$
Pitetti et al (2002)*	13M + 38F 8-15 years	Total laps	0.88	$VO_{2max} = 0.35(\text{TL}) - 0.59(\text{BMI}) - 4.61(\text{sex ;M=1; F=2}) + 50.6$
Matsuzaka et al (2004)	62M + 70F 8-17	Max speed	0.81	1) $VO_{2max} = 25.9 - 2.21(\text{sex: M=0; F=1}) - 0.0449(\text{age}) - 0.831(\text{BMI}) + 4.12(\text{MS});$
		Total laps	0.80	2) $VO_{2max} = 61.1 - 2.20(\text{sex: M=0; F=1}) - 0.462(\text{age}) - 0.862(\text{BMI}) + 0.192(\text{TL})$

*Validation study of the equation developed by Fernhall et al (1998) for children with mild and moderate mental retardation, in nondisabled, healthy youths.

VALIDATION STUDIES OF THE 20-M MSRT AS A PERFORMANCE OUTPUT PREDICTOR

More than a few previously reported investigations in the literature have attempted to identify the physiological attributes associated with distance running success. There are some evidences gathered in the past years: 1) the possession of a large VO_{2max} has frequently been connected with successful running performance (Astrand, Stromme et al. 2003); 2) differences in running economy, which has been defined as the oxygen consumption for a given submaximal treadmill running velocity, discriminate between individuals possessing similar VO_{2max} values in terms of performance (Bar-Or 1983); 3) the ability to utilize a large fraction of the VO_{2max} , the proportion of slow twitch fibers in the running musculature, and the rate of accumulation of lactate in the plasma or blood have also been recognized as factors which may determine distance running success (Ramsbottom, Nute et al. 1987).

Table 15 assembles a few examples of 20-m MSRT validation studies as a performance output predictor in addition to its ability to estimate VO_{2max} as previously referred.

Table 15: 20-m MSRT Validation Studies as a Performance Output Predictor.

Author	Participants	VO ₂ test criterion	Performance variable	Correlation Coefficient	VO _{2max}
Paliczka et al (1987)	9M 26-47 years	Treadmill Mean speed of the 10 km run 2% ↑/1min	Total laps	0.93	59.0 ± 9.9
Ramsbottom et al (1988)	36M, 38F 19-36 years	Treadmill Exhaustion speed 7-10 min 3.5% initial and then 2.5% ↑/3min	Total stages	0.82M 0.89F 0.92 M&F	M: 58.5 ± 47.0 F: 47.4 ± 6.1
Berthoin et al (1994)	17 M&F 22.6 ± 5.6 years	Treadmill 8 km.h ⁻¹ F 10 km.h ⁻¹ M 2 km.h ⁻¹ ↑/4min	Max Speed at the last completed stage	0.82	
Aziz et al (2005)	21M 17 ± 0.4 years	Treadmill 10 Km. h ⁻¹ 0% ↑/2min, then 2% ↑/1min, until 12%. Then 1 Km. h ⁻¹ ↑/1min	Total distance	0.86	59.1 ± 4.8

VO_{2max} in ml.kg⁻¹.min⁻¹.

Paliczka et al (1987) assessed the validity of a 20-m MSRT as both a field test of CR endurance and as a predictor of competitive performance in a 10 km race. Nine male subjects (age 35.4 ± 5.8 years) underwent a laboratory test of maximum oxygen uptake on a treadmill (VO_{2max} 59.0 ± 9.9 ml.kg⁻¹.min⁻¹), completed the 20-m MSRT (score 105 ± 23.7 laps/11.4 ± 2.7 paliers) and competed in a 10 km race (finishing time 41.8 ± 7.3 min). Analysis using Pearson product-moment correlation revealed high correlations between these variables (20-MSRT vs. VO_{2max}, $r=0.93$; 20-MSRT vs. 10 km, $r=-0.93$; VO_{2max} vs. 10 km, $r=-0.95$). These results confirm that the 20-MSRT is a valid field test of CR endurance and suggest that it can additionally be used to predict relative running performance over 10 km.

Ramsbottom et al (1988) examined the validity of using a 20-MSRT to estimate VO_{2max}. Running ability was described as the final level attained on the shuttle run test and as time on a 5 km run. Maximal oxygen uptake was determined directly for 74 volunteers (36 men, 38 women) who also completed the 20-MSRT. Maximal oxygen uptake values were 58.5 ± 7.0 and 47.4 ± 6.1 ml.kg⁻¹.min⁻¹ for the men and women respectively (mean ± SD, $p<0.01$). The levels attained on the 20-MSRT were 12.6 ± 1.5 (men) and 9.6 ± 1.8 (women; $p<0.01$). The correlation between VO_{2max} and shuttle level was 0.92. The correlation between VO_{2max} and the 5 km run was -0.94 and the correlation between both field tests was -0.96. The results of this study suggest that the 20-m MSRT provides a valid estimate of VO_{2max} and indicates 5 km running potential in active men and women.

Berthoin et al (1994) carried out the measurement of maximal aerobic speed (MAS) and the prediction of VO_{2max} by means of field tests on 17 students studying physical education. The subjects underwent a continuous multi-stage track test (Leger and Boucher 1980), shuttle test (Leger and Gadoury 1989) and VO_{2max} measurement on a treadmill. The VO_{2max} values estimated using the track test (56.8 ± 5.8 ml.kg⁻¹.min⁻¹) were not significantly different from the values measured in the treadmill test (56.8 ± 7.1 ml.kg⁻¹.min⁻¹), but were higher than those estimated using the shuttle test (51.1 ± 5.9 ml.kg⁻¹.min⁻¹). The means of the MAS observed in the track test (15.8 ± 1.9 km.h⁻¹) and in the treadmill test (15.9 ± 2.6 km.h⁻¹) were not significantly different ($p>0.10$). The mean of the shuttle test MAS (13.1 ± 1 km.h⁻¹) was significantly lower ($p<0.01$) than those of the other tests. However, the MAS of the shuttle test and track test were linked. The equation for linear regression between MAS values in these two tests is $MAS_{track} = 1.81 \times MAS_{shuttle} - 7.86$ ($r=0.91$), allowing estimation of one of these MAS values when the other is known. Thus these values may be used within diversified training.

Aziz et al (2005) compared the performances obtained during soccer-specific field tests of the 20-m MSRT and the Yo-Yo intermittent endurance test (YIET), with the measured VO_{2max} obtained in both field tests as well as that obtained in the traditional test of running to exhaustion on a treadmill (TM), in young trained soccer players. Twenty-one national-level youth players performed, in random order, the 20-m MSRT and YIET to determine the relationship between the two field tests. From these, eight randomly chosen players performed their field tests as well as a TMT, equipped with an ambulatory gas exchange measurement device. Pearson product-moment correlation coefficient analysis showed that the players' performance (i.e. distance covered) in the 20-m MSRT and YIET was correlated ($r=0.65$, $p<0.01$). Players' performance in the YIET was not significantly correlated with the measured VO_{2max} obtained in the same YIET nor with the measured VO_{2max} obtained in the 20-m MSRT and in the TMT (all $p>0.05$). In contrast, significant correlations were observed between the players' performance in the 20-m MSRT with the measured VO_{2max} obtained in the same 20-m MSRT and in the YIET (both $p<0.05$); and attained almost statistical significance with the measured VO_{2max} in the TMT ($p=0.06$). The lack of association between distances covered in the YIET with all the measured VO_{2max} values suggest that measured VO_{2max} per se may not be suitable to characterize soccer players' intermittent endurance performance. In comparison with the 20-m MSRT, the YIET may be a more favourable field-based assessment of soccer player's endurance performance.

RELIABILITY OF THE 20-M MSRT IN CHILDREN AND YOUTH

The reliability of a test it is closely a requisite to its validity. For a test to be reliable it is necessary to have similar outputs in at least two different occasions. If that does not occur, the values are not well correlated with the standard measure, which in case of the VO_{2max} corresponds to a direct measure using a gas analyser.

The first test-retest reliability coefficients ($r=0.89$) of the 20-m MSRT in children (139 boys and girls 6-16 years old) were given by Leger et al (1988) (see Table 16). Liu et al (1992) determined the test-retest reliability of the 20-m MSRT (number of laps) on American children 12-15 years old. An intraclass coefficient of 0.93 was obtained on 20 students (12 males; $r=0.91$ and 8 females; $r=0.87$) who completed the test twice, 1 week apart, identical to the reliability found by Oliveira (1998) in 157 Portuguese adolescents. McVeigh et al (1995) examined the reliability of the 20-m MSRT as a predictor of VO_{2peak} in Edinburgh school children and observed a moderate correlation between trials ($r=0.75$). A similar correlation was observed by Mahoney (1992) with a total sample of 103 12 years old boys and girls (retest $n=20$; $r=0.73$, boys; $r=0.88$, girls; $p<0.01$). At last, Pitetti et al (2002) predicting the CVF based on the 20-m MSRT for nondisabled youth and for youth with mild mental retardation (ages 8 to 15 years; 38 females and 13 males) found high test-retest correlations for the 20-m MSRT (n° of laps; $r=0.89$).

Table 16: Reliability of the 20-m MSRT in Children and Youth.

Author	Participants	VO_2 test criterion	Reliability
Leger et al (1988)	139M&F 6-16 years	Retroextrapolation at the end of MSRT	0.89M&F
Liu et al (1992)	22M + 26F 12-15 years	Treadmill progressive	0.93M&F 0.87F 0.91M
Mahoney (1992)	53M + 50F 12 years	Treadmill M: 8; 9,4 Km. h ⁻¹ 5% ↑/2min F: 4.8 Km. h ⁻¹ 5% ↑/2min	0.73M 0.88F
McVeigh et al (1995)	15M + 18F M: 13.68 ± 0.3 F: 13.68 ± 0.3	Treadmill 7-8 Km. h ⁻¹ 0% /2min, then 9-10 Km. h ⁻¹ 2% /2min, until exhaustion.	0.75
Oliveira (1998)	77M + 80 F 15.0 ± 0.8	Portable gas analyser Cortex Metamax, CBS	0.93
Pitetti et al (2002)	13M + 38F 8-15 years	Treadmill Modified Bruce Protocol	0.89M&F

Studies in adults demonstrate higher correlations than in children, independent of the 20-m MSRT version in use (see Table 17).

Leger and Lambert (1982) testing 91 adults (32 females and 59 males, aged 27.3 ± 9.2 and 24.8 ± 5.5 year respectively and with VO_{2max} (mean \pm SD) of 39.3 ± 8.3 and 51.6 ± 7.8 ml. kg^{-1} . min^{-1} respectively, reported a high test-retest correlation ($r=0.975$, $SEE=2.0$, $n=50$). The authors concluded that the 20-m MSRT is a reliable test for the prediction of the VO_{2max} of male and female adults, individually or in groups, on most gymnasium surfaces, as they would later reconfirm in a sample of 81 adults (20-45 years) with a correlation coefficient of 0.95 (Leger, Mercier et al. 1988). Sproule et al (1993) in a sample of twenty Asian subjects (16 male, 4 female), all of whom were physical education students, and using a modified form of Leger and Lambert (1982) 20-m MSRT, confirmed a test-retest reliability for both direct and indirect measurements ($r=0.90$ and $r=0.91$ respectively). Differences for test-retest were found to be not significant. Wilkinson et al (1999) analysis of the 20-m MSRT in 208 elite female netball players and 381 elite male lacrosse players found a mean difference between a test-retest condition ($n=20$) for peak shuttle running speed (-0.03 ± 0.01 m. s^{-1}) and HR_{max} (0.4 ± 0.1 bpm) on the incremental test. The mean difference showed no bias ($p>0.05$).

Despite the reliability values cited above, Lamb and Rogers (2007) began the investigation stating that most studies that have set out to quantify the test-retest reliability of the 20-m MSRT have typically used inappropriate statistics (correlation coefficients) and/or one of two possible calculation methods for deriving predictions of VO_{2max} , and have not considered the impact of habituation on performance in the 20-m MSRT and whether its reliability improved beyond a single repeat trial. In this context, their study assessed the reliability of the 20-m MSRT over three trials. Thirty-five healthy and active university games players (22 males and 13 females) performed the 20-m MSRT on three occasions, 1 week apart. Individual total numbers of completed shuttles were recorded and VO_{2max} estimated values were derived via established equation and table methods. Analysis revealed that the overall mean VO_{2max} score from the equation method (52.5 ± 7.8 ml. kg^{-1} . min^{-1}) was significantly higher than that for the table (46.9 ± 8.9 ml. kg^{-1} . min^{-1}) whilst the mean trial 2 and trial 3 scores were significantly higher than that for trial 1, but not different to each other. The Limits of Agreement (LA) for the table method were -1.4 ± 5.0 (trial 1-trial 2) and 0.0 ± 5.5 ml. kg^{-1} . min^{-1} (trial 2-trial 3), and for the equation method -1.1 ± 4.7 (trial 1-trial 2) and 0.0 ± 5.0 ml. kg^{-1} . min^{-1} (trial 2-trial 3). These results suggest that systematic bias is eliminated after the first trial (due to habituation), but a considerable amount of random error remains, regardless of the type of score calculated. Hence, the authors concluded that the 20-m MSRT does not appear to be reliable enough for the purpose of monitoring changes in VO_{2max} due to non-random reasons.

Table 17: Reliability of the 20-m MSRT in Adults.

Author	Participants	VO_2 test criterion	Reliability
Leger and Lambert (1982)	59M + 32F M: 24.8 ± 5.5 F: 27.3 ± 9.2	Retroextrapolation at the end of MSRT	0.98
Leger et al (1988)	81M&F 20-45 years	Retroextrapolation at the end of MSRT	0.95
Sproule et al (1993)	16M + 4F M: 21-35 years F: 20-21 years	Treadmill Taylor Protocol	0.91
Wilkinson et al (1999)	25M + 2F 22.6 ± 3.4 years	Treadmill 4 min: Exhaustion speed 8-12 min $1\% \uparrow/1min$ and then 0.1 km. $h^{-1} \uparrow/5$ seg	0.34

THE 20-M MSRT IN THE WORLD ACROSS THE YEARS

The 20-m MSRT has been widely studied since its development. Table 18 intends to elucidate to the magnitude of its power as a study focus.

Table 18: Normative Studies of the 20-m MSRT Around the World. Adapted from T. Olds et al (2006).

Reference	Country	Year	Sex	Age range	Sample sizes
ACHPER (1996)	Australia	1994	M&F	9 – 18	39 – 104
Australian Sports Commission (1994)	Australia	1993	M&F	12 – 17	301 – 542
Booth et al. (1997)	Australia	1997	M&F	9, 11, 13, 15	399 – 634
Brewer et al. (1988)	Australia	1991	M&F	11 – 14	57 – 256
Cooley and McNaughton (1999)	Australia	1998	M&F	11 – 16	339 – 636
Western Australian Institute of Sport (pers. comm., 2002)	Australia	1999 – 2001	M&F	12 – 15	951 – 3461
Tasmanian Institute of Sport (pers. comm., 2002)	Australia	1996 – 2001	M&F	12 – 15	1520 – 3304
Gibbs et al. (1995)	Australia	1995	F	9.5+0.7	36 – 38
Jill Hage (pers. comm., Dec. 1999)	Australia	1999	F	13, 14, 16	173 – 107
Hands (2000)	Australia	1999	M&F	6 – 12	14 – 37
Jenner et al. (1992)	Australia	1988	M&F	12+0.4	527 – 565
A. Lindquist (pers. comm., Dec. 1999)	Australia	1998 – 1999	F	7 – 14	2 – 40
Lloyd & Antonas (2000)	Australia	1999?	M&F	11 – 12	35 – 53
Shane Masson (pers. comm., Dec. 1999)	Australia	1999	M&F	14, 16	6 – 11
Sara Mulkearns (pers. comm., Aug. 2001)	Australia	1996 – 1998	F	8 – 10	45 – 79
Okely et al. (1997)	Australia	1996	M	14	51
Annette Raynor (pers. comm., Sept. 2001)	Australia	1999	M&F	6 – 12	5 – 20
South Australian Sports Institute (pers. comm., 2002)	Australia	2001	M&F	11 – 15	8 – 1116
Tomkinson et al. (2003b)	Australia	1995 – 2000	M&F	12 – 15	1018 – 4451
Vandongen et al. (1995)	Australia	1990	M&F	10 – 12	485 – 486
Mr. Colin Walker (pers. comm., 2002)	Australia	1998	M	13	13
Sue Woolard (pers. comm., 1999)	Australia	1999	F	8 – 13	20
Baquet et al. (2000)	Belgium	1997	M&F	11 – 18	6 – 88
Beunen et al. (1991)	Belgium	1990	M&F	6 – 18	84 – 423
Lefèvre et al. (1998)	Belgium	1993, 1997	M&F	12 – 18	40 – 288
Nupponen & Telama (1998)	Belgium	1994 – 1995	M&F	12, 15	423 – 700
Pirnay (1995)	Belgium	1991 – 1992	M&F	10 – 18	90 – 247
Poortmans et al. (1986)	Belgium	1984 – 1985	M&F	6 – 19	9 – 82
Cazorla et al. (1987)	Benin	1987	M&F	8 – 18	139 – 425
Gouthon (1987)	Benin	1987	M&F	13 – 18	16 – 35
Falgairrette et al. (1994)	Bolivia	1992	M	10, 13	12 – 16
Hobold (2003)	Brazil	2002?	M&F	7 – 17	90 – 153
Pieta (2000)	Brazil	1999?	M&F	7 – 14	83 – 148
Léger et al. (1984)	Canada	1981	M&F	6 – 17	112 – 404
Massicotte (1990)	Canada	1989 – 1990	M&F	6 – 17	75 – 402
Cazorla et al. (1985)	Côte d'Ivoire	1985	M&F	7 – 18	57 – 314
Gajda (1994)	Czech Republic	1992 – 1993	M&F	7 – 14	83 – 188
Lars Bo Andersen (pers. comm., Sept. 2001)	Denmark	1996 – 1997	M&F	15 – 19	183 – 3198
Cazorla et al. (1986b)	Djibouti	1986	M&F	11 – 18	20 – 182
Jurimae & Volbekiene (1998)	Estonia	1993 – 1994	M&F	11 – 17	133 – 294
Kull & Jurimae (1994)	Estonia	1992 – 1993	M&F	17+0.3	132 – 207
Oja & Jurimae (1998)	Estonia	1996 – 1997	M&F	6	133 – 161
Raudsepp & Jurimae (1996)	Estonia	1994 – 1995	M	7 – 10	45 – 55
Raudsepp & Jurimae (1997)	Estonia	1994 – 1995	F	7 – 10	48 – 57
Viljanen et al. (2000)	Finland	1993	M&F	7, 10, 14, 16	63 – 249
Nupponen & Telama (1998)	Finland	1995	M&F	12, 15	254 – 307
Baquet et al. (1999)	France	1992	M&F	6 – 11	14 – 49
Baquet et al. (2001)	France	1997	M&F	10 – 15	8 – 88
Baquet et al. (2000)	France	2000	M&F	8 – 10	3 – 12
Blonc et al. (1992)	France	1991?	M&F	11 – 16	20 – 124
Cazorla (1987)	France	1987	M&F	7 – 12	119 – 693
Cazorla et al. (1997)	France	1996 – 1997	M&F	7 – 11	47 – 110
Carzorla et al. (1999)	France	1999	M&F	11 – 18	85 – 234
Colin Walker (pers. comm., March 2002)	France	1992	M	10, 13	28 – 41
Levarlet-Joye & Fievetz (1990)	France	1989?	M	11 – 14	7 – 33
Van Praagh et al. (1987)	France	1985 – 1986	M&F	10 – 11	64 – 75
Van Praagh et al. (1988)	France	1986	M	7, 12	15 – 18
Nupponen & Telama (1998)	Germany	1994 – 1995	M&F	12, 15	216 – 264
Georgiadis (1993)	Greece	1990 – 1991	M&F	6 – 18	103 – 297
Manios et al. (1999)	Greece	1992 – 1993	M&F	6	264 – 305
Eston et al. (1998)	Hong Kong	1995 – 1996	M&F	7 – 15	5 – 32
Macfarlane et al. (1998)	Hong Kong	1995 – 1996	M&F	8 – 11	28 – 31
Barabás & Sebestyén (1995)	Hungary	1992 – 1994	M	18 – 19	330 – 475
Nupponen & Telama (1998)	Hungary	1994 – 1995	M&F	12, 15	100 – 114
Gunnarsson & Sigriksón (1999)	Iceland	1998	M&F	6 – 15	255 – 528
MacDonncha et al. (1999)	Ireland	1998?	M	15	22
Bellucci (1997)	Italy	1994	M&F	12 – 14	10 – 23
Cilia & Bellucci (1993)	Italy	1992	M&F	12 – 14	189 – 259
Cilia et al. (1996)	Italy	1995	M&F	12 – 15	23 – 120
Cilia et al. (1997)	Italy	1997	M&F	12 – 19	73 – 404

Cilia et al. (1998)	Italy	1997	M&F	14 – 19	16 – 68
Council of Europe (1986)	Italy	1985 – 1986	F	14 – 18	44 – 79
MECSST (1999 – 2002)	Japan	1998 – 2001	M&F	6 – 19	52 – 1123
Matsuzaka et al. (2004)	Japan	2004	M&F	18 – 23	62 – 70
Jurimae & Volbekiene (1998)	Lithuania	1993	M&F	11 – 17	66 – 150
Cazorla et al. (1986a)	Mauritius	1986	M&F	7 – 17	137 – 160
van Mechelen et al. (1986)	Netherlands	1983	M&F	12	41
van Mechelen et al. (1987)	Netherlands	1987	M&F	12 – 16	101 – 285
Maciaszek & Osinski (2001)	Poland	2000?	M&F	10 – 14	245 – 371
Mleczo & Ozimek (2000)	Poland	1991 – 1992	M&F	15 – 19	150 – 303
Przeweda & Dobosz (2003)	Poland	1999	M&F	7 – 19	623 – 3470
Mota et al. (2002)	Portugal	1998 – 2000	M&F	9 – 15	17 – 115
Miguel Oliveira (pers. comm., 2002)	Portugal	2001 – 2002	M&F	10 – 19	28 – 250
Present study	Portugal	2008	M&F	8 – 10	90
Cazorla et al. (1988)	Sénégal	1988	M&F	7 – 18	72 – 163
Cazorla et al. (1990)	Seychelles	1990	M&F	7 – 18	142 – 259
Ms. Annette Raynor (pers. comm., July 2002)	Singapore	1999	M&F	6 – 11	4 – 17
Kasa & Majherova (1997)	Slovakia	1996	M&F	11 – 14	24 – 45
Kyselovicova (2000)	Slovakia	1993 – 1995	F	15 – 16	15 – 38
Moravec (1996)	Slovakia	1993	M&F	7 – 19	54 – 554
du Randt (1996)	South Africa	1995 – 1996	M&F	11 – 16	45 – 210
Garcia Baena (1999)	Spain	1999	M&F	13 – 18	10 – 70
Brito Ojeda et al. (1995)	Spain	1993?	M&F	10 – 19	35 – 50
Prat et al. (1998)	Spain	1984 – 1985	M&F	10 – 18	140 – 267
Javier Rivas (pers. comm., June 2002)	Spain	1987	M&F	9 – 18	6 – 303
Sainz (1992)	Spain	1986 – 1989	M&F	10 – 15	37 – 180
Sainz (1996)	Spain	1990 – 1994	M&F	9 – 17	54 – 671
Lieveld et al. (1993)	Suriname	1992	M	14	70 – 115
Cauderay et al. (2000)	Switzerland	1996 – 1997	M&F	9 – 19	36 – 214
Calis et al. (1992)	Turkey	1991	M	15	19 – 30
Armstrong et al. (1988)	UK (England)	1987?	M	12	77
McVeigh et al. (1995)	UK (Scotland)	1994?	M&F	13	15 – 18
Eston et al. (1998)	UK (Wales)	1995 – 1996	M&F	7 – 15	7 – 15
Boreham et al. (1987)	UK (Northern Ireland)	1986	M&F	13, 15	13 – 32
Boreham et al. (1990)	UK (Northern Ireland)	1988	M&F	15	18 – 23
Boreham et al. (2001)	UK (Northern Ireland)	1989 – 1990	M&F	12, 15	251 – 258
Mahoney & Boreham (1992)	UK (Northern Ireland)	1991	M&F	7, 9, 11	91 – 109
Mahoney (1992)	UK (Northern Ireland)	1990 – 1991	M&F	12	50 – 53
Nichols & Riddoch (1986)	UK (Northern Ireland)	1986	M&F	12 – 15	32 – 163
Riddoch (1990)	UK (Northern Ireland)	1988 – 1989	M&F	11 – 18	87 – 237
Twisk et al. (1999)	UK (Northern Ireland)	1992 – 1993	M&F	15	229 – 230
Michael Beets (pers. comm., 2003)	USA	1999 – 2001	M&F	8 – 19	5 – 63
Michael Beets (pers. comm., 2003)	USA	2003	M&F	14 – 19	5 – 32
Chun et al. (2000)	USA	1997	M&F	12	116 – 120
Dinschel (1994)	USA	1993?	M&F	9 – 12	3 – 39
Liu et al. (1992)	USA	1990	M&F	13 – 14	22 – 26
Mahar et al. (1997)	USA	1995	M&F	10	113 – 144
Pitetti et al. (2002)	USA	2001?	M&F	10, 11	13 – 28
Wolford (1998)	USA	1997	M&F	10	97 – 102

METHODOLOGY

INTRODUCTION

In this content we intend to describe the experimental concept of this investigation. We further explain the preparation procedures to PACER test performance plus the statistical treatment of the data of interest to the present study.

The aim of this study was to compare the validity of the FITNESSGRAM's MSRT software reports and previously reported equations as predictors of CVF in Lisbon healthy elementary school children, by measuring the VO_{2peak} during the actual 20-m MSRT performance. We additionally assessed the reliability of the 20-m MSRT in a sub-group of subjects that performed the test twice.

EXPERIMENTAL CONCEPT

In an attempt to follow an ecological approach to the validation of the PACER test in Lisbon healthy elementary school children, 90 subjects performed this field test as described in The Cooper Institute for Aerobics Research (2002). The test environment was as authentic as possible, which is in fact the main purpose of these kinds of tests. In the same manner, we opted for the presence of a single investigator, who represented the figure of the teacher, who is mostly certain to be alone while teaching.

The VO_{2peak} results obtained at the end of the PACER test were compared to the estimated FITNESSGRAM's software reports as a predictor of VO_{2max} . In addition, we compared the measured directly VO_{2peak} values with the predicted VO_{2max} derived from previously reported equations for children and youth.

In order to assess the reliability of the PACER test performed by the participants in this study, a sub-group of 12 volunteer children randomly selected by age and gender completed the PACER test twice, with at least 6 days interval.

PARTICIPANTS

This study used a small convenience sample on a local basis in which children volunteered to participate. Ninety children 8-10 years old from an elementary school of the great Lisbon district were subsequently selected to participate accordingly to their age and sex, allowing the harmonization of the sub-groups (see Table 19).

Table 19: Sample's Selection.

Age (years)	Boys	Girls
8	15	15
9	15	15
10	15	15

INSTRUMENTS AND PROTOCOLS

In the following section we are going to present and describe the instruments and protocols used in the present study.

ANTHROPOMETRIC MEASURES

Height and abdominal circumference were measured using a tape measure to the nearest 0.1 cm. Light upward pressure was applied under the jaw and occipital with two hands to provide maximum extension of the spine. Weight was measured using standard weighing scales (Seca, Hamburg) to the nearest 100 g. Triceps and calf skinfolds were measured using a "Slimguide Skinfold Caliper". The values were introduced in the software and retrieved from the child's fitness report.

Following the indication of The Cooper Institute for Aerobics Research (2002), the triceps skinfold was measured on the back of the right arm over the triceps muscle, midway between the elbow and the acromion process of the scapula. The skinfold site was vertical. The calf skinfold was measured on the inside of the right leg at the level of maximal calf girth. The right foot

was placed flat on an elevated surface with the knee flexed at a 90° angle. The vertical skinfold was grasped just above the level of maximal girth and the measurement made below the grasp. Each skinfold was measured three times and the value registered was the median of those measures.

20-M MSRT TEST PROTOCOL

The PACER test is a multistage fitness test adapted from the 20-meter shuttle run test (20-m MSRT) published by Leger and Lambert (1982) and revised in 1988 (Leger, Mercier et al. 1988). The test is progressive in intensity, it is easy at the beginning (8.0 km.h⁻¹) and gets harder at the end increasing 0.5 km.h⁻¹ every stage after the first minute (see Table 20). The progressive nature of the test provides a built-in warm-up and helps children to pace themselves effectively.

The test has also been set to music to create a valid, fun alternative to the customary distance run test for measuring aerobic capacity. The PACER test CD has two music versions: the beep version and the cadence version. Each version of the test gives a 5 sec countdown and tells the students when to start. Regarding sonority clearness, in the present study all the subjects performed the PACER test with the beep version.

Participants should run across the 20-m distance and touch the line with their foot by the time the beep sounds. At the sound of the beep, they turn around and run back to the other end. If some students get to the line before the beep they must wait for the beep before running the other direction.

A single beep will sound at the end of the time for each lap. A triple beep sounds at the end of each minute. The triple beep serves the same purpose as the single beep does and also alerts the runners that the pace will get faster.

The first time a student does not reach the line by the beep, the student stops where he or she is and reverses direction immediately, attempting to get back on pace. The test is completed for a student the next time (second time) he or she fails to reach the line by the beep (the two misses do not have to be consecutive; the test is over after two total misses).

In the PACER test, a lap is 1 x 20-meter distance (from one end to the other). Have the scorer record the lap number (crossing off each lap number) on a PACER score sheet. The recorded score is the total number of laps completed by the participant.

As a matter of further interest, the protocol of this test suggests that students ages 5 to 9 years until grade 3 do not have to receive a score by the software or the teacher, they may simply participate in the activity. With the software, it is possible to enter the actual number of laps or enter a score of 0 laps to indicate that they successfully participated in the PACER run. Regardless of the entry, the performance will not be evaluated against a criterion standard. Nine-year-olds in grade 4 may receive a score, and it will be evaluated against a criterion standard. All 10-year-old students receive a score regardless of grade level.

Table 20: Running Speed (Km.h⁻¹) and Number of Laps for Each Stage of the PACER Test According to The Cooper Institute for Aerobics Research (2002).

Stage	PACER Test (The Cooper Institute for Aerobics Research 2002)	
	Speed	Laps
1	8.0	7
2	9.0	8
3	9.5	8
4	10.0	9
5	10.5	9
6	11.0	10
7	11.5	10

OXYGEN UPTAKE TRANSIENTS MEASURES

The determination of the ventilatory and pulmonary gas exchange responses to exercise requires measurement of either volume or airflow and the fractional concentrations of oxygen (FEO₂) and carbon dioxide (FECO₂) in the expired air. From measurements of these 3 signals, valuable exercise variables and parameters can be determined and/or calculated. Among them are minute ventilation (V_E), respiratory frequency (F_R), and tidal volume (V_T), as well as oxygen uptake (VO₂) and carbon

dioxide output (VCO_2). Other useful exercise variables that can be derived from these responses include ventilatory equivalent for oxygen (V_E/VO_2) and carbon dioxide (V_E/VCO_2), the end-tidal partial pressure of carbon dioxide ($PETCO_2$), and respiratory exchange ratio (RER) (Paridon, Alpert et al. 2006). To accomplish the purpose of this study, VO_{2max} was obviously the major parameter determined in open circuit.

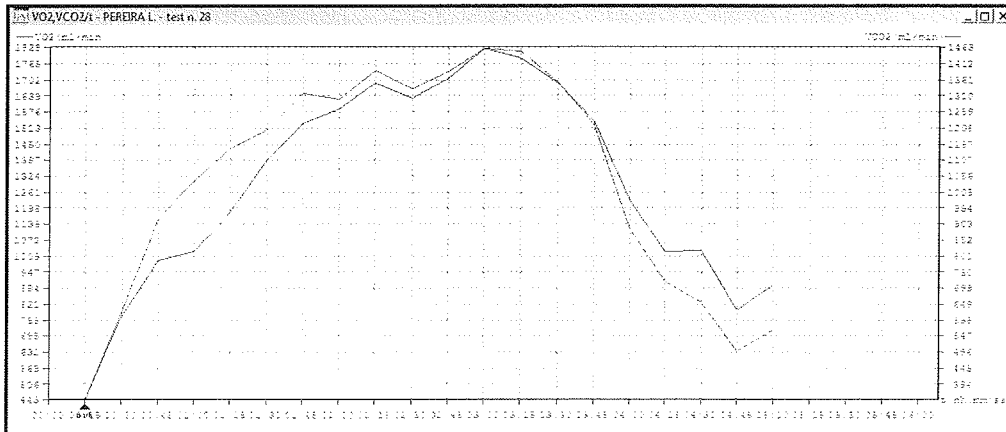


Figure 1: Typical (VO_2 and VCO_2 ; $ml \cdot min^{-1} / t$) line graphic representation after averaging and smoothing the data measured through K4b².

During the graded test, inspired and expired gases were continuously monitored, breath by breath, through a portable gas analyser (K4b², Cosmed, Rome, Italy) (see Figure 1), which had been previously validated (McLaughlin, King et al. 2001). Before each test, the analysis systems of O_2 and CO_2 were calibrated using ambient air and standard gases of known concentration (16.7% O_2 and 5.7% CO_2). The calibration of the turbine flowmeter of the K4b² was performed using a 3-l syringe (Quinton Instruments, Seattle, Wash., USA) according to the manufacturer's instructions. Both VO_2 and VCO_2 values were averaged to 15 s. Heart Rate was continuously monitored (*Polar T31 coded™ transmitter*). This compact device was easy to attach without constricting children's movements.

The K4b² (see Figure 2) weights 475g and was not expected to affect significantly the energy demands of the subjects. In fact, none of the subjects had any negative remarks of the system's weight and on their mobility and vision during the test (see Figure 3, Figure 4, Figure 5, Figure 6). A pilot study of Flouris et al (2005) using five subjects (age: 21.6 ± 1.3 ; BMI: 24.3 ± 1.5) was conducted in order to investigate additional energy demands and ensure that significant agreement existed between two gas analyzers employed: an automated gas analyser (Vmax 29, SensorMedics, Yorba Linda, CA) and the referred portable gas analyser (K4b², Cosmed, Rome, Italy). The subjects performed a treadmill test twice using both gas analyzers. Results showed no significant difference ($p > 0.05$) between the mean VO_{2max} value recorded by the stationary (Vmax 29, SensorMedics) and the portable (K4b², Cosmed, Rome, Italy) gas analyzer (48.7 ± 3.1 vs. 49.1 ± 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively), with an average absolute error of 0.51 ± 0.18 $ml \cdot kg^{-1} \cdot min^{-1}$.

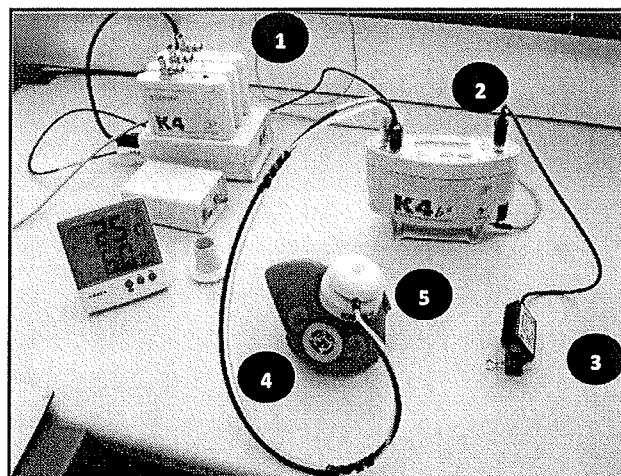


Figure 2: The K4b2 Metabolic Measuring System: 1) Battery + Antenna; 2) K4b2 unit; 3) HR probe; 4) Mask; 5) Flowmeter.

CD AUDIO PLAYER

During the entire research period, only one copy of the PACER test audio CD and one CD player model *CFD-S03CPSONY* were used. The copy of the PACER test audio CD was gently conceded by the Faculty of Human Kinetics, Lisbon.

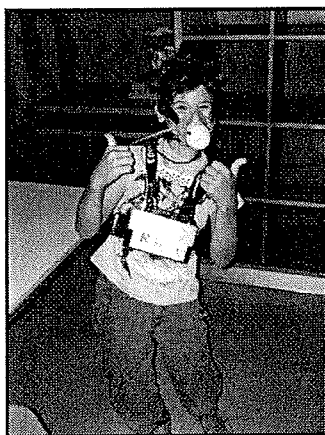


Figure 3: Front image of a participant wearing still the K4b², Cosmed

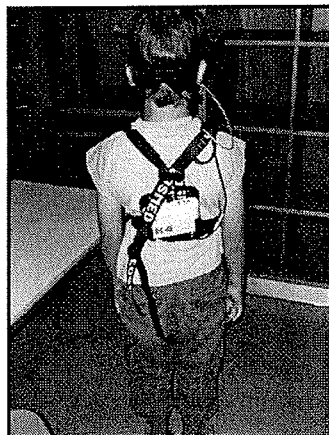


Figure 4: Back image of a participant wearing still the K4b², Cosmed

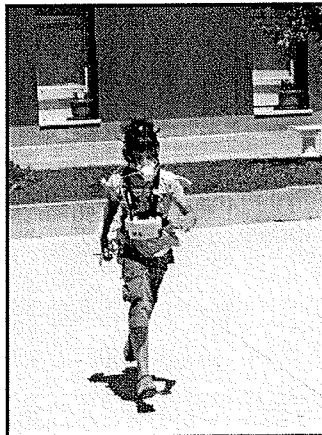


Figure 5: Front image of a participant wearing the K4b², Cosmed while running



Figure 6: Back image of a participant wearing the K4b², Cosmed while running

As described in The Cooper Institute for Aerobics Research (2002), the PACER test audio CD was calibrated previously to each trial and was consistently within the 0.5 sec interval.

TIME AND SPACE RESOURCES

The PACER test was performed in an outdoor playground (see Figure 7) of a school of the Great Lisbon district (E.B 1 J.I Agualva 1), during the summer months of July, August and September 2008.

The reliability of the PACER test in the present study was established in a sub-group of 12 volunteer children randomly selected by age and gender who completed the PACER test twice, within a 6 days period.

The concrete surface and the children's trainers ensured safe and outperforming PACER trials, even though it also increased significantly the testing temperature as it reflected the sunlight. Attending to children well-being, the PACER test was only performed in two possible periods a day: from 09h00 to 11h00 and from 15h30 to 18h30 avoiding the peak sunlight hours. The environmental temperature oscillated between 26^o-29^oC, the atmospheric pressure varied between 744-749 mmHg and the humidity was around 50% across the entire research period.

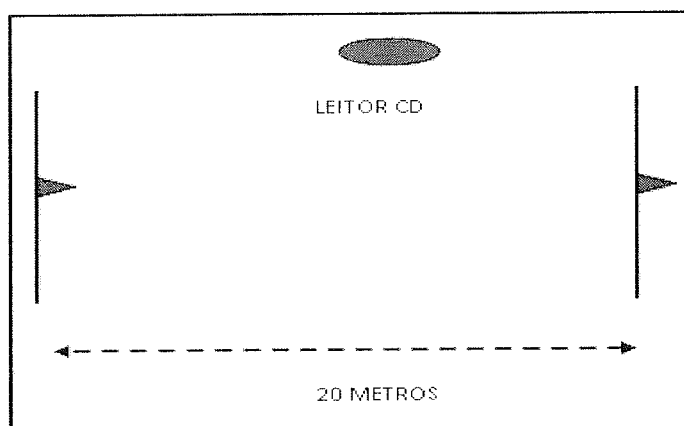


Figure 7: Disposition of the different elements of the test.

DATA ASSORTMENT

The spirometric variables (VO_2 , VCO_2 , V_E , RER, ...) were collected using a portable gas analyser (K4b², Cosmed, Rome, Italy) and later synchronized with the Cosmed software (version 8.0b) via a USB plug. This equipment also collected data regarding the atmospheric pressure, humidity and temperature. The HR data were also collected by the K4b² with a *Polar T31 coded™ transmitter*. Beyond these variables, the total number of complete shuttle runs was registered.

DATA PREPARATION

Almost all the desired data were integrally provided by the Cosmed Software (version 8.0b) and by the FITNESSGRAM software. Before each test, participant's age, height and weight were introduced in the K4b² module and after the synchronization with the PC, the spirometric outputs were presented. Every subject was added to the FITNESSGRAM software in order to have the respective PACER test performance, estimated VO_{2max} , and %Fat reports printed.

These records were then added to a Microsoft Excel database where other crucial data like the total number of shuttle runs completed, the duration of the test, the distance accomplished, the stage number, and so on, had already been previously introduced.

The estimated VO_{2max} values (mean \pm SD; $ml.Kg^{-1}.min^{-1}$) provided from several 20-m MSRT equations for children and youth were chosen accordingly to which best fitted in the age spectrum of this study (8-10 year old) or in the nationality of the subjects as in Oliveira (1998). The variables were added to the previously reported VO_{2max} estimation equations accordingly to the indications of each author. As follows, in the VO_{2max} estimation equation by Leger et al (1988), the last stage number announced was used to predict VO_{2max} ($ml.kg^{-1}.min^{-1}$) from the speed ($km.h^{-1}$) corresponding to that stage. In the equations of (Barnett, C. et al. 1993; Matsuzaka, Takahashi et al. 2004), maximal speed (MS) was calculated from the total number of laps (TL) attained. In all other lasting equations (Fernhall, Pitetti et al. 1998; Oliveira 1998; Matsuzaka, Takahashi et al. 2004), the TL attained was introduced as a variable in the respective equations without further treatment.

At last, a simple export of the desired data to *SPSS Statistics 17.0* and *MedCalc version 10.1.2.0* allowed its statistical treatment.

PROCEDURES

In the following sections we are going to report the procedures carried out before, during and after the PACER test performance.

BEFORE THE TEST

For the realization of the PACER test it was necessary to limit the 20-m course with a yellow tape and red chalk. To increase the course visibility colored field marker cones were utilized.

Age and anthropometrical data were collected previously to the test performance. Triceps and calf skinfold measurements were introduced in the software of the FITNESSGRAM to estimate participants' body fat, whilst abdominal circumference was registered to supplementary health-related analysis. In a later exam we verified that the outputs provided by the software were identical to the ones provided by the equation of Slaughter et al (1988).

The PACER test protocol as described in The Cooper Institute for Aerobics Research (2002) was clearly explained to each participant, despite the fact that the participants had already performed the test in their Physical Education classes. Emphasis was given to the need of a slow start in order to run as much time as possible and the participants were informed that when the triple beep sounds they should not stop but continue the test by turning and running towards the other end of the area instead.

As it was only permitted to test a child at a time, the next participant was invited to assist the performance of the previous participant in order to strengthen up the comprehension of the task. All the participants had their trainers on and tighten.

Even though there were several silicon face masks available, all the participants fit best with the face mask size 3. To assure it, each individual performed a forced expiration with an exit occlusion in order to detect possible escapes of expired gases. All the participants wore a *Polar T31 coded™ transmitter*.

Before each test, the K4b² analysis systems of O₂ and CO₂ were calibrated using ambient air and standard gases of known concentration (16.7% O₂ and 5.7% CO₂). The calibration of the turbine flowmeter of the K4b² was performed using a 3-l syringe (Quinton Instruments, Seattle, Wash., USA) according to the manufacturer's instructions.

After a short general warming, the test began.

DURING THE TEST

All the inaccurate events during the PACER test performance were immediately corrected. Inadequate fast start, bad synchronization between the audio signal and the arrival or departure time of the subject to one of the 20-m course extremities, untimed inversion of direction and utilization of a curvilinear course instead of a stop-and-go rectilinear course, and a mistaken stop at the sound of triple beeps, are a few examples of mistaken events.

Subjects were highly encouraged to achieve their maximum effort through verbal support. During the test the investigator was deeply focused on any sign of intolerance to the effort. In such case, the test would be interrupted without delay as it only happened once.

AFTER THE TEST

The end of the test was determined by the investigator, while the total number of laps was confirmed with an assistant who noted accurately the participant's performance.

The 20-m MSRT performed was only considered maximum when verified at least two of the following: 1) voluntary exhaustion; 2) a HR greater than 90% of the HR_{max} predicted (220-age); 3) a RER greater than 1.0; or 4) subjective opinion of experienced testers (Barnett, C. et al. 1993).

Two minutes was the minimum active recovery time of the participants.

STATISTICAL ANALYSIS

Paired Student's *t*-test was used to compare the mean spirometric values between the measured directly VO_{2peak} values and the estimated FITNESSGRAM's software reports as well as with all other previously reported VO_{2max} predictive equations. This technique was also employed to compare mean spirometric and performance values in test and retest trials where the normality of the distribution was assumed, allowing a discussion around the learning effect. In those cases where the normal distribution was not established, Wilcoxon non-parametric tests were performed.

Independent Student's *t*-test and the Mann-Whitney test were used to compare mean anthropometric, spirometric and performance values between boys and girls when the normal distribution was confirmed, or not, respectively.

The One-way ANOVA was used to test for differences in mean VO_{2peak} between age groups of each gender and to compare mean differences between measured directly VO_{2peak} and FITNESSGRAM software reports for the same groups. As the One-way ANOVA pointed out to significant differences between the means, Fisher's least significant difference (LSD) multiple comparisons were performed to detect in which ages those differences were significant in boys, girls and both. The assumption of normality was valid for each population.

The Friedman test which is a non-parametric statistical test similar to the parametric repeated measures ANOVA was used to perceive differences among previously reported VO_{2max} predictive equations in children and youth of several authors who dedicate their studies to this issue since Mercier et al (1983). As the Friedman test results pointed out to significant differences between the equations, Dunnett's procedure of a control group to other groups was used to find out which equations results were similar to the VO_{2peak} obtained by the K4b² in this study, and which were not. The non-parametric procedures were used since the normality assumption was rejected.

Shapiro-Wilk test was used before parametric and non-parametric tests to verify if the data being used had a normal distribution. Significant values were considered when $p < 0.05$.

Pearson product-moment correlation coefficients (r) were used to indicate the strength and direction of the linear relationship between measured directly VO_{2peak} and total number of laps (TL) or maximal speed (MS) attained. Similar relationships were sought between measured directly VO_{2peak} and BMI. They also served the same purpose when analyzing the relationship between measured directly VO_{2peak} and the previously reported VO_{2max} predictive equations. Spearman's rank correlation coefficient was used to indicate the relationship of measured directly VO_{2peak} and total number of laps (TL) with age.

Another effective technique used in the assessment of the relationship and reproducibility of the measurements was the concordance correlation coefficient (r_c). The VO_{2peak} values measured directly by the K4b² and the estimated VO_{2max} values from the FITNESSGRAM software reports were plotted in a graphic where the perfect reproducibility of assay would be manifested by the data falling on a 45° line intercepting the origin of the graph. Later on, the concordance correlation coefficient was added to all the tables where the relationship between measured directly VO_{2peak} and the previously reported VO_{2max} predictive equations was tested in order to compare the different magnitude of the correlations given by r and r_c .

We wanted to know how much the previously reported VO_{2max} estimation equations were likely to differ from the measured directly VO_{2peak} , if this was or was not enough to cause problems in the interpretation and consequently replace or not one method by the other or use the two interchangeably. Therefore, a Bland-Altman approach which is in its essence a plot of the difference between two methods against their mean was performed.

The statistical analyses were saved, computed and analyzed using the *SPSS Statistics 17.0*, *MedCalc 10.1.2.0* and *Microsoft Office Excel 2007*.

RESULTS

INTRODUCTION

The aim of this study was to compare the validity of the FITNESSGRAM's MSRT software reports and previously reported equations as predictors of CVF in Lisbon healthy elementary school children, by measuring the VO_{2peak} during the actual 20-m MSRT performance. Ninety subjects, 45 boys and 45 girls aged 8-10 years (in a balanced experimental design), performed the MSRT and their VO_{2peak} was measured directly in field using a portable gas analyzer (K4b², Cosmed, Rome, Italy).

We additionally assessed the reliability of the 20-m MSRT in a sub-group of subjects that performed the test twice.

EXPLORATORY ANALYSIS

This study used a small convenience sample on a local basis in which children volunteered to participate. Ninety children 8-10 years old were subsequently selected to participate accordingly to their age (8: n=30; 9: n=30; 10: n=30) and sex (boys: n=45; girls: n=45), allowing the harmonization of the sub-groups. Only a few non-caucasian children took part in this study, not reaching 5% of the whole sample, reflecting the school population in this age group. In Table 21 it is possible to glimpse the sample's mean ages and mean morphological characteristics.

Table 21: Age (years) and Morphologic Characteristics (Mean \pm SD): Height (m), Weight (Kg), Body Mass Index (BMI; kg/m^2) and %Fat.

	Boys (n=45)	Girls (n=45)	Total (n=90)
Age	9.0 \pm 0.8	9.0 \pm 0.8	9.0 \pm 0.8
Height	1.37 \pm 0.6	1.36 \pm 0.1	1.37 \pm 0.1
Weight	33.3 \pm 7.4	34.3 \pm 7.5	33.8 \pm 7.4
BMI	17.7 \pm 3.1	18.3 \pm 2.5	18.0 \pm 2.8
%Fat	20.0 \pm 7.9**	24.4 \pm 6.6	22.2 \pm 7.6

* Difference between the means of boys and girls ($p < 0.05$); ** Difference between the means of boys and girls ($p < 0.001$).

The Independent Student's *t*-test, used when a normal distribution was verified and the non-parametric Mann-Whitney test when it was not, revealed no differences between both boys and girls in all the parameters in Table 21, with exception for the percentage of fat (%Fat) ($p < 0.001$).

Table 22 shows the spirometric and cardiovascular variables (mean \pm SD) assessed at the end of the PACER test. The mean VO_{2peak} was 49.6 ± 9.0 $ml.kg^{-1}.min^{-1}$. Not surprisingly, boys had higher mean absolute ($l.min^{-1}$) and relative ($ml.kg^{-1}.min^{-1}$) VO_{2peak} values than girls (1.7 ± 0.5 ; 53.3 ± 9.2 vs. 1.4 ± 0.4 ; 46.0 ± 7.3 , respectively). The distribution of the values is observable in Figure 8, and the Shapiro-Wilk test verified that the distributions were not significantly different from normal distribution ($p > 0.05$) in all the variables of Table 22. Mean RER_{peak} was higher in girls (0.99 ± 0.1) than in boys (0.88 ± 0.1). This is interesting in the way that, accordingly to the independent Student's *t*-tests, mean HR_{max} was not significantly different between boys and girls ($p = 0.73$). On the contrary, mean absolute ($l.min^{-1}$) and relative VO_{2peak} ($ml.kg^{-1}.min^{-1}$) values and mean RER_{peak} were all found to be significantly different between genders ($p < 0.05$).

Table 22: Spirometric and Cardiovascular Variables (Mean \pm SD) Assessed During the PACER Test.

	Boys (n=45)	Girls (n=45)	Total (n=90)
VO_{2peak} ($l.min^{-1}$)	1.7 \pm 0.5*	1.4 \pm 0.4	1.6 \pm 0.5
VO_{2peak} ($ml.kg^{-1}.min^{-1}$)	53.3 \pm 9.2**	46.0 \pm 7.3	49.6 \pm 9.0
RER_{peak}	0.88 \pm 0.1**	0.99 \pm 0.1	0.94 \pm 0.1
HR_{max} (bpm)	195.1 \pm 10.1 ¹	195.9 \pm 8.0 ²	195.5 \pm 9.0 ³

1: (n=38); 2: (n=42); 3: (n=80).

* Difference between the means of boys and girls ($p < 0.05$); ** Difference between the means of boys and girls ($p < 0.001$).

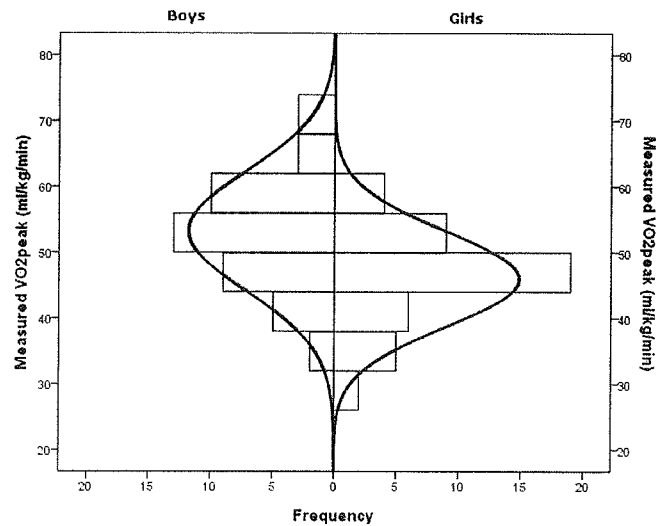


Figure 8: Gender based histogram of measured directly VO_{2peak} with normal curve displayed.

Boys had systematically superior VO_{2peak} values ($ml \cdot kg^{-1} \cdot min^{-1}$) than girls across all age spectrums (see Table 23). The uppermost measured directly VO_{2peak} values ($ml \cdot kg^{-1} \cdot min^{-1}$) were observed at 10 years old boys (56.8 ± 8.4) and 8 year old girls (48.9 ± 5.9) and it is rather interesting to notice that the smaller values were found at 9 years old for both boys and girls (50.4 ± 7.9 vs. 42.3 ± 7.3). The closest range values between boys and girls was recognizable at 8 years old (52.6 ± 10.4 vs. 48.9 ± 5.9 , respectively) and the widest at the age of 10 (see Figure 9). Attention should also be afforded to the fact that the VO_{2peak} variability tends to be higher in boys than in girls at all considered ages.

Table 23: Gender/Age Based Distribution of Measured Directly VO_{2peak} ($ml \cdot kg^{-1} \cdot min^{-1}$; Mean \pm SD).

	8	9	10	Total
Boys (n=45)	52.6 ± 10.4	50.4 ± 7.9	56.8 ± 8.4	53.3 ± 9.2
Girls (n=45)	$48.9 \pm 5.9^*$	$42.3 \pm 7.3^*$	46.7 ± 7.3	46.0 ± 7.3
Total (n=90)	50.7 ± 8.5	$46.4 \pm 8.5^*$	$51.8 \pm 9.3^*$	49.6 ± 9.0

* Difference between the means of each gender across ages 8, 9 and 10 years old ($p < 0.05$); ** Difference between the means of each gender across ages 8, 9 and 10 years old ($p < 0.001$).

The One Way Anova analysis revealed differences among the mean VO_{2peak} values in girls across the ages ($p = 0.03$) as well as when considering the entire sample ($p = 0.04$). Accordingly to Fisher's least significant difference (LSD) procedure, significant differences were found between 8 and 9 year old girls ($p = 0.01$) and between 9 and 10 year old participants when allowing for the entire sample.

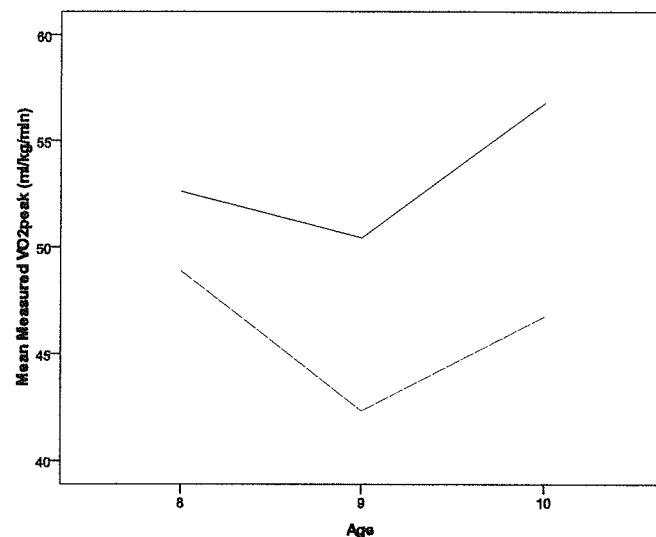


Figure 9: Mean measured directly VO_{2peak} in boys (full line) and girls (dashed line) across ages.

In relation to the performance variables attained during the PACER test trials, boys accomplished a greater number of laps than girls (31.2 ± 11.3 ; 21.6 ± 7.2 , respectively). The same was true to MS (10.2 ± 0.6 ; 9.6 ± 0.4), test time ($04:04 \pm 01:20$; $02:56 \pm 00:53$) and distance attained (624.4 ± 225.7 ; 432 ± 144.6), due to the intrinsic relation between the variables. The difference between the values of boys and girls was confirmed by the Mann-Whitney test in all the cited variables ($p < 0.001$) (see Table 24).

Table 24: Performance Variables Assessed (Mean \pm SD) during the PACER Test: Total Laps (n^o), Maximal Speed (Km.h⁻¹) Test Time (mm:ss) and Distance (m).

	Boys (n=45)	Girls (n=45)	Total (n=90)
Total laps	31.2 \pm 11.3**	21.6 \pm 7.2	26.4 \pm 10.6
Maximal speed	10.2 \pm 0.6**	9.6 \pm 0.4	9.9 \pm 0.6
Test time	04:04 \pm 01:20**	02:56 \pm 00:53	03:30 \pm 01:15
Distance attained	624.4 \pm 225.7**	432 \pm 144.6	528.2 \pm 211.8

* Difference between the means of boys and girls ($p < 0.05$); ** Difference between the means of boys and girls ($p < 0.001$).

VALIDITY OF THE 20-M MSRT AS A PREDICTOR OF VO_{2PEAK} IN LISBON ELEMENTARY SCHOOL CHILDREN

High significant positive Pearson product-moment correlation coefficients were seen between measured directly VO_{2peak} and total number of laps (TL) and maximal speed (MS) attained, except for girls whose r was only significantly moderate in both categories. Distance (total laps) rather than MS provided a better discriminatory measure of aerobic performance in boys ($r=0.75$ vs. $r=0.74$), girls ($r=0.69$ vs. $r=0.63$) and when taken as a single group ($r=0.77$ vs. $r=0.75$) (see Table 25).

Table 25: Pearson Product-Moment Correlations Between Measured Directly VO_{2peak} (ml.kg⁻¹.min⁻¹) and Total Number of Laps (n^o) and Between Measured Directly VO_{2peak} (ml.kg⁻¹.min⁻¹) and Maximal Speed (Km.h⁻¹).

	Total Laps	Maximal Speed
Boys (n=45)	0.75*	0.74*
Girls (n=45)	0.69*	0.63*
Total (n=90)	0.77*	0.75*

* Correlation is significant at the 0.05 level; ** Correlation is significant at the 0.001 level.

Pearson product-moment correlation coefficients between measured directly VO_{2peak} and weight, height, and BMI, as well Spearman's rank correlation coefficient between measured directly VO_{2peak} and age are shown in Table 26. Correlation between total number of laps for the PACER test and same morphologic variables were also evaluated. Even if some correlations were significant, they were all weak. Positive Spearman's rank correlation coefficient was obtained between the total number of laps for the PACER test and age ($r=0.27$), while negative relationships existed among both total number of laps for the PACER test and measured directly VO_{2peak}, with weight ($r=-0.31$ and $r=-0.45$, respectively) and BMI ($r=-0.46$ and $r=-0.55$, respectively) suggesting that both performance on the PACER test and measured directly VO_{2peak} are superior as BMI decreases. Non significant relationships were seen between the total number of laps for the PACER test and height and between the measured directly VO_{2peak} and age or height ($p > 0.05$).

Table 26: Pearson Product-Moment Correlations Between Measured Directly VO_{2peak} (ml.kg⁻¹.min⁻¹) and Total Laps (n^o), with Weight, Height, and BMI; and Spearman's Rank Correlation Between Measured Directly VO_{2peak} (ml.kg⁻¹.min⁻¹) and Total Laps (n^o) with Age.

	Age	Height	Weight	BMI
Measured VO _{2peak}	0.07	-0.10	-0.45*	-0.55*
Total laps	0.27*	0.06	-0.31*	-0.46*

* Correlation is significant at the 0.05 level; ** Correlation is significant at the 0.001 level.

VALIDITY OF THE FITNESSGRAM'S PACER TEST SOFTWARE REPORTS AS A PREDICTOR OF VO_{2PEAK} IN LISBON ELEMENTARY SCHOOL CHILDREN

The highest mean measured VO_{2peak} values (ml.kg⁻¹.min⁻¹) were obtained in 10 year old boys (56.8 ± 8.4) and the lowest in 9 year old girls (42.3 ± 7.3). From 8 to 10 years old and independently of sex, the VO_{2max} estimated by FITNESSGRAM consistently underpredicted VO_{2peak} scores with exception to 9 year old girls whose mean measured directly VO_{2peak} values

were below the mean estimated VO_{2max} values (48.9 ± 5.9 vs. 47.7 ± 1.6 , respectively). This event had necessarily a repercussion on the overestimation of the VO_{2peak} of the entire 9 year old population (46.4 ± 8.5 vs. 47.4 ± 3.0).

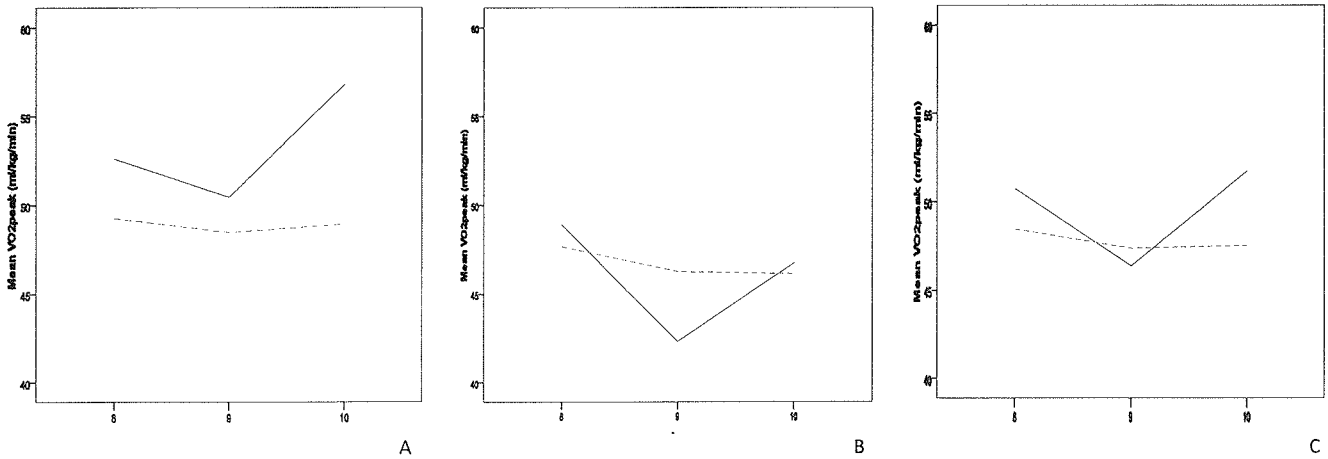


Figure 10: Mean measured directly VO_{2peak} (full line) and mean estimated VO_{2max} by FITNESSGRAM (dashed line) in: A) boys; B) girls; C) total, across ages.

In relation to the FITNESSGRAM's estimated VO_{2max} values across the ages, the line in Figure 10 tends to smooth the smaller values obtained at the age of 9 (all 3 categories included), in comparison with the measured directly values at the same age where it seems evident the existence of 3 clearly defined peaks (8, 9 and 10 years old). This line follows a much more rectilinear tendency than their measured directly counterpart assured in part by the narrower range of values between maximum and minimum VO_{2max} scores as we will later see in Figure 11.

The estimated VO_{2max} lines of boys (see Figure 10A) and girls (see Figure 10B) draw attention to a decrease of the aerobic capacity in this group of children across the ages. As mentioned before, this is not the case with the measured directly VO_{2peak} tendency line. However, it is observable a slightly increase in maximal aerobic power at the age of 10 when the entire sample is considered.

A deeper look into boys' line graphic representation (see Figure 10A) reveals the absence of any interception between the measured and estimated lines which means that, in average, the estimation VO_{2max} values of boys are constantly below the measured VO_{2peak} values.

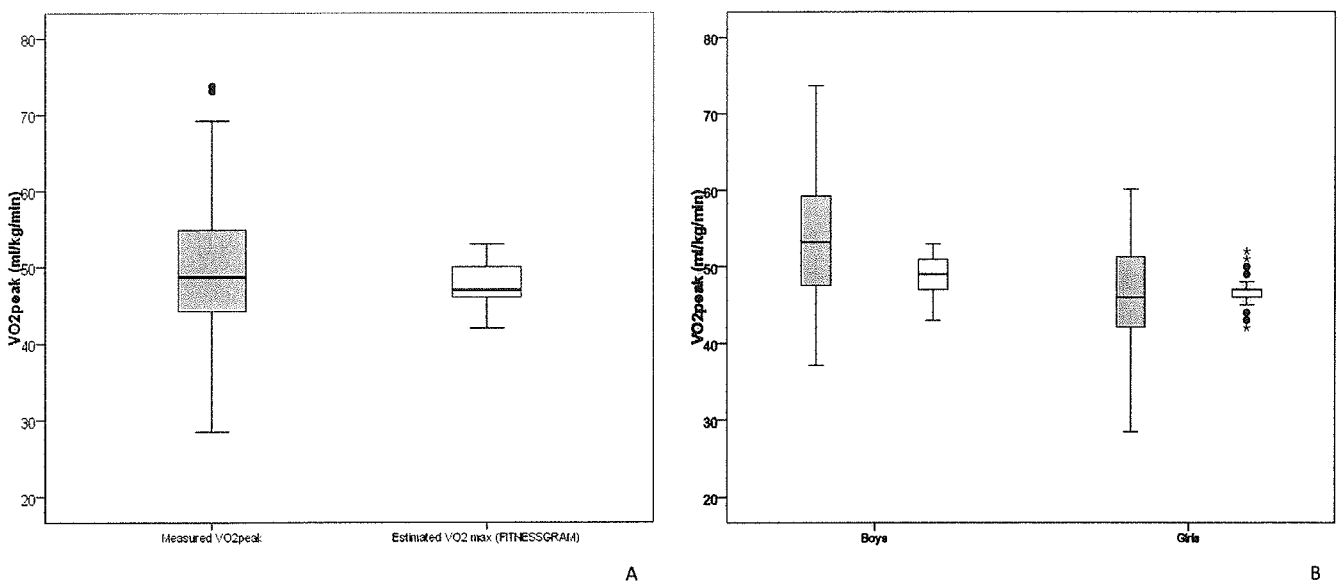


Figure 11: Boxplots of the measured directly VO_{2peak} (grey) and the estimated VO_{2max} by FITNESSGRAM (white) in: A) whole sample; B) split by gender.

Figure 11 provides a unique vision of the distinctly range of values between the measured directly VO_{2peak} and the estimated VO_{2max} values given by the FITNESSGRAM's software reports. In Figure 11A the entire sample is considered and it is noticeable the great concentration of values ($ml \cdot kg^{-1} \cdot min^{-1}$) of the estimated VO_{2max} by FITNESSGRAM ($Q_1=46.00$; $Q_3=50.00$), in

opposition to the fairly large dispersion of the VO_{2peak} values obtained directly ($Q_1=44.15$; $Q_3=54.97$). This pattern was exacerbated when both measured directly and estimated VO_{2peak} values were split by gender (see Figure 11B). If one consider boys for instance, the estimated VO_{2max} boxplot exhibits a superior concentration of values ($ml.kg^{-1}.min^{-1}$) ($Q_1=47.00$; $Q_3=51.00$) to the measured directly VO_{2peak} boxplot which is incomparably wider ($Q_1=47.38$; $Q_3=59.28$). As in for girls, although the variation of the measured directly VO_{2peak} is somewhat inferior ($Q_1=41.77$; $Q_3=51.64$), the eye gauging difference between the directly measured VO_{2peak} boxplot and the estimated VO_{2max} boxplot ($Q_1=46.00$; $Q_3=47.00$) appears to be even most pronounced than in boys. This is undoubtedly interesting in the way that such variation could not be imagined just by looking into the means of measured directly and estimated VO_{2peak} in Table 27.

Paired Student's *t*-test detected differences between both mean measured directly and estimated VO_{2peak} values ($ml.kg^{-1}.min^{-1}$) when considering the whole sample (49.6 ± 9.0 vs. 47.8 ± 2.7 ; $p=0.02$; $n=90$) and only boys (53.3 ± 9.2 vs. 48.9 ± 2.7 ; $p<0.001$; $n=45$). Girls' estimated and measured directly VO_{2peak} values were not significantly different (46.0 ± 7.3 vs. 46.7 ± 2.2 ; $p=0.45$; $n=45$).

Despite some variation in the measured directly VO_{2peak} and estimated VO_{2max} values referred above, paired student's *t*-test revealed no significant differences between the means of measured directly and estimated VO_{2peak} values of 8 and 9 years old boys and girls ($p>0.05$). On the other hand, at the age of 10, significant differences were detected among boys (56.8 ± 8.4 vs. 48.9 ± 2.8 ; $p<0.001$) and between both genders together (51.8 ± 9.3 vs. 48.9 ± 5.9 ; $p<0.001$).

Ongoing with the analysis, we went towards the age distribution of the mean differences (d ; $ml.kg^{-1}.min^{-1}$) between measured directly VO_{2peak} values and predicted VO_{2max} by FITNESSGRAM. The highest d was found in 10 year old boys (7.9 ± 6.8) and the smallest, also in 10 year old subjects, but this time in girls (0.6 ± 5.9). Girls' maximum disagreement values were found at 9 years old, where as previously reported, FITNESSGRAM's VO_{2max} values overestimated their aerobic capacity. In respect to total difference values per age, the lowest predictable power of FITNESSGRAM's reports was seen at 10 year old subjects. An overall view allow us to see that boys presented the highest estimation default value (4.4 ± 7.5) in comparison with girls whose slight overestimation most matched the VO_{2peak} (-0.7 ± 6.1). Thus, the mean whole sample difference among measured directly VO_{2peak} and the software reports by FITNESSGRAM was only $1.84 \pm 7.3 ml.kg^{-1}.min^{-1}$.

The One Way Anova analysis revealed differences among the mean differences (d) between measured directly VO_{2peak} values and FITNESSGRAM's predicted VO_{2max} in girls across the ages ($p=0.04$) as well as when considering the entire sample ($p=0.02$). Accordingly to Fisher's least significant difference (LSD) procedure, significant differences were found in the d of 8 and 9 year old girls ($p=0.02$) and also in 9 and 10 year old girls ($p=0.04$). When allowing for the whole sample, significant differences were also seen between the d of 9 and 10 year old participants ($p=0.01$).

Table 27: Gender/Age Based Measured Directly VO_{2peak} and Estimated VO_{2max} by FITNESSGRAM (Mean \pm SD; $ml.kg^{-1}.min^{-1}$); Mean Differences ($d \pm$ SD; $ml.kg^{-1}.min^{-1}$), Pearson Product-Moment Correlation (r) and Concordance Correlation (r_c) Between Measured Directly and Estimated VO_{2peak} .

Age	Gender	Measured VO_{2peak}	FITNESSGRAM	d	r	r_c
Total	Boys (n = 45)	53.3 ± 9.2	48.9 ± 2.7	$4.4 \pm 7.5^{**}$	0.70	0.31
	Girls (n = 45)	46.0 ± 7.3	46.7 ± 2.2	-0.7 ± 6.1	0.61	0.34
	Total (n = 90)	49.6 ± 9.0	47.8 ± 2.7	$1.8 \pm 7.3^*$	0.72	0.38
8	Boys (n = 15)	52.6 ± 10.4	49.3 ± 2.1	3.3 ± 9.0	0.72	0.26
	Girls (n = 15)	48.9 ± 5.9	47.7 ± 1.6	1.2 ± 4.7	0.79	0.39
	Total (n = 30)	50.7 ± 8.5	48.5 ± 2.0	2.3 ± 7.1	0.75	0.37
9	Boys (n = 15)	50.4 ± 7.9	48.5 ± 3.2	2.0 ± 5.5	0.84	0.56
	Girls (n = 15)	42.3 ± 7.3	46.3 ± 2.4	-4.0 ± 6.7	0.40	0.19
	Total (n = 30)	46.4 ± 8.5	47.4 ± 3.0	-1.0 ± 6.7	0.71	0.44
10	Boys (n = 15)	56.8 ± 8.4	48.9 ± 2.8	$7.9 \pm 6.8^{**}$	0.69	0.23
	Girls (n = 15)	46.7 ± 7.3	46.1 ± 2.4	0.6 ± 5.9	0.70	0.41
	Total (n = 30)	51.8 ± 9.3	47.5 ± 2.9	$4.2 \pm 7.3^{**}$	0.77	0.31

* Difference between the means ($p<0.05$); ** Difference between the means ($p<0.001$).

Table 27 resumes the correlations between the measured directly VO_{2peak} values and the FITNESSGRAM's VO_{2max} software reports. Pearson product-moment correlation coefficient (r) was higher when considering the entire sample ($r=0.72$), instead of the split by sex ($r=0.70$ for boys and $r=0.61$ for girls). When accounting for age, the highest r was found in 9 year old boys ($r=0.84$) and the smallest in girls of the same age ($r=0.40$).

In validation processes it is widely common the use of the Pearson product-moment correlation coefficient, the paired t -test, the least squares analysis of slope ($=1$) and intercept ($=0$), the coefficient of variation, or the intraclass correlation coefficient. There are drawbacks to all of these, however, in that none alone can fully assess the desired reproducibility characteristics (Lin 1989). Trailing the hint of Lin (1989), we additionally used a reproducibility index called concordance correlation coefficient (r_c), which evaluates the agreement between two readings (from the same sample) by measuring the variation from the 45° line through the origin (the concordance line).

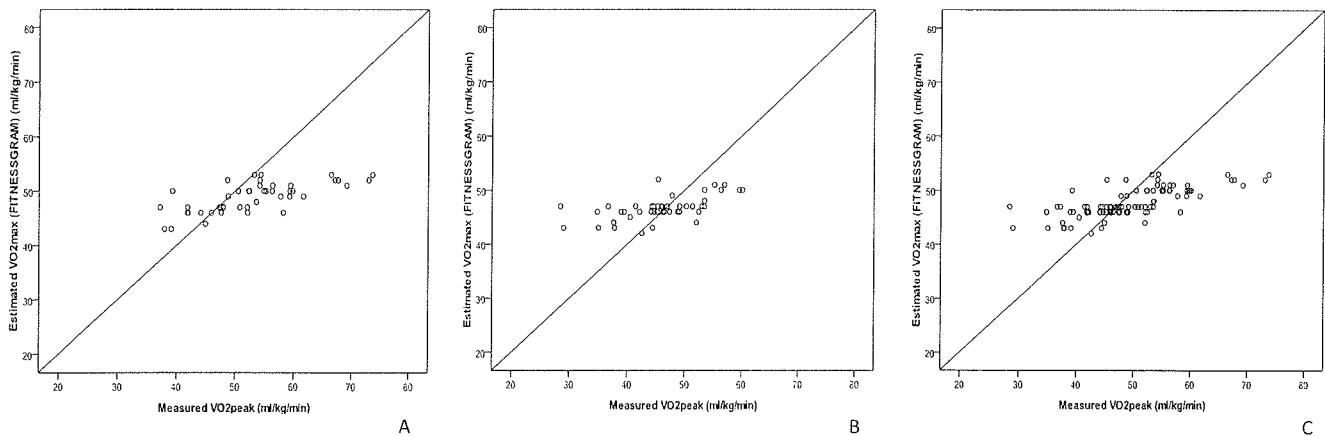


Figure 12: Scatterplot of estimated VO_{2max} by FITNESSGRAM against measured directly VO_{2peak} values in: A) boys; B) girls; C) both.

When we plot the predicted VO_{2max} and the measured directly VO_{2peak} values into a scatterplot, the dissimilarities between the two methods are immediately evident (see Figure 12). It may be important to make clear that both measured directly VO_{2peak} and estimated VO_{2max} by FITNESSGRAM would only match if the values were coincident with the 45° line drawn. Furthermore, the values that are to the right of the line suggest an underestimation of the predictive equation, in this specific case, the VO_{2max} software reports by FITNESSGRAM. The inverse is possible to deduce when the values are to the left of the 45° line. As we might have guessed from the interpretation of the difference values counting boys, girls and the entire sample of all ages (see Table 27), the variation of these values is somewhat inferior when considering girls (Figure 12B) and the whole sample (Figure 12C) in opposition to boys (Figure 12A). In the present study, the VO_{2peak} tends to be underestimated in boys (Figure 12A), overestimated in girls (Figure 12B) and without a clear tendency in both (Figure 12C).

Therefore, a different pattern was followed when analyzing the relationship between the measured directly VO_{2peak} values and FITNESSGRAM's predicted VO_{2max} values using the concordance correlation coefficient (r_c). Even though the highest correlation value was found when considering both boys and girls together ($r_c=0.38$), this time it were girls who had the highest correlation value ($r_c=0.31$ for boys and $r_c=0.34$ for girls). The utmost score was seen among 9 year old boys ($r_c=0.56$), and the smallest between 9 year old girls ($r_c=0.19$). All the concordance correlation coefficients were manifestly inferior to the Pearson product-moment correlation coefficients.

This disagreement between the values of r and r_c was already expectable from the first impression of Figure 12. Pearson product-moment correlation measures the strength of the relation between two variables, not the agreement between them. It is clearly evident the divergence between the two methods from the line of equality drawn in boys, girls and both as a single group.

COMPARING SEVERAL REGRESSION EQUATIONS THAT PREDICT VO_{2PEAK} USING THE 20-M MSRT IN LISBON ELEMENTARY SCHOOL CHILDREN

Table 28 puts in evidence the estimated VO_{2max} values (mean ± SD; ml.Kg⁻¹.min⁻¹) provided from several 20-m MSRT equations that predict VO_{2peak} in children and youth, selected accordingly to which best fitted the age spectrum of this study (8-10 year old) or in the nationality of the subjects as in Oliveira (1998). It may be important to reinforce the idea that all the specific variables introduced in each equation were based exclusively on the PACER test performance of the subjects and in anthropometric variables evaluated earlier to the test.

When considering the entire sample, the higher estimating values (53.9 ± 4.6; n=90) were obtained using the equation nº 2 of Barnett et al (1993) [VO_{2max} = 28.3 - 2.1(sex: M=0; F=1) - 0.7(Tric_{Skinf}) + 2.6(MS)]. In fact, this equation overestimated not only the VO_{2max} when considering the entire sample, but also when it was split by sex (49.6 ± 1.4 for girls and 58.1 ± 2.0 for boys). On the other hand, the smallest prediction values (41.0 ± 4.9; n=90) were seen using the equation of Oliveira (1998) [VO_{2max} = 25 + 6.6(sex: M=2; F=1) + 0.23(TL)]. Even when the split by sex was performed, the equation of Oliveira (1998) still underestimated the most the VO_{2max} values in girls (36.6 ± 1.7; n=45) and in boys (45.4 ± 12.6; n=45).

FITNESSGRAM's VO_{2max} software reports were the values that most closely predicted the measured directly VO_{2peak}. Yet, this is only true for girls (46.7 ± 2.2) and for both genders together (47.8 ± 2.7). Boys' best predictive values (49.6 ± 1.8) were obtained with the equation nº 3 of Barnett et al (1993) [VO_{2max} = 24.2 - 5.0(sex: M=0; F=1) - 0.8(age) + 3.4(MS)].

Table 28: Gender/Age Based Measured VO_{2peak} and Estimated VO_{2max} Values (mean ± SD; ml.Kg⁻¹.min⁻¹) Derived from Several 20-m MSRT Estimating Equations for Children and Youth.

Age	Gender	Measured VO _{2peak}	FITNESSGRAM	Leger et al (1988)	Barnett et al (1993) ¹	Barnett et al (1993) ²	Barnett et al (1993) ³	Fernhall et al (1998)	Oliveira (1998)	Matsuzaka et al (2004) ¹	Matsuzaka et al (2004) ²
Total	Boys (n=45)	53.3 ± 9.2	48.9 ± 2.7	46.4 ± 2.9	46.5 ± 4.9	58.1 ± 2.0	49.6 ± 1.8	46.5 ± 5.1	45.4 ± 2.6	46.9 ± 4.5	47.7 ± 4.1
	Girls (n=45)	46.0 ± 7.3	46.7 ± 2.2	43.4 ± 2.9	40.6 ± 4.1	49.6 ± 1.4	42.7 ± 1.4	38.1 ± 3.4	36.6 ± 1.7	41.6 ± 3.8	43.1 ± 3.0
	Total (n=90)	49.6 ± 9.0	47.8 ± 2.7	44.9 ± 3.3	43.6 ± 5.4	53.9 ± 4.6	46.1 ± 3.8	42.3 ± 6.0	41.0 ± 4.9	44.3 ± 4.9	45.4 ± 4.3
8	Boys (n=15)	52.6 ± 10.4	49.3 ± 2.1	47.0 ± 1.0	46.8 ± 3.5	57.2 ± 1.4	49.5 ± 1.4	45.2 ± 3.1	44.3 ± 1.7	46.8 ± 2.7	47.7 ± 2.6
	Girls (n=15)	48.9 ± 5.9	47.7 ± 1.6	45.1 ± 2.0	40.0 ± 3.2	49.4 ± 0.9	43.3 ± 0.9	38.3 ± 2.8	36.3 ± 1.3	42.5 ± 3.1	44.1 ± 2.6
	Total (n=30)	50.7 ± 8.5	48.5 ± 2.0	46.1 ± 2.1	44.4 ± 4.1	53.3 ± 4.1	46.4 ± 3.3	41.7 ± 4.6	40.3 ± 4.3	44.7 ± 3.6	45.9 ± 3.2
9	Boys (n=15)	50.4 ± 7.9	48.5 ± 3.2	45.7 ± 3.7	45.8 ± 5.0	57.8 ± 2.2	49.2 ± 2.2	45.3 ± 5.7	44.9 ± 2.8	45.6 ± 5.6	46.6 ± 4.7
	Girls (n=15)	42.3 ± 7.3	46.3 ± 2.4	42.8 ± 3.0	39.9 ± 4.6	49.3 ± 4.5	42.4 ± 1.5	37.1 ± 3.3	36.2 ± 1.5	40.4 ± 4.1	42.2 ± 3.1
	Total (n=30)	46.4 ± 8.5	47.4 ± 3.0	44.2 ± 3.6	42.8 ± 5.6	53.5 ± 4.7	45.8 ± 3.9	41.2 ± 6.2	40.5 ± 5.0	43.0 ± 5.5	44.4 ± 4.5
10	Boys (n=15)	56.8 ± 8.4	48.9 ± 2.8	46.6 ± 2.8	47.0 ± 6.2	59.3 ± 1.9	50.0 ± 1.9	49.0 ± 5.3	47.0 ± 2.5	48.4 ± 4.7	48.7 ± 4.6
	Girls (n=15)	46.7 ± 7.3	46.1 ± 2.4	42.5 ± 3.1	39.9 ± 4.3	50.2 ± 1.7	42.5 ± 1.7	39.0 ± 4.0	37.2 ± 2.0	41.8 ± 4.0	43.1 ± 3.2
	Total (n=30)	51.8 ± 9.3	47.5 ± 2.9	44.5 ± 3.6	43.4 ± 6.4	54.7 ± 5.0	46.2 ± 4.2	44.0 ± 6.9	42.1 ± 5.4	45.1 ± 5.4	45.9 ± 4.8

Prediction Equations for VO_{2max}: Leger et al (1988): VO_{2max} = 31.025 + 3.238(Speed) - 3.248(Age) + 0.1536(Age)(Speed). Barnett et al (1993): 1) VO_{2max} = 28.3 - 2.1(sex: M=0; F=1) - 0.7(Tric_{Skinf}) + 2.6(MS); 2) VO_{2max} = 25.8 - 6.6(sex: M=0; F=1) - 0.2(weight) + 3.2(MS); 3) VO_{2max} = 24.2 - 5.0(sex: M=0; F=1) - 0.8(age) + 3.4(MS). Oliveira (1998): VO_{2max} = 25 + 6.6(sex: M=2; F=1) + 0.23(TL). Fernhall et al (1998): VO_{2max} = 0.35(TL) - 0.59(BMI) - 4.61(sex: M=1; F=2) + 50.6. Matsuzaka et al (2004): 1) VO_{2max} = 25.9 - 2.21(sex: M=0; F=1) - 0.0449(age) - 0.831(BMI) + 4.12(MS); 2) VO_{2max} = 61.1 - 2.20(sex: M=0; F=1) - 0.462(age) - 0.862(BMI) + 0.192(TL).

A deeper look into the distribution of the estimated VO_{2max} values among different ages, exposes the overestimating power of the equation nº 2 of Barnett et al (1993). This miscalculation was found at 8 (57.2 ± 1.4; 49.4 ± 0.9; 53.3 ± 4.1), 9 (57.8 ± 2.2; 49.3 ± 4.5; 53.5 ± 4.7) and 10 (59.3 ± 1.9; 50.2 ± 1.7; 54.7 ± 5.0) year old boys, girls and both together, respectively. On the contrary, the most pronounced underestimations were originated by the predictive equation of Oliveira (1998) also

simultaneously at 8 (44.3 ± 1.7 ; 36.3 ± 1.3 ; 40.3 ± 4.3), 9 (44.9 ± 2.8 ; 36.2 ± 1.5 ; 40.5 ± 5.0) and 10 (47.0 ± 2.5 ; 37.2 ± 2.0 ; 42.1 ± 5.4) year old boys, girls and both together, respectively.

Although nearly a single equation (FITNESSGRAM) emerged as the best predictor of VO_{2peak} when considering the entire sample or the split by sex, the same did not occurred when additionally splitting by age. In fact, FITNESSGRAM's software reports were only considered as the best estimators in 8 year old boys (49.3 ± 2.1) and in both genders with the same age (48.5 ± 2.0). The equation nº 3 of Barnett et al (1993) assumed a particular relevance in 9 year old subjects (boys: 49.2 ± 2.2 ; girls: 42.4 ± 1.5 ; both: 45.8 ± 3.9). At last, at the age of 10 the equation nº 2 of Barnett et al (1993) was the stronger predictor in boys (59.3 ± 1.9) and in both genders (54.7 ± 5.0). Girls' best predictive values were derived from FITNESSGRAM's software reports (46.1 ± 2.4). It is certainly important to refer that this estimations fail greatly in respect to the variability (\pm SD), in other words, the large dispersion of values in the direct measurement is not represented in any way by the correspondent estimated values.

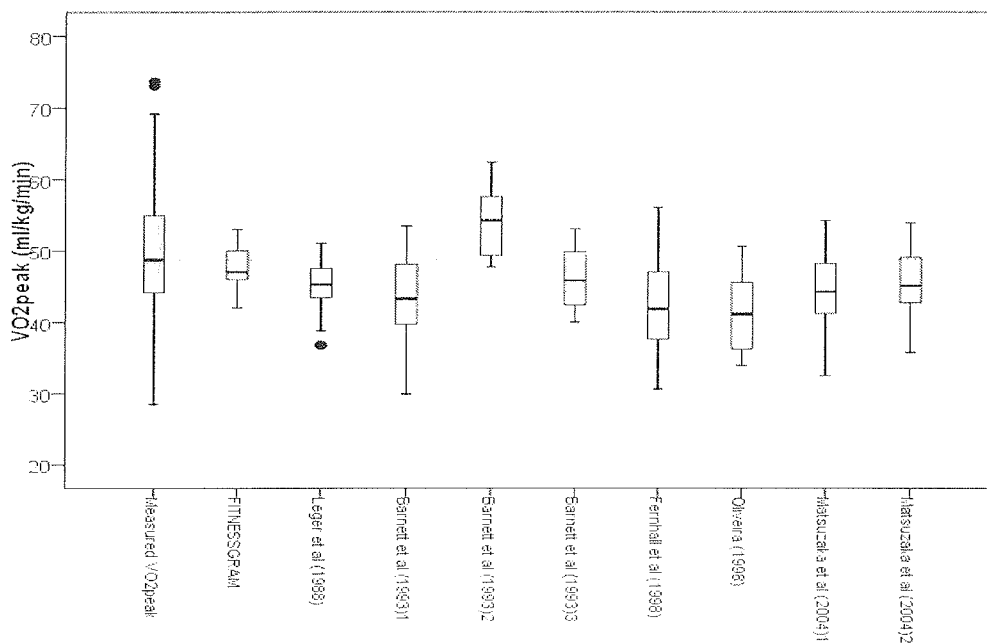


Figure 13: Boxplots of measured directly VO_{2peak} and estimated VO_{2max} from previously reported equations.

Special attention should be given to Figure 13. Thanks to the boxplots representations it is possible to see the wide distribution of the directly measured VO_{2peak} values. Even though the proximity between the median VO_{2max} lines of the measured VO_{2peak} and FITNESSGRAM is perceptible, the width of each box is totally distinctive. The most contracted boxes are indeed FITNESSGRAM's software reports and the equation scores of Leger et al (1988) [$VO_{2max} = 31.025 + 3.238(\text{Speed}) - 3.248(\text{Age}) + 0.1536(\text{Age})(\text{Speed})$]. This means that the estimated VO_{2max} values did not diverge much from the median. Although none of the boxplots may be roughly compared to the large disparity of values visible in measured VO_{2peak} , it is still well worth it to draw attention to the wider distribution of equations such as Fernhall et al (1998) [$VO_{2max} = 0.35(\text{TL}) - 0.59(\text{BMI}) - 4.61(\text{sex: M=1; F=2}) + 50.6$] and the equation nº 1 of Matsuzaka et al (2004) [$VO_{2max} = 25.9 - 2.21(\text{sex: M=0; F=1}) - 0.0449(\text{age}) - 0.831(\text{BMI}) + 4.12(\text{MS})$].

Table 29: Minimum, Maximum and Mean Differences (Mean \pm SD; ml.kg⁻¹.min⁻¹) Between Measured Directly VO_{2peak} and Estimated VO_{2max} from Previously Reported Equations.

	Boys (n = 45)	Girls (n = 45)	Total (n = 90)	Minimum	Maximum
FITNESSGRAM	4.4 \pm 7.5	-0.7 \pm 6.2	1.8 \pm 7.3	-18.50	21.06
Leger et al (1988)	6.9 \pm 7.4	2.5 \pm 5.7	4.7 \pm 6.9*	-10.27	23.35
Barnett et al (1993) ¹	6.8 \pm 7.2	5.4 \pm 4.9	6.1 \pm 6.1*	-11.53	21.66
Barnett et al (1993) ²	-4.8 \pm 7.8	-3.7 \pm 6.4	-4.2 \pm 7.1*	-19.26	13.92
Barnett et al (1993) ³	3.7 \pm 7.9	3.3 \pm 6.4	3.5 \pm 7.1**	-12.31	21.66
Fernhall et al (1998)	6.8 \pm 6.4	7.8 \pm 4.9	7.3 \pm 5.7*	-9.05	21.69
Oliveira (1998)	7.9 \pm 7.4	9.4 \pm 6.2	8.7 \pm 6.9*	-6.10	25.66
Matsuzaka et al (2004) ¹	6.3 \pm 6.9	4.4 \pm 4.7	5.4 \pm 5.9*	-11.41	21.70
Matsuzaka et al (2004) ²	5.6 \pm 7.1	2.9 \pm 5.0	4.2 \pm 6.3*	-12.67	20.57

* Difference between measured VO_{2peak} and each other group ($p < 0.05$); ** Difference between measured VO_{2peak} and each other group ($p < 0.01$).

Prediction Equations for VO_{2max}: Leger et al (1988): VO_{2max} = 31.025 + 3.238(Speed) - 3.248(Age) + 0.1536(Age)(Speed). Barnett et al (1993): 1) VO_{2max} = 28.3 - 2.1(sex: M=0; F=1) - 0.7(Tric_{skinf}) + 2.6(MS); 2) VO_{2max} = 25.8 - 6.6(sex: M=0; F=1) - 0.2(weight) + 3.2(MS); 3) VO_{2max} = 24.2 - 5.0(sex: M=0; F=1) - 0.8(age) + 3.4(MS). Oliveira (1998): VO_{2max} = 25 + 6.6(sex: M=2; F=1) + 0.23(TL). Fernhall et al (1998): VO_{2max} = 0.35(TL) - 0.59(BMI) - 4.61(sex: M=1; F=2) + 50.6. Matsuzaka et al (2004): 1) VO_{2max} = 25.9 - 2.21(sex: M=0; F=1) - 0.0449(age) - 0.831(BMI) + 4.12(MS); 2) VO_{2max} = 61.1 - 2.20(sex: M=0; F=1) - 0.462(age) - 0.862(BMI) + 0.192(TL).

Another perspective of analysis is to look at the mean differences (d ; ml.kg⁻¹.min⁻¹) between the measured directly VO_{2peak} and all other estimated VO_{2max} values from the previously reported equations (see Table 29). The higher d (boys: 6.8 \pm 7.2; girls: 5.4 \pm 4.9; both: 6.1 \pm 6.1) were all derived from the equation n° 1 of Barnett et al (1993). On the contrary, the lowest d were detected mostly on FITNESSGRAM's software reports (girls: -0.7 \pm 6.2; both: 1.8 \pm 7.3) and on the equation n° 3 of Barnett et al (1993) (boys: 3.7 \pm 7.9). There were only two estimating equations that overpredicted the VO_{2peak} scores which were: equation n° 2 of Barnett et al (1993) in boys (-4.8 \pm 7.8), girls (-3.7 \pm 6.4) and both (-4.2 \pm 7.1) and, at last, the VO_{2max} derived from FITNESSGRAM in girls (-0.7 \pm 6.2). When taking into consideration the whole group, the most extreme minimum difference value was found in the equation n° 2 of Barnett et al (1993) (-19.26 ml.kg⁻¹.min⁻¹) and the outermost maximum difference value was seen in the equation of Oliveira (1998) (25.66 ml.kg⁻¹.min⁻¹). On going with the analysis of the variable difference, Figure 14 draws gender based individual differences (n=90) between the measured directly VO_{2peak} and the estimated VO_{2max} from several 20-m MSRT VO_{2max} equations for children and youth.

From the 9 graphics below, only the equation n° 2 of Barnett et al (1993) (see Figure 14D) clearly overestimates VO_{2peak}. That is possible to see based on the negative pattern followed by the individual bars. In FITNESSGRAM's software reports (see Figure 14A), in the equation of Leger et al (1988) (see Figure 14B) and in the equation n° 3 of Barnett et al (1993) (see Figure 14E), it is also possible to verify a considerable number of overestimating individual bars, but with a inferior emphasis. On the contrary, the equation n° 2 of Barnett et al (1993) (see Figure 14D), the equation of Fernhall et al (1998) (see Figure 14F) and both equations of Matsuzaka et al (2004) (see Figure 14H-I) seem undoubtedly underestimators.

The analysis of the individual differences according to gender is attention-grabbing. In FITNESSGRAM's software reports (see Figure 14A) it is obvious a twofold pattern among boys and girls. Boys' VO_{2peak} values are clearly underestimated by the reports, whilst in girls the reports are pretty close within the \pm 9.0 ml.kg⁻¹.min⁻¹ correspondent to the measured directly VO_{2peak} SD. A relatively similar model can be found when considering the differences in the VO_{2max} values by the equation of Leger et al (1988) (see Figure 14B), and in a moderately faraway approach, when considering the differences in the VO_{2max} values by the equation of Matsuzaka et al (2004) (see Figure 14H).

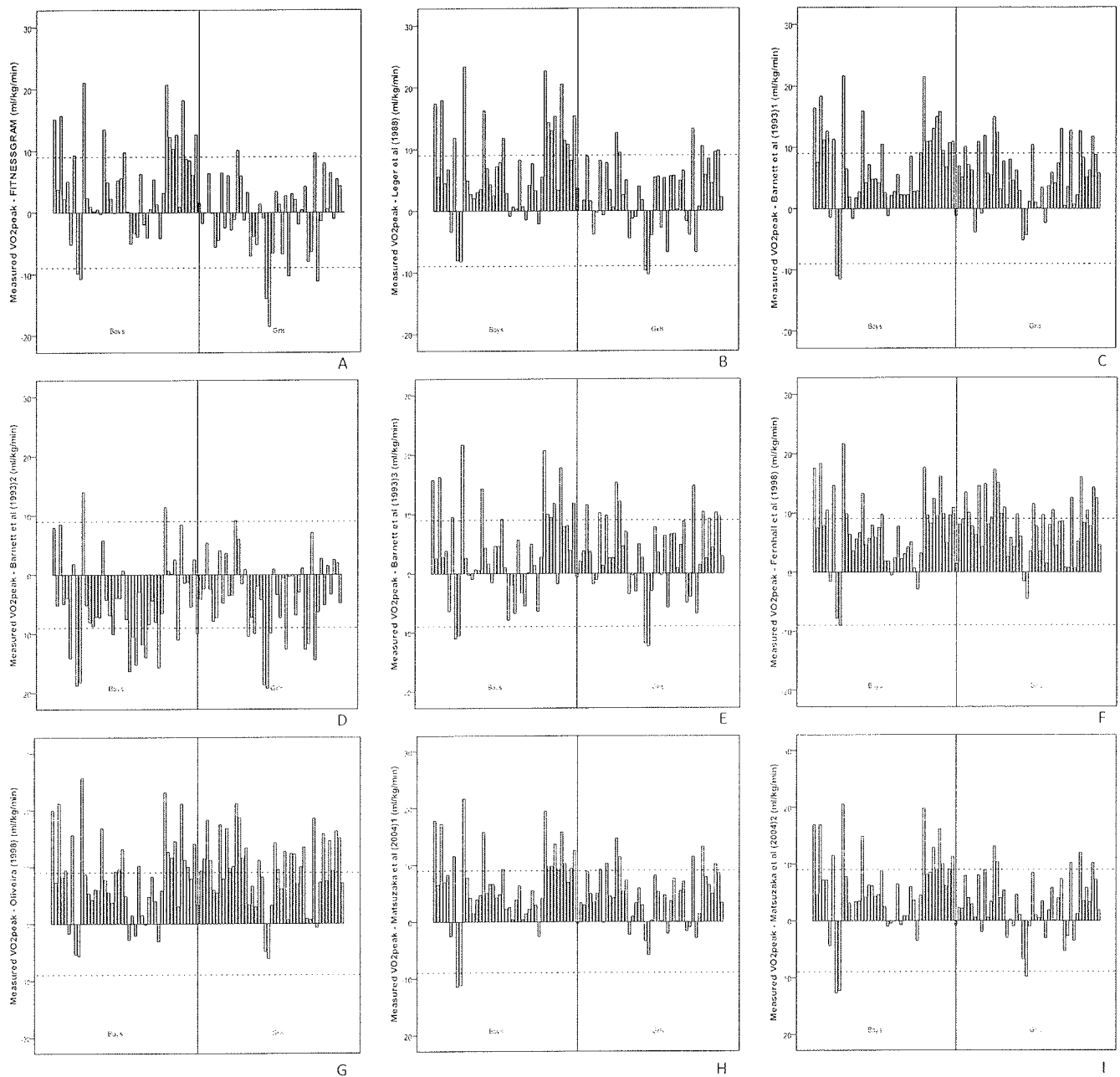


Figure 14: Gender based bar chart of the individual difference (n=90) between measured directly VO_{2peak} and estimated VO_{2max} from: A) FITNESSGRAM; B) Leger et al (1988); C) D) E) Barnett et al (1993); F) Fernhall et al (1998); G) Oliveira (1998) and H) I) Matsuzaka et al (2004). The two horizontal reference lines in each graphic represent the \pm SD of mean measured directly VO_{2peak} ($49.6 \pm 9.0 \text{ ml.kg}^{-1}.\text{min}^{-1}$).

One interesting approach is to observe the percentage of individuals whose difference scores (measured directly VO_{2peak} - equation VO_{2max}) fell inside the $9 \text{ ml.kg}^{-1}.\text{min}^{-1}$ correspondent to the measured VO_{2peak} SD ($49.6 \pm 9.0 \text{ ml.kg}^{-1}.\text{min}^{-1}$). Doing so, it is possible to access a certain degree of association. The higher percentages (77.8%) were found using FITNESSGRAM's software reports and the equation nº 2 of Matsuzaka et al (2004). A close percentage (75,6%) was obtained by the equation of Leger et al (1988), considerably distant from the equation of Oliveira (1998) which only 51.1% of the difference values fell inside the $9 \text{ ml.kg}^{-1}.\text{min}^{-1}$ of the measured VO_{2peak} SD.

Friedman test which is a non-parametric statistical procedure similar to the parametric repeated measures ANOVA, was used to detect differences across the previously reported 20-m MSRT VO_{2max} prediction equations and measured directly VO_{2peak} since the normality was not assumed in all the equations by the Shapiro-Wilk test ($p < 0.05$). As the Friedman test results pointed out to significant differences between the equations ($p < 0.001$), Dunnett's procedure of measured directly VO_{2peak}

values to each of the estimated VO_{2max} prediction equations was performed. This procedure is described in (Zar 1999), and was implemented in an *Excel 2007* spreadsheet as it is not available at the *SPSS Statistics 17.0*.

As we might see in Table 29, only FITNESSGRAM's software reports are not significantly different from measured directly VO_{2peak} values ($p>0.05$). Following this information, we went further on the intention to better illustrate the comparisons between measured directly VO_{2peak} and the VO_{2max} estimated from previously reported equations for children and youth and to create valid and comparable statistical data to that reported in previous researches, hoping to enrich the next chapter of the discussion of the results.

Table 30 resumes the Pearson product-moment correlations coefficients and the concordance correlation coefficients between measured directly VO_{2peak} and predicted VO_{2max} from previously reported equations. In boys, the highest Pearson product-moment correlation coefficient was found between the measured VO_{2peak} and the VO_{2max} estimated from the equation of Oliveira (1998) ($r=0.75$, $p<0.001$), while in girls it was seen between the measured VO_{2peak} and the VO_{2max} estimated from the equations of Fernhall et al (1998) and Matsuzaka et al (2004)² ($r=0.84$; $p<0.001$). When taking into account to the whole group, the highest correlations were found between the measured VO_{2peak} and the equation of Fernhall et al (1998) and Matsuzaka et al (2004)¹ ($r=0.79$; $p<0.001$).

Pearson product-moment correlations measure the strength of a relation between two variables, not the agreement between them. We have perfect agreement only if the points lie along the line of equality, but we will have perfect correlation if the points lie along any straight line. Therefore, it is not surprising to see that the concordance correlation coefficients (r_c) were comparatively inferior to the Pearson product-moment correlations (r). The correlations were no longer moderate to high as seen using the Pearson product-moment correlations but low or very low indeed when considering the concordance correlation. In boys, the highest concordance correlation coefficient was found between the measured VO_{2peak} and the VO_{2max} estimated from the equation of Fernhall et al (1998) ($r_c=0.44$), while in girls and when considering the whole group was seen between the measured VO_{2peak} and the VO_{2max} estimated from the equation nº 1 of Matsuzaka et al (2004) ($r_c=0.52$ for both). Whereas the highest Pearson product-moment correlation coefficient was seen in boys, the highest concordance correlation coefficient was seen in girls.

A curious observation is on the fact that both equations nº 1 and nº 2 of Matsuzaka et al (2004) and the equations nº 2 and nº 3 of Barnett et al (1993) present very similar Pearson product-moment correlations coefficients and concordance correlation coefficients for boys, girls and both.

Table 30: Pearson Product-Moment Correlations and Concordance Correlation Between Measured Directly VO_{2peak} and Estimated VO_{2max} from Previously Reported Equations.

	Boys (n = 45)		Girls (n = 45)		Both (n = 90)	
	<i>r</i>	<i>r_c</i>	<i>r</i>	<i>r_c</i>	<i>r</i>	<i>r_c</i>
FITNESSGRAM	0.70	0.31	0.61	0.34	0.72	0.38
Leger et al (1988)	0.71	0.27	0.68	0.42	0.75	0.39
Barnett et al (1993) ¹	0.63	0.37	0.76	0.46	0.74	0.49
Barnett et al (1993) ²	0.74	0.24	0.63	0.19	0.62	0.43
Barnett et al (1993) ³	0.74	0.25	0.68	0.22	0.65	0.41
Fernhall et al (1998)	0.67	0.44	0.84	0.32	0.79	0.50
Oliveira (1998)	0.75	0.23	0.69	0.11	0.66	0.32
Matsuzaka et al (2004) ¹	0.69	0.39	0.83	0.52	0.79	0.52
Matsuzaka et al (2004) ²	0.67	0.38	0.84	0.52	0.78	0.51

Prediction Equations for VO_{2max} : Leger et al (1988): $VO_{2max} = 31.025 + 3.238(\text{Speed}) - 3.248(\text{Age}) + 0.1536(\text{Age})(\text{Speed})$. Barnett et al (1993): 1) $VO_{2max} = 28.3 - 2.1(\text{sex: M=0; F=1}) - 0.7(\text{Tric}_{\text{skinf}}) + 2.6(\text{MS})$; 2) $VO_{2max} = 25.8 - 6.6(\text{sex: M=0; F=1}) - 0.2(\text{weight}) + 3.2(\text{MS})$; 3) $VO_{2max} = 24.2 - 5.0(\text{sex: M=0; F=1}) - 0.8(\text{age}) + 3.4(\text{MS})$. Oliveira (1998): $VO_{2max} = 25 + 6.6(\text{sex: M=2; F=1}) + 0.23(\text{TL})$. Fernhall et al (1998): $VO_{2max} = 0.35(\text{TL}) - 0.59(\text{BMI}) - 4.61(\text{sex: M=1; F=2}) + 50.6$. Matsuzaka et al (2004): 1) $VO_{2max} = 25.9 - 2.21(\text{sex: M=0; F=1}) - 0.0449(\text{age}) - 0.831(\text{BMI}) + 4.12(\text{MS})$; 2) $VO_{2max} = 61.1 - 2.20(\text{sex: M=0; F=1}) - 0.462(\text{age}) - 0.862(\text{BMI}) + 0.192(\text{TL})$.

A simple plot of the results of one equation against those of the VO_{2peak} though without a regression line is a useful start but usually the data points will be clustered near the line and it will be difficult to assess between-method differences. If we want to know by how much a prediction equation is likely to differ from the measured directly VO_{2peak} the plot of difference

against the mean is more informative (Bland and Altman 1986). It would be a mistake to plot the difference against either value separately because the difference would be related to each, a well-known statistical artefact (Gill JS, Zezulka AV et al. 1985), even though in this study it was indifferent to a certain extent the use of the mean or the true value as we can see in Figure 15.

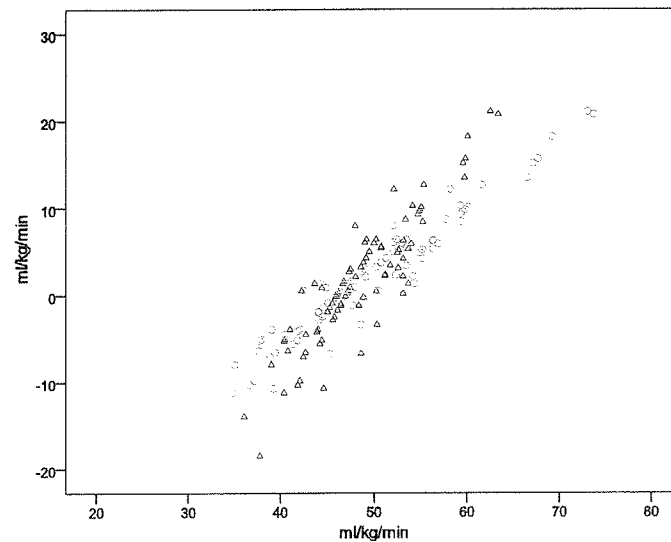


Figure 15: Scatterplot of mean difference (measured VO_{2peak} - FITNESSGRAM; Y-axis) against the average value (triangles; mean: measured VO_{2peak} , FITNESSGRAM; X-axis) and mean difference (measured VO_{2peak} - FITNESSGRAM; Y-axis) against measured directly VO_{2peak} (circles) (X-axis).

As follows, using the Bland-Altman approach (Bland and Altman 1986) (see Figure 16), the limits of agreement (average difference ± 1.96 SD of the difference; $ml.kg^{-1}.min^{-1}$) and the slope of the mean difference (d ; Y-axis; $ml.kg^{-1}.min^{-1}$) against the average value of measured VO_{2peak} and predicted VO_{2max} (X-axis; $ml.kg^{-1}.min^{-1}$) were examined. Figure 16 helps the eye in gauging the degree of agreement between measurements. In each graphic, three reference lines are drawn representing the mean difference of the two measurements (d) and the limits of agreement (LA). The LA ($ml.kg^{-1}.min^{-1}$) are only estimates of the values which apply to the whole population and in fact, the graphics display considerable lack of agreement between the measured directly VO_{2peak} and the VO_{2max} estimated from the previously reported equations, with discrepancies of up to $22 ml.kg^{-1}.min^{-1}$ (Oliveira 1998) (see Figure 16G).

If the differences are within $d \pm 1.96SD$ there is no significant difference between measured directly VO_{2peak} value and the estimated VO_{2max} from the considered reported equation. However, in this study these intervals ($d - 1.96SD$ and $d + 1.96SD$) are wide, reflecting the great variation of the differences and all graphics exhibit several differences outside the LA.

It is well noticeable a positive linear tendency of the distribution of the values in FITNESSGRAM's software reports (see Figure 16A), in the equation of Leger et al (1988) (see Figure 16B) and in the equation nº 2 of Matsuzaka et al (2004) (see Figure 16D). On the contrary, a wider dispersion of the differences are visible in the equation nº 1 of Barnett et al (1993) (see Figure 16C), in the equation of Fernhall et al (1998) (see Figure 16F) and in the equation nº 1 of Matsuzaka et al (2004) (see Figure 16H). Essentially the equations overpredicted VO_{2peak} particularly for aerobically unfit children, and underpredicted VO_{2peak} in the fit children.

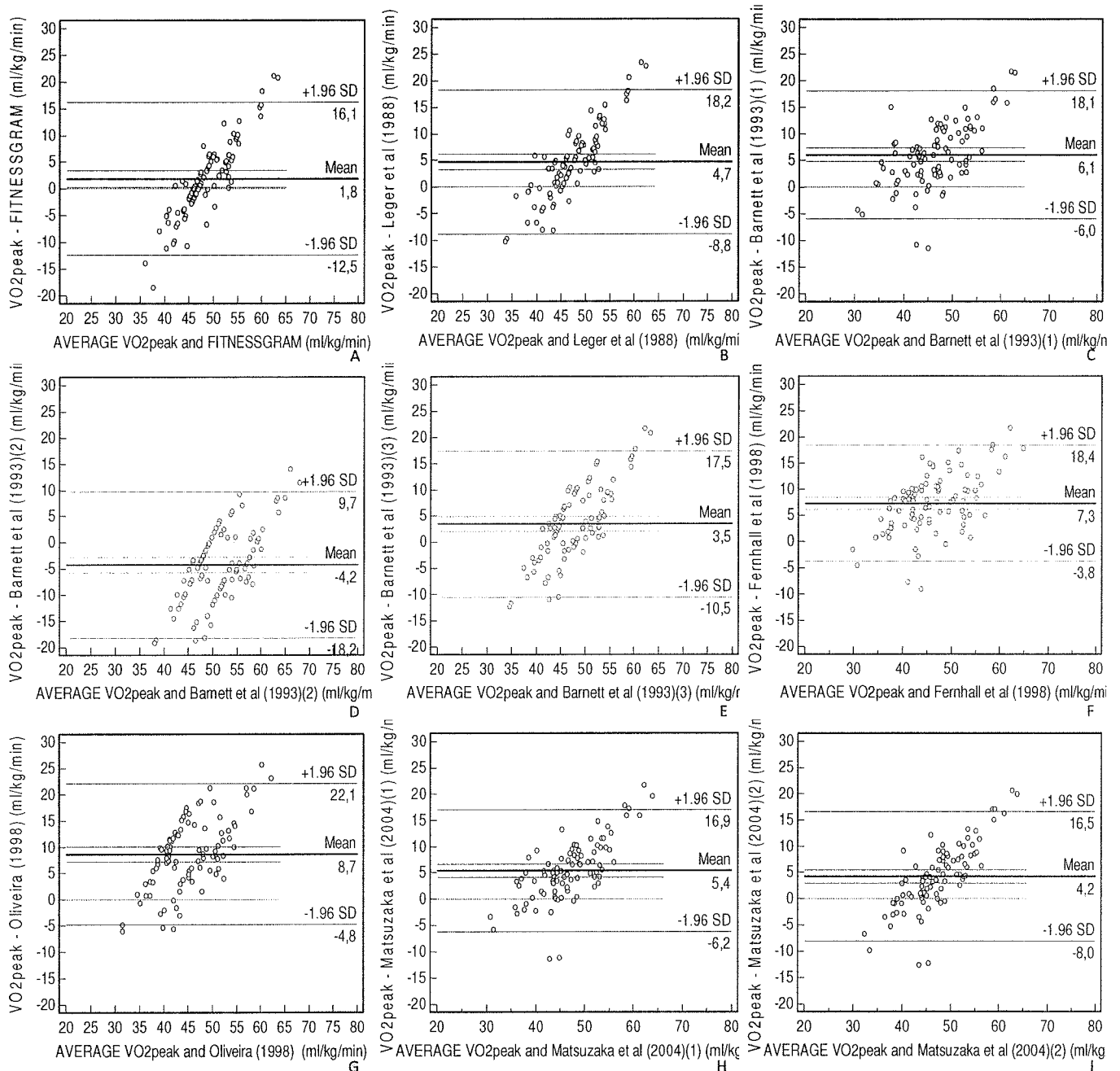


Figure 16: Bland-Altman plot of measured directly VO_{2peak} and the estimated VO_{2max} from: A) FITNESSGRAM; B) Leger et al (1988); C) D) E) Barnett et al (1993); F) Fernhall et al (1998); G) Oliveira (1998) and H) I) Matsuzaka et al (2004).

As previously stated the FITNESSGRAM's software reports yielded the smallest d ($ml \cdot kg^{-1} \cdot min^{-1}$) between measured directly and estimated VO_{2peak} values (1.84). Even so, it presented the highest range of values between upper and lower LA (28.6) and the highest positive slope (1.20) meaning that the equation overpredicted VO_{2peak} in unfit participants and underpredicted VO_{2peak} in fit participants. The lowest slope (0.44) was seen between the measured directly VO_{2peak} and the equation of Fernhall et al (1998) as well as the least lower LA (-3.80) concurring consequently to the lowest range ($ml \cdot kg^{-1} \cdot min^{-1}$) between upper and lower LA (22.23). The lowest upper LA was found between the measured directly VO_{2peak} and the equation nº 2 of Barnett et al (1993) (9.73) but was seriously influenced by the highest observed lower LA (-18.21), also observed between the measured directly VO_{2peak} and the equation nº 2 of Barnett et al (1993), which resulted in one of the highest range (27.94) between upper and lower LA. The appraisal of these values show that this is in fact a equation that tends to overpredict the VO_{2peak} values, confirmed by the d value in Table 31 (-4.24). The comparison between the measured directly VO_{2peak} and the equation of Leger et al (1988) provided a reasonably large d (4.69), one of the highest range (27.08) between the upper (18.23) and lower (-8.85) LA and the second uppermost slope (1.04). The equation of Matsuzaka et al (2004) that included MS (1; maximal speed) showed a higher d (5.38) than the other equation of Matsuzaka et al (2004) using TL (2; total laps)

($d=4.23$) when face up against the measured directly VO_{2peak} . However, the first demonstrated a smaller range (23.14) between upper (16.95) and lower (-6.19) LA, and a gentle slope (0.64). When some attention is afforded to the three equations of Barnett et al (1993), the equation that includes the $Tric_{skinf}$ and maximal speed (MS) showed the highest d to the measured directly VO_{2peak} (6.07) but the smallest range (24.09) between upper (18.12) and lower (-5.97) LA and the less steeper slope (0.57) (see Table 31).

Table 31: Bland-Altman Approach to Compare Measured Directly VO_{2peak} ($ml.kg^{-1}.min^{-1}$) and Estimated VO_{2max} ($ml.kg^{-1}.min^{-1}$) from Previously Reported Equations.

	d	LA		Range	Slope
FITNESSGRAM	1.84	-12.46	16.14	28.60	1.20
Leger et al (1988)	4.69	-8.85	18.23	27.08	1.04
Barnett et al (1993) ¹	6.07	-5.97	18.12	24.09	0.57
Barnett et al (1993) ²	-4.24	-18.21	9.73	27.94	0.78
Barnett et al (1993) ³	3.49	-10.48	17.45	27.93	0.95
Fernhall et al (1998)	7.32	-3.80	18.43	22.23	0.44
Oliveira (1998)	8.65	-4.77	22.09	26.86	0.69
Matsuzaka et al (2004) ¹	5.38	-6.19	16.95	23.14	0.64
Matsuzaka et al (2004) ²	4.23	-8.05	16.50	24.55	0.79

Mean differences (d), Limits of Agreement (LA) and Range (upper LA - lower LA) are reported in $ml.kg^{-1}.min^{-1}$.

Prediction Equations for VO_{2max} : Leger et al (1988): $VO_{2max} = 31.025 + 3.238(\text{Speed}) - 3.248(\text{Age}) + 0.1536(\text{Age})(\text{Speed})$. Barnett et al (1993): 1) $VO_{2max} = 28.3 - 2.1(\text{sex: M=0; F=1}) - 0.7(\text{Tric}_{skinf}) + 2.6(\text{MS})$; 2) $VO_{2max} = 25.8 - 6.6(\text{sex: M=0; F=1}) - 0.2(\text{weight}) + 3.2(\text{MS})$; 3) $VO_{2max} = 24.2 - 5.0(\text{sex: M=0; F=1}) - 0.8(\text{age}) + 3.4(\text{MS})$. Oliveira (1998): $VO_{2max} = 25 + 6.6(\text{sex: M=2; F=1}) + 0.23(\text{TL})$. Fernhall et al (1998): $VO_{2max} = 0.35(\text{TL}) - 0.59(\text{BMI}) - 4.61(\text{sex: M=1; F=2}) + 50.6$. Matsuzaka et al (2004): 1) $VO_{2max} = 25.9 - 2.21(\text{sex: M=0; F=1}) - 0.0449(\text{age}) - 0.831(\text{BMI}) + 4.12(\text{MS})$; 2) $VO_{2max} = 61.1 - 2.20(\text{sex: M=0; F=1}) - 0.462(\text{age}) - 0.862(\text{BMI}) + 0.192(\text{TL})$.

Once the Bland-Altman approach is performed based on gender (see Table 32), a not so coincident pattern is drawn to that of the sample considered as a whole. The FITNESSGRAM's software reports still yielded the smallest d ($ml.kg^{-1}.min^{-1}$) between measured directly and estimated VO_{2peak} values (-0.71) but only in girls. It did not presented the highest slope as previously reported to the entire sample, but it was a high value as well (1.24) meaning that the equation overpredicted VO_{2peak} in the most unfit girls and underpredicted VO_{2peak} in the fittest girls, although the mean values were rather matching. The lowest slope (0.63) was seen between measured directly VO_{2peak} and the equation nº 3 of Barnett et al (1993), though with a moderate d (5.39), but one of the lowest range between upper and lower LA (19.28). The equation of Leger et al (1988) had higher d (2.54) when compared to the one obtained between the measured directly VO_{2peak} and FITNESSGRAM's software reports (-0.71), but lower range between LA (22.30) and lower slope (0.98).

In boys the lowest d (3.71) was seen in the equation nº 3 of Barnett et al (1993), 4.42 $ml.kg^{-1}.min^{-1}$ higher than the lowest d in girls (FITNESSGRAM), but had also the highest range (30.91) between upper and lower LA and the highest slope (1.43). The lowest slope (0.65) was seen between the measured directly VO_{2peak} and the equation of Fernhall et al (1998) as well as the lowest range between upper and lower LA (25.02). The equation of Leger et al (1988) had higher d (6.84) when compared to the one obtained between measured VO_{2peak} and FITNESSGRAM's software reports (4.39), but lower range between LA (28.98) and lower slope (1.16), just as similar to the pattern reported in girls.

Table 32: Gender Based Bland-Altman Approach to Compare Measured Directly VO_{2peak} ($ml.kg^{-1}.min^{-1}$) and Estimated VO_{2max} ($ml.kg^{-1}.min^{-1}$) from Previously Reported Equations.

	Boys					Girls				
	<i>d</i>	LA		Range	Slope	<i>d</i>	LA		Range	Slope
FITNESSGRAM	4.39	-10.32	19.11	29.43	1.22	-0.71	-12.80	11.38	24.18	1.24
Leger et al (1988)	6.84	-7.65	21.33	28.98	1.16	2.54	-8.61	13.69	22.30	0.98
Barnett et al (1993) ¹	-4.81	-20.07	10.45	30.52	1.38	-3.67	-16.30	8.96	25.26	1.49
Barnett et al (1993) ²	6.76	-7.27	20.79	28.06	0.72	5.39	-4.25	15.03	19.28	0.63
Barnett et al (1993) ³	3.71	-11.74	19.17	30.91	1.43	3.26	-9.20	15.72	24.92	1.46
Fernhall et al (1998)	6.79	-5.71	19.31	25.02	0.65	7.84	-1.70	17.38	19.08	0.78
Oliveira (1998)	7.90	-6.64	22.44	29.08	1.22	9.41	-2.80	21.62	24.42	1.39
Matsuzaka et al (2004) ¹	6.33	-7.10	19.77	26.87	0.78	4.42	-4.69	13.53	18.22	0.69
Matsuzaka et al (2004) ²	5.59	-8.35	19.53	27.88	0.89	2.87	-6.93	12.67	19.60	0.88

Mean differences (*d*), Limits of Agreement (LA) and Range (upper LA - lower LA) are reported in $ml.kg^{-1}.min^{-1}$.

Prediction Equations for VO_{2max} : Leger et al (1988): $VO_{2max} = 31.025 + 3.238(\text{Speed}) - 3.248(\text{Age}) + 0.1536(\text{Age})(\text{Speed})$. Barnett et al (1993): 1) $VO_{2max} = 28.3 - 2.1(\text{sex: M=0; F=1}) - 0.7(\text{Tric}_{skini}) + 2.6(\text{MS})$; 2) $VO_{2max} = 25.8 - 6.6(\text{sex: M=0; F=1}) - 0.2(\text{weight}) + 3.2(\text{MS})$; 3) $VO_{2max} = 24.2 - 5.0(\text{sex: M=0; F=1}) - 0.8(\text{age}) + 3.4(\text{MS})$. Oliveira (1998): $VO_{2max} = 25 + 6.6(\text{sex: M=2; F=1}) + 0.23(\text{TL})$. Fernhall et al (1998): $VO_{2max} = 0.35(\text{TL}) - 0.59(\text{BMI}) - 4.61(\text{sex: M=1; F=2}) + 50.6$. Matsuzaka et al (2004): 1) $VO_{2max} = 25.9 - 2.21(\text{sex: M=0; F=1}) - 0.0449(\text{age}) - 0.831(\text{BMI}) + 4.12(\text{MS})$; 2) $VO_{2max} = 61.1 - 2.20(\text{sex: M=0; F=1}) - 0.462(\text{age}) - 0.862(\text{BMI}) + 0.192(\text{TL})$.

RELIABILITY OF THE 20-M MSRT IN THE PRESENT STUDY

In Figure 17A it is possible to see a reasonably large difference between test and retest medians (52.05 vs. 45.75, respectively; $ml.kg^{-1}.min^{-1}$) and quartiles ($Q_1=43.85$; $Q_3=56.00$ vs. $Q_1=44.77$; $Q_3=60.96$; respectively; $ml.kg^{-1}.min^{-1}$) of the measured directly VO_{2peak} ($ml.kg^{-1}.min^{-1}$). A larger spectrum of maximum values (59.81 vs. 69.27, respectively; $ml.kg^{-1}.min^{-1}$) and a similarity between minimum values (28.49 vs. 28.38, respectively; $ml.kg^{-1}.min^{-1}$) was also observed. A different pattern was manifested when analysing the total number of laps (see Figure 17B) since both test and retest boxplots were very similar in respect to medians (26.50 vs. 24.00, respectively), quartiles ($Q_1=19.25$; $Q_3=34.25$ vs. $Q_1=18.00$; $Q_3=31.75$), maximum (13 vs. 10, respectively) and minimum (44 vs. 42, respectively) laps attained.

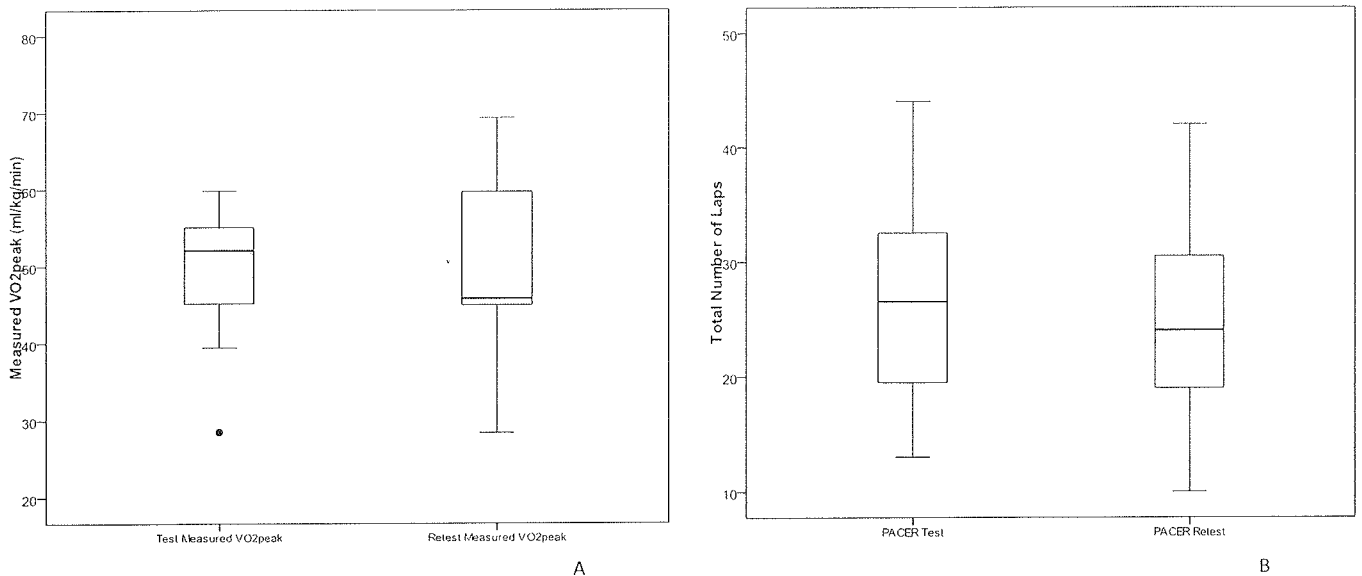


Figure 17: Boxplots representations of measured directly VO_{2peak} values (A) and total number of laps (B) in test-retest trials.

In the present study, the reliability of the PACER test was found to be high ($r=0.82$). In respect to the correlation among the two trials, high Pearson product-moment correlation coefficients (r) were obtained between total number of laps ($r=0.97$), VO_{2peak} scores ($r=0.82$) and HR_{max} ($r=0.70$). Yet, when looking to the concordance correlation (r_c), the correlation between VO_{2peak} scores was slightly inferior ($r_c=0.78$), as well as between total number of laps ($r_c=0.95$), HR_{max} ($r_c=0.68$) and even RER

($r_c=0.16$ vs. $r=0.18$) when each variable was compared to their r counterpart. Figure 18 is an illustration of the concordance correlation between VO_{2peak} , TL, RER and HR_{max} in test and retest trials. It is evident the high agreement between values and the 45° line. Only RER presents a spreading pattern proper of a low r_c .

As it is possible to see in Table 33, peak oxygen uptake, RER, total number of laps and HR_{max} were all considered not significantly different between trials ($p>0.05$) accordingly to paired student's t -test (total comparisons) and Wilcoxon non-parametric tests (gender comparisons). When considering VO_{2peak} , the p -values were very high, specifically in girls ($p=1.00$). On the contrary, the lowest p -values were found in girls' RER (test: 1.01 ± 0.1 vs. retest: 0.90 ± 0.1 ; $p=0.13$) and TL (test: 26.1 ± 9.4 vs. retest: 24.9 ± 8.6 ; $p=0.13$).

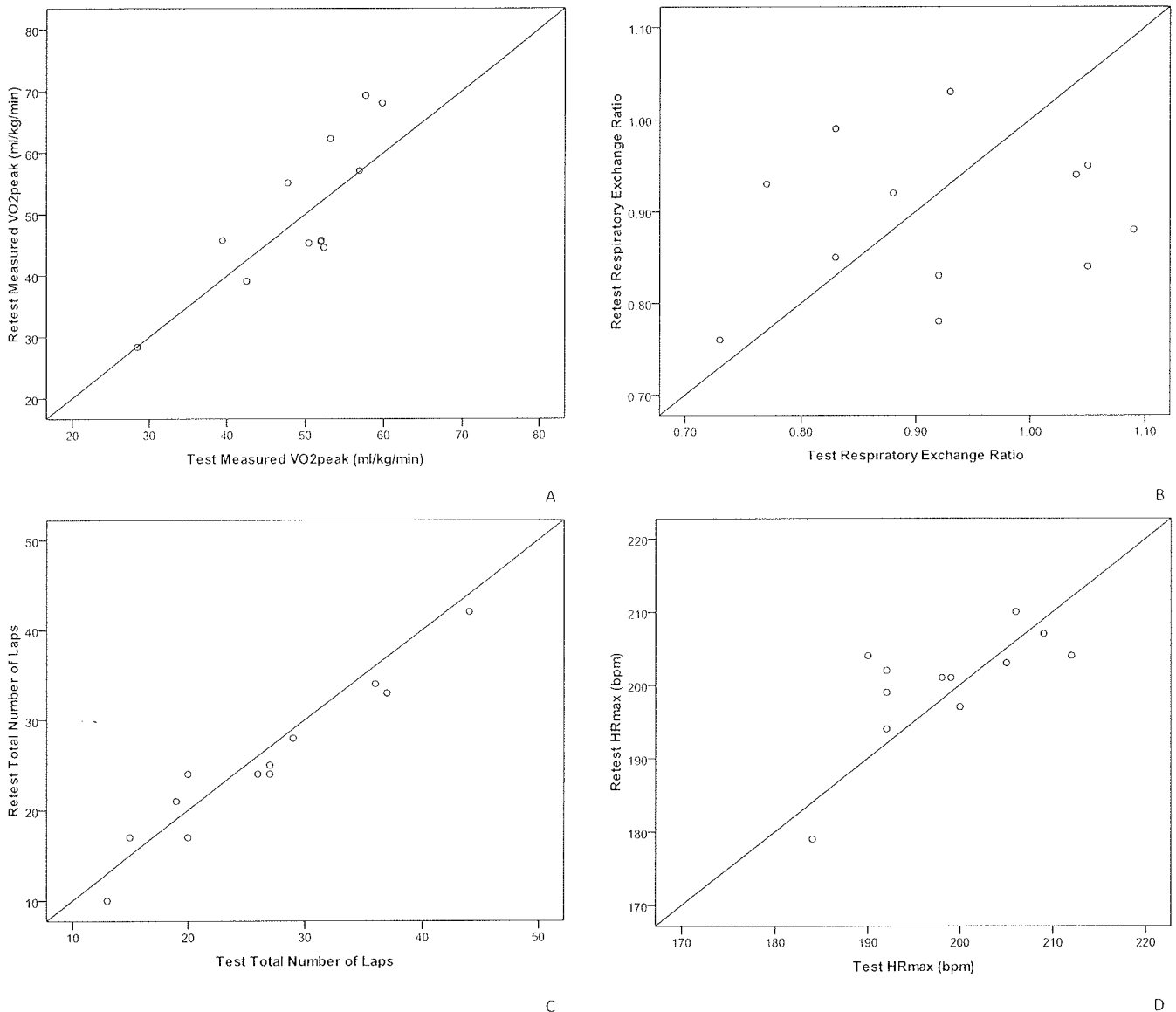


Figure 18: Scatterplots of Test against Retest of: A) measured directly VO_{2peak} ; B) RER; C) Total Number of Laps; D) HR_{max} .

Of particular interest is the fact that although the mean VO_{2peak} retest values in boys, girls and both were higher than the previous trial, this pattern was not followed by the mean number of laps, even though the difference was not significant (boys: $p=0.16$; girls: $p=0.75$; both: $p=0.13$). This is quite extraordinary in the way that when considering the previous reported 20-m MSRT prediction VO_{2peak} equations, it is assumed that when the performance increases the VO_{2peak} increases as well.

Table 33: Gender Based Spirometric (VO_{2peak} ; $ml.kg^{-1}.min^{-1}$ and RER_{peak}), Performance (Total laps; n°) and Cardiovascular (HR_{max} ; bpm) Outputs in Test and Retest Trials (Mean \pm SD); Pearson Product-Moment Correlation (r) and Concordance Correlation (r_c) Between Test and Retest Trials.

	Gender	TEST	RETEST	r	r_c
VO_{2peak}	Boys (n = 6)	54.2 \pm 3.7	56.0 \pm 3.7		
	Girls (n = 6)	44.6 \pm 10.1	45.0 \pm 10.6		
	Total (n=12)	49.4 \pm 8.8	50.5 \pm 12.1	0.82	0.78
RER_{peak}	Boys (n = 6)	0.83 \pm 0.1	0.88 \pm 0.0		
	Girls (n = 6)	1.01 \pm 0.1	0.90 \pm 0.1		
	Total (n=12)	0.92 \pm 0.1	0.89 \pm 0.1	0.18	0.16
Total laps	Boys (n = 6)	31.5 \pm 9.1	29.7 \pm 8.0		
	Girls (n = 6)	20.7 \pm 6.4	20.0 \pm 6.7		
	Total (n=12)	26.1 \pm 9.4	24.9 \pm 8.6	0.97	0.95
HR_{max}	Boys (n = 6)	202.7 \pm 8.0	203.8 \pm 4.5		
	Girls (n = 6)	193.4 \pm 7.1	196.3 \pm 9.1		
	Total (n=12)	198.3 \pm 8.5	200.1 \pm 7.9	0.70	0.68

* Difference between the means ($p < 0.05$); ** Difference between the means ($p < 0.001$).

DISCUSSION

The aim of this study was to compare the validity of the FITNESSGRAM's MSRT software reports and previously reported equations as predictors of CVF in Lisbon healthy elementary school children, by measuring the VO_{2peak} during the actual 20-m MSRT performance. Ninety subjects, 45 boys and 45 girls aged 8-10 years (in a balanced experimental design), performed the MSRT and their VO_{2peak} was measured directly in field using a portable gas analyzer (K4b², Cosmed, Rome, Italy).

We additionally assessed the reliability of the 20-m MSRT in a sub-group of subjects that performed the test twice.

ANTHROPOMETRIC, SPIROMETRIC AND PERFORMANCE VARIABLES

The anthropometric characteristics of the subjects in this study, in particular the BMI, were closely related to the values reported by Pitetti et al (2002) in boys (17.7 ± 3.1 vs. 17.7 ± 1.9) and in girls (18.3 ± 2.5 vs. 18.4 ± 3.4), even though the age range of the authors' sample was bigger (8-15 years old). Taking into consideration the ages of 8 to 10 years old, we come across with the study of Matsuzaka et al (2004) who reported comparable BMI values in boys (17.0 ± 2.1) and in girls (17.4 ± 2.2).

Boys' and girls' mean HR_{max} values obtained in the present study (195.1 ± 10.1 ; 195.9 ± 8.0 , respectively) were similar to the ones reported by Matsuzaka et al (2004) (196.5 ± 7.0 ; 194.7 ± 9.5 , respectively) in subjects of the same age (8-10 years old) and to Liu et al (1992) (194.8 ± 14.6 ; 195.5 ± 10.3 ; respectively) in subjects a little older (12-15 years old).

To verify that participants had reached their VO_{2peak} in the PACER test, the criterions of subjective symptoms, RER_{peak} and HR_{peak} were used. Although a $VO_{2plateau}$ has been generally recognized as the most reliable criterion for adults, it was not adopted in this study because, as shown in previous reports, the $VO_{2plateau}$ is not necessary as a criterion in children and adolescents (Rivera-Brown, Alvarez et al. 2001). In fact, these authors found that a high proportion of adolescents exhibit subjective and objective indicators of maximal performance without showing a plateau in VO_2 . Also, they have summarized previous studies reporting that the frequency of achievement of the plateau in adolescents ranged from 8-56%. Because the participants who did not meet at least two of the other criteria in this study would be excluded, it suggests that the VO_{2peak} obtained in this study is valid.

Motivation and rhythm did not seem to constitute a setback as it is evident from the high mean HR_{max} (195.5 ± 9.0), and high mean RER_{peak} (0.94 ± 0.1) registered during the test. This information is particularly relevant when it is nowadays acknowledged that the speed increment and the verbal information truly have an impact in the terminus of the test. Tests with sudden and progressive changes in the initial speed at each new stage, may encourage the participants to prematurely put an end to the test in order to avoid an undesired increase of speed when near the exhaustion breaking point (Wilkinson, Fallowfield et al. 1999).

The mean VO_{2peak} was 49.6 ± 9.0 ml.kg⁻¹.min⁻¹ in conformity with the ranging values of 45-60 ml.kg⁻¹.min⁻¹ found in literature for school age children (Krahenbuhl and Pangrazi 1983). Not surprisingly, boys had highest mean absolute and relative VO_{2peak} values than girls (1.7 ± 0.5 ; 53.3 ± 9.2 vs. 1.4 ± 0.4 ; 46.0 ± 7.3 , respectively) and the values were significantly different ($p < 0.001$). In the study of Rutenfranz et al (1982) boys exhibited as well clearly superior exercise fitness during all years of childhood, this being in contrast to the widely accepted concept that no sex difference exists in exercise fitness before puberty.

Accordingly to G.S Krahenbuhl et al (1989) if we compare boys' and girls' VO_{2peak} (l.min⁻¹) at various chronological ages, data suggest that boys' values are about 12% higher than girls' at 10 years and that the sex difference becomes more pronounced during the teen years, typically reaching 37% at age 16 years. In the present study, the results suggest that girls' VO_{2peak} (l.min⁻¹) values are 21% inferior to the obtained in boys. Gender differences in body composition (Mota, Guerra et al. 2002), levels of physical activity (Anderssen, Cooper et al. 2007) and anaerobic capacity (Rivera-Brown, Alvarez et al. 2001) may partially explain these variations.

Although the mean measured directly $\text{VO}_{2\text{peak}}$ ($\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) were apparently high, particularly in boys (53.3 ± 9.2), there are similar scores reported in previous researches. In three validation studies of the 20-m MSRT in children and youth from the Netherlands, China and Japan, van Mechelen et al (1986), Barnett et al (1993) and Matsuzaka et al (2004) reported analogous scores in boys (53.2 ± 5.4 ; 53.9 ± 5.6 ; 52.1 ± 6.2 , respectively). In girls, Barnett et al (1993) and Matsuzaka et al (2004) once more reported $\text{VO}_{2\text{peak}}$ values (45.1 ± 4.6 ; 49.3 ± 5.5 , respectively) comparable to the ones obtained in the present study (46.0 ± 7.3).

However, not only in validation studies of the 20-m MSRT in children and youth it is possible to gather comparable $\text{VO}_{2\text{peak}}$ data. Massicotte et al (1985) in an attempt to predict $\text{VO}_{2\text{max}}$ from the running performance in children aged 10-17 years obtained 52.7 to 57.4 $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ in boys and 41.9 to 48.3 $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ using a cycle ergometer VO_2 test criterion. Although reporting inferior values most evident in boys, Cureton et al (1995) had yet related values to the present study (boys: 49.6; girls: 45.4) using a treadmill VO_2 test criterion in order to assess the determinants of distance running performance in 8-17 year old children.

These results are unmistakably in line with the conclusions of Ahmaidi et al (1992) and Flouris et al (2005) in which the 20-m MSRT was considered clearly less economical than the progressive, continuous running in track. This may explain in part the smaller resemblances between the $\text{VO}_{2\text{peak}}$ measured directly during the actual 20-m MSRT and the running treadmill protocols. In some 20-m MSRT validation studies for instance, the wind resistance is simulated in the treadmill by introducing a 3% grade, since it is well proved that the VO_2 is dependent of the air resistance (Berthoin, Gerbeaux et al. 1994). The same can be applied to the 20-m MSRT validation studies performed in cycle ergometer. Leger et al (1988) for example reported manifestly inferior $\text{VO}_{2\text{peak}}$ scores for 8-11 year old Canadian boys and girls (40.5 ± 7.1 ; 41.3 ± 9.0) using a cycle ergometer VO_2 test criterion which is well known for being centrally less demanding (Washington, Bricker et al. 1994).

VALIDITY OF THE 20-M MSRT AS A PREDICTOR OF $\text{VO}_{2\text{PEAK}}$ IN LISBON ELEMENTARY SCHOOL CHILDREN

The entire sample ($n=90$) produced high significant validity scores for the 20-m MSRT ($r=0.77$). The enlargement of the variance of the $\text{VO}_{2\text{peak}}$ values when both genders are analysed as a whole group causes the strengthening of the linear dependence between two variables (Barnett, C. et al. 1993). Similar scores were seen in boys ($r=0.75$) and slightly inferior in girls ($r=0.69$). The high correlations between the PACER test and $\text{VO}_{2\text{peak}}$ suggest that greater than 56% of the variance for boys and 48% of that for girls can be explained by the performance (total laps) in the 20-m MSRT. These values are both inferior to the values obtained by Mahoney (1992) in a multiracial UK population of 12 year olds ($r=0.83$, boys; $r=0.76$, girls), however slightly superior or even identical to the values reported by van Mechelen et al (1986) in 12-14 year old Dutch boys ($r=0.68$), girls ($r=0.69$) and when considering the entire sample ($r=0.76$). The results indicate that the 20-m MSRT (PACER test) is a valid tool for the evaluation of $\text{VO}_{2\text{max}}$ in Lisbon healthy elementary school children and we express no reservation regarding the use of this test as a predictor of $\text{VO}_{2\text{max}}$. This is emphasized if one considers the homogeneity of our population with regard to age and body build.

Distance (total laps) rather than maximal speed (MS; $\text{km}\cdot\text{h}^{-1}$) provided a better discriminatory measure of aerobic performance ($r=0.77$ vs. $r=0.75$) in line with the findings of Liu et al (1992), Matsuzaka et al (2004) and Pitetti et al (2002). Falgairette et al (1994), Leger et al (1988) and McVeigh et al (1995) used MS and this would appear to provide the stronger predictor of $\text{VO}_{2\text{peak}}$. The correlations found between the total number of laps and the $\text{VO}_{2\text{peak}}$ were significant and strong, even when the split by sex was performed (boys: $r=0.75$; girls: $r=0.69$). Matsuzaka et al (2004) reported similar values for the children engaging in their study ($r=0.75$). The values in the present study were fairly higher than the correlations attained by female ($r=0.51$), boys ($r=0.65$) and both participants ($r=0.69$) in the study of Liu et al (1992), and much higher when considering the values presented by Pitetti et al (2002) ($r=0.53$), but considerably inferior ($r=0.87$) when bearing in mind the study of Boreham et al (1990) in 15 year old adolescents. McVeigh et al (1995) taking into account the MS as the best single predictor of $\text{VO}_{2\text{peak}}$ reported quite lower Pearson product-moment correlation coefficients for boys ($r=0.60$) and girls ($r=0.65$) in comparison to the obtained in this study (boys: $r=0.74$; girls: $r=0.63$). The magnitude of these correlation coefficients can be questioned since the studies did not followed similar $\text{VO}_{2\text{peak}}$ criterion test (i.e., continuous modified Bruce TM test vs. measured directly 20-m MSRT performance) and the same 20-m MSRT protocol. Therefore, methodological differences are not to be discarded from constituting bias factors to our findings.

Like in Pitetti et al (2002), significant positive Spearman's rank correlation coefficient existed between the total number of laps (TL) for the PACER test and age ($r=0.27$), meaning that the PACER Test performance increased with age. However, no significant relationships were seen between the measured directly VO_{2peak} and age. Possibly this is triggered by the fact that the performance in the PACER test and the VO_{2max} obtained depends also on the individual's body dimensions (Bar-Or 1983; Mota, Guerra et al. 2002), running economy (Leger, Mercier et al. 1988), anaerobic capacity (Anderson 1992), turning skill (Grant, Corbett et al. 1995), motivation (Cairney, Hay et al. 2008), and its maturity. In fact, there is some evidence from longitudinal studies that the greatest increase in VO_{2peak} is associated with the attainment of peak height velocity (PHV) in both boys and girls, generally accepted as an age reference of biological age and maturation (Kobayashi, Kitamura et al. 1978; Rutenfranz, Lange Andersen et al. 1982). Accordingly to Malina et al (1991), boys reach their PHV approximately at 13 years old while girls precede them by 2 years (11 years old).

The development of maximal aerobic power as a function of maturation was examined by Rutenfranz et al (1982) during annual examinations of representative samples of 56 boys and 56 girls in Norway and Western Germany. Before the PHV the mean VO_{2max} was the same in two cohorts of children both in absolute values as well as in values relative to total and lean body mass. At and after PHV the boys were similar in their absolute values, but the German girls decreased their exercise fitness and became inferior to the Norwegian girls during later adolescence. This may explain the relevance of PHV in the triggering of maturational gender/age VO_{2max} differences. As none of the participants in the present study was over 10 years old, it would be expected that none of them had attained their PHV and consequently it can be considered normal the lack of significant correlation between VO_{2peak} and age.

Once again, and similarly to Pitetti et al (2002), significant negative relationships were found between the measured directly VO_{2peak} and weight and BMI meaning that the VO_{2peak} was superior as the BMI decreased. Previous studies (Leger, Mercier et al. 1988; Barnett, C. et al. 1993; McVeigh, Payne et al. 1995; Boreham, Twisk et al. 2001) have been equivocal regarding the significance of the gender factor in VO_{2max} prediction for children. Until the effects of gender on the relationship between 20-m MSRT performance and VO_{2max} are rigorously assessed, separate analysis for the gender groups should be maintained. No significant relationships were seen between the TL for the PACER test or measured directly VO_{2peak} , and height.

VALIDITY OF THE FITNESSGRAM'S PACER TEST SOFTWARE REPORTS AS A PREDICTOR OF VO_{2PEAK} IN LISBON ELEMENTARY SCHOOL CHILDREN

The main purpose of this investigation was to examine the validity of the FITNESSGRAM's 20-m MSRT (PACER test) software reports as a predictor of VO_{2max} in Lisbon healthy elementary school children, by measuring the VO_{2peak} during the actual 20-m MSRT performance. Significant differences were found when considering the entire sample ($d=1.8 \pm 7.3$; $p=0.02$) and boys ($d=4.4 \pm 7.5$; $p<0.001$). In girls on the other hand, it were not found significant differences between estimated and measured directly VO_{2peak} values ($d=-0.7 \pm 6.2$; $p=0.45$). These results suggest that the FITNESSGRAM's MSRT software reports underestimate the VO_{2peak} values determined by the K4b² in boys, but are suitable for girls. It is interesting to see that Boreham et al (1990) while comparing the validity of two field tests of aerobic fitness (PWC170 and 20-MSRT) for predicting VO_{2max} in a group of adolescent schoolchildren (boys: 15.6 ± 0.6 ; $n=24$; girls: 15.4 ± 0.7 ; $n=24$) indicated that the 20-m MSRT may be preferable for use with girls. A underestimating tendency of this field test has already been previously discussed by several authors (Sproule, Kunalan et al. 1993; Berthoin, Gerbeaux et al. 1994; Grant, Corbett et al. 1995) in particularly when VO_{2max} is beyond the healthy untrained subjects as in the original Leger and Lambert (1982) study.

Even though all the articles reviewed used the Pearson product-moment correlation coefficient (r) between the results of two or more measurement methods as an indicator of agreement, it is no such thing. The Pearson product-moment correlation coefficient measures the strength of a relation between two variables, not the agreement between them. We have perfect agreement only if the points lie along the line of equality, but we will have perfect correlation if the points lie along any straight line. In fact, when attention is afforded to both concordance correlation coefficient and Pearson product-moment correlation coefficient in this specific assessment of the validity of the FITNESSGRAM's 20-m MSRT (PACER test) software reports as a predictor of VO_{2max} in Lisbon healthy elementary school children, it is fairly evident that all the concordance correlation coefficients were manifestly inferior to the Pearson product-moment correlation coefficients. The correlations were no longer high (whole group: $r=0.72$; boys: $r=0.70$) or moderate (girls: $r=0.61$) but they were rather low

indeed (whole group: $r_c=0.38$; boys: $r_c=0.31$ and girls: $r_c=0.34$). If we take into consideration these values, we cannot keep suggesting that the FITNESSGRAM's MSRT software reports underestimate the VO_{2peak} measured directly in boys, but are suitable for girls because they are not. It urges the necessity to reflect on this expressed low agreement to the VO_{2peak} test criterion.

COMPARING SEVERAL REGRESSION EQUATIONS THAT PREDICT VO_{2PEAK} USING THE 20-M MSRT IN LISBON ELEMENTARY SCHOOL CHILDREN

Until a few years ago, the prediction equation of Leger et al (1988) was the only one available in the literature to convert the 20-m MSRT performance data to VO_{2max} ($ml.kg^{-1}.min^{-1}$) for children and adolescents 8-19 years old. The ability to convert 20-m MSRT performance data to VO_{2max} values is extremely practical in comparison to other field tests (i.e., PWC₁₇₀, step tests and endurance runs) for children and adolescents (Matsuzaka, Takahashi et al. 2004).

In order to determine the compatibility of previously reported prediction equations for these 90 Lisbon healthy elementary school children, we compared the predicted VO_{2max} using previously published formulas (Leger, Mercier et al. 1988; Barnett, C. et al. 1993; Fernhall, Pitetti et al. 1998; Oliveira 1998; Matsuzaka, Takahashi et al. 2004) with those measured directly in this study.

All the Pearson product-moment correlation coefficients between the measured directly VO_{2peak} and each one of the previously reported VO_{2max} estimation equations used in this study were significant ($p<0.01$). The correlation coefficient found between measured directly VO_{2peak} and the scores from the equation of Leger et al (1988) were 0.71 in boys, 0.68 in girls and 0.75 when considering both as a single group. Those studies that also used the equation of Leger et al (1988) reported correlation coefficients of 0.71 (Leger, Mercier et al. 1988), 0.72 (Barnett, C. et al. 1993), 0.72 (Liu, Plowman et al. 1992), 0.93 (Falgairrette, Bedu et al. 1994), 0.57 (Pitetti, Fernhall et al. 2002), and 0.77 (Matsuzaka, Takahashi et al. 2004).

The regression equations that best predicted VO_{2peak} reported by Barnett et al (1993) included triceps skinfold ($r=0.85$, $SEE=3.7$), gender and weight ($r=0.84$, $SEE=3.7$), and gender and age ($r=0.82$, $SEE=4.0$). The formula of Fernhall et al (1998) showed significantly higher regression coefficients ($r=0.88$) and lower SEE (4.5) for children and adolescents with MR when compared to the Pitetti et al (2002) study in nondisabled, healthy youths ($r=0.66$; $SEE=6.0$). At last, Matsuzaka et al (2004) with their predictive equation including gender, age, BMI and TL reported high correlation coefficients ($r=0.80$; $SEE=3.4$). However, a different picture is drawn when we pay close attention to the scores derived from the above reported VO_{2peak} prediction equations in this study. In relation to the regressions equations reported by Barnett et al (1993), we observed inferior Pearson product-moment correlation coefficient in the triceps skinfold equation ($r=0.74$), in the gender and weight equation ($r=0.62$), and finally in the gender and age equation ($r=0.65$). In comparison to the scores reported by Pitetti et al (2002) higher correlation coefficients were seen ($r=0.79$). At last, inferior correlation coefficients were found derived from the predictive equation of Matsuzaka et al (2004) including gender, age, BMI and TL ($r=0.78$).

Indeed, for the participants in this study, multiple regression equations using age, sex and MS (Barnett, C. et al. 1993), and age, sex and TL (Fernhall, Pitetti et al. 1998), showed consistently lower Pearson product-moment correlation coefficient than seen in other studies. In the investigation of Pitetti et al (2002), a moderate Pearson product-moment correlation coefficient ($r=0.66$) was found between the equation of Fernhall et al (1998) and the treadmill directly measured VO_{2peak} . Matsuzaka et al (2004) for instance, stated high Pearson product-moment correlation coefficients for boys ($r=0.81$) and girls ($r=0.87$) when testing the relationships between the treadmill directly measured VO_{2peak} and the equation of Barnett et al (1993) which included gender and age, and when correlating the treadmill directly measured VO_{2peak} and the equation of Fernhall et al (1998) ($r=0.82$ for boys; $r=0.81$ for girls).

Since BMI is a significant predictor of VO_{2peak} in the regression formula of Fernhall et al (1998), it would be expected, or at least justified, dissimilar correlation coefficients in those studies whose participants had anthropometric characteristics outside the spectrum values of the original study. Nonetheless, the BMI of the children participating in the present study (17.7 ± 3.1 in boys and 18.3 ± 2.5 in girls) was quite similar to that of boys and girls from other investigations (around 17.7 ± 1.9 in boys and 18.4 ± 3.4 in girls) (Pitetti, Fernhall et al. 2002; Matsuzaka, Takahashi et al. 2004). Therefore, the lower r seen

that motivation and rhythm constituted a setback since it is evident from the high mean HR_{max} (200.1 ± 7.9 ; $p=0.70$) registered during the retest, that the participants were well focused and intrinsically motivated for the test.

CONCLUSIONS

The results of the present study may be summarized of the following manner:

1. The 20-m MSRT (PACER test) is a valid tool for the evaluation of VO_{2max} in Lisbon healthy elementary school children;
2. Weight and BMI were related to VO_{2peak} or total number of laps for the 20-m MSRT;
3. FITNESSGRAM's 20-m MSRT software reports underestimated the measured directly VO_{2peak} values in boys, but were suitable for girls. Additionally, it tends to overpredict VO_{2peak} in unfit participants and underpredict VO_{2peak} in fit participants;
4. Previously reported equations might be unsuitable for Lisbon healthy elementary school children. We suggest that the FITNESSGRAM's software reports or the equation of Fernhall et al (1998) may be the most appropriate. However, they still yield unsatisfactory accuracy. Our findings advocate that published equations predicting VO_{2peak} from MSRT may provide reasonable group estimates, but poor individual predictions of VO_{2peak} .
5. Test-retest correlation coefficients for the 20-m MSRT were high and the differences between the means were not significant.

FUTURE RECOMENDATIONS

In future investigations on this topic, we consider that eventually it is important to attend to some of the following points:

1. There is a critical need for standardization in the way that the 20-m MSRT is administered and the results analysed and reported.
2. Pearson product-moment correlations (r) measure the strength of a relation between two variables, not the agreement between them. We have perfect agreement only if the points lie along the line of equality, but we will have perfect correlation if the points lie along any straight line. Therefore, we suggest a broader use of the concordance correlation coefficient (r_c). We further advance that until the effects of gender on the relationship between 20-m MSRT performance and VO_{2peak} are rigorously assessed, separate analysis for the gender groups should be maintained.
3. It might be interesting to use the maturational age instead of the chronological age of the children in a future study on this subject.
4. The results from present study enhance and urge the establishment of a regression equation specific for the Portuguese elementary school children. This future study should come up with sampling procedures that obviously do not raise issues of representativeness. Future recommendations on the difference between total number of laps and maximal speed as measures of 20-m MSRT performance and the physiological implications for VO_{2max} prediction is required. The inclusion of skinfold measurement in the regression equation may help to account for differences in running economy and running styles between individuals with different body composition or reduce the requirement of repeat testing.
5. At last, we step forward and launch the initiative of a coordinated global approach to systematically monitor changes in children's fitness. Mechanisms should be put in place to facilitate data-pooling, such as an internet-based data repository.

REFERENCES

- ACSM (2005). ACSM's Guidelines for exercise testing and prescription. Baltimore, Lippincott Williams & Wilkins.
- Ahmaidi, S., K. Collomp, et al. (1992). "The effect of shuttle test protocol and the resulting lactacidaemia on maximal velocity and maximal oxygen uptake during the shuttle exercise test." Eur J Appl Physiol Occup Physiol **65**(5): 475-9.
- Andersen, R. E., C. J. Crespo, et al. (1998). "Relationship of physical activity and television watching with body weight and level of fatness among children: results from the Third National Health and Nutrition Examination Survey." JAMA **279**(12): 938-42.
- Anderson, G. S. (1992). "The 1600-m run and multistage 20-m run as a predictive test of aerobic capacity in children." Pediatr Exerc Sci **4**: 312-318.
- Anderssen, S. A., A. R. Cooper, et al. (2007). "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 **14**: 526-531.
- Ardern, C. I., P. T. Katzmarzyk, et al. (2005). "Revised Adult Treatment Panel III guidelines and cardiovascular disease mortality in men attending a preventive medical clinic." Circulation **112**(10): 1478-85.
- Armstrong, N. and J. R. Welsman (2000). "Development of aerobic fitness during childhood and adolescence." Pediatr Exerc Sci **12**(2): 128-149.
- Astrand, P.-O., S. Stromme, et al. (2003). Textbook of work physiology - physiological bases of exercise. U.K, Human Kinetics.
- Aziz, A., F. H. Y. Tan, et al. (2005). "A pilot study comparing two field tests with the treadmill run test in soccer players." J Sports Sci Med **4**: 105-112.
- Baquet, G., S. Berthoin, et al. (2002). "Effects of high intensity intermittent training on peak VO₂ in prepubertal children." Int J Sports Med **23**(6): 439-44.
- Bar-Or, O. (1983). Pediatric sports medicine for the practitioner. From physiologic principles to clinical applications. New-York, USA, Springer-Verlag.
- Bar-Or, O. (2000). "New and old in pediatric exercise physiology." Int J Sports Med **21** Suppl 2: S113-6; discussion S117.
- Bar-Or, O., R. J. Shephard, et al. (1971). Cardiac output of 10-13 year old boys and girls during submaximal exercise. Journal of applied physiology. **30** (2): 219-223.
- Barnett, A., L. Y. S. C., et al. (1993). "A preliminary study of the 20-m multistage shuttle run as a predictor of peak vo₂ in hong kong chinese students." Pediatr Exerc Sci **5**(1): 42-50.
- Bassett, D. R., Jr. and E. T. Howley (2000). "Limiting factors for maximum oxygen uptake and determinants of endurance performance." Med Sci Sports Exerc **32**(1): 70-84.
- Bergh, U., B. Sjodin, et al. (1991). "The relationship between body mass and oxygen uptake during running in humans." Med Sci Sports Exerc **23**(2): 205-11.
- Berthoin, S., M. Gerbeaux, et al. (1994). "Comparison of two field tests to estimate maximum aerobic speed." J Sports Sci **12**(4): 355-62.
- Blair, S. N., H. W. Kohl, 3rd, et al. (1989). "Physical fitness and all-cause mortality. A prospective study of healthy men and women." JAMA **262**(17): 2395-401.
- Bland, J. M. and D. G. Altman (1986). "Statistical methods for assessing agreement between two methods of clinical measurement." Lancet **1**(8476): 307-10.
- Boreham, C., V. J. Paliczka, et al. (1990). "A comparison of the PWC170 and 20-MST tests of aerobic fitness in adolescent schoolchildren." J Sports Med Phys Fitness **30**(1): 19-23.
- Boreham, C., J. Twisk, et al. (2001). "Fitness, fatness, and coronary heart disease risk in adolescents: the Northern Ireland Young Hearts Project." Med Sci Sports Exerc **33**(2): 270-4.
- Borms, J. (1986). "The child and exercise: an overview." J Sports Sci **4**(1): 3-20.
- Bouziotas, C., Y. Koutedakis, et al. (2001). "The prevalence of selected modifiable coronary heart disease risk factors in 12-year-old greek boys and girls " Pediatr Exerc Sci **13**(2): 173-184.
- Buck, S. M., C. H. Hillman, et al. (2008). "The relation of aerobic fitness to stroop task performance in preadolescent children." Med Sci Sports Exerc **40**(1): 166-72.

- Buono, M., J. Roby, et al. (1991). "Validity and reliability of predicting maximum oxygen uptake via field tests in children and adolescents." *Pediatr Exerc Sci* **3**(3): 250-255.
- Cairney, J., J. A. Hay, et al. (2008). "Generalized self-efficacy and performance on the 20-metre shuttle run in children." *Am J Hum Biol* **20**(2): 132-8.
- Caspersen, C. J., K. E. Powell, et al. (1985). "Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research." *Public Health Rep* **100**(2): 126-31.
- Cauderay, M., F. Narring, et al. (2000). "A cross-sectional survey assessing physical fitness of 9- to 19-year-old girls and boys in Switzerland." *Pediatr Exerc Sci* **12**(4): 398-412.
- Chun, D. M., C. B. Corbin, et al. (2000). "Validation of criterion-referenced standards for the mile run and progressive aerobic cardiovascular endurance tests." *Res Q Exerc Sport* **71**(2): 125-34.
- Cole, T., M. Bellizzi, et al. (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey. *British Medical Journal*. **320 (7244)**: 1240-1243.
- Coleman, K. J., C. L. Tiller, et al. (2005). "Prevention of the epidemic increase in child risk of overweight in low-income schools: the El Paso coordinated approach to child health." *Arch Pediatr Adolesc Med* **159**(3): 217-24.
- Cooke, C. B., M. J. McDonagh, et al. (1991). "Effects of load on oxygen intake in trained boys and men during treadmill running." *J Appl Physiol* **71**(4): 1237-44.
- Cooley, D. and L. McNaughton (1999). "Aerobic fitness of Tasmanian secondary school children using the 20-m shuttle run test." *Percept Mot Skills* **88**(1): 188-98.
- Cooper, K. (1970). *The new aerobics*. New-York, Bantam Books Inc.
- Council of Europe (1988). *Eurofit: handbook for the Eurofit tests of physical fitness*. Rome, Council of Europe.
- Cumming, G. R. (1978). "Recirculation times in exercising children." *J Appl Physiol* **45**(6): 1005-8.
- Cureton, K. J., R. A. Boileau, et al. (1977). "Determinants of distance running performance in children: analysis of a path model." *Res Q* **48**(2): 270-9.
- Cureton, K. J., M. A. Sloniger, et al. (1995). "A generalized equation for prediction of VO₂peak from 1-mile run/walk performance." *Med Sci Sports Exerc* **27**(3): 445-51.
- Daniels, J., N. Oldridge, et al. (1978). "Differences and changes in VO₂ among young runners 10 to 18 years of age." *Med Sci Sports* **10**(3): 200-3.
- Davies, C. T. (1968). "Limitations to the prediction of maximum oxygen intake from cardiac frequency measurements." *J Appl Physiol* **24**(5): 700-6.
- Davies, C. T. (1980). "Metabolic cost of exercise and physical performance in children with some observations on external loading." *Eur J Appl Physiol Occup Physiol* **45**(2-3): 95-102.
- Eriksson, B. O., G. Grimby, et al. (1971). "Cardiac output and arterial blood gases during exercise in pubertal boys." *J Appl Physiol* **31**(3): 348-52.
- Falgairrette, G., M. Bedu, et al. (1994). "Evaluation of physical fitness from field tests at high altitude in circumpubertal boys: comparison with laboratory data." *Eur J Appl Physiol Occup Physiol* **69**(1): 36-43.
- Fernhall, B., K. H. Pitetti, et al. (1998). "Validation of cardiovascular fitness field tests in children with mental retardation." *Am J Ment Retard* **102**(6): 602-12.
- Flouris, A. D., G. S. Metsios, et al. (2005). "Enhancing the efficacy of the 20 m multistage shuttle run test." *Br J Sports Med* **39**(3): 166-70.
- Freedson, P. S., T. B. Gilliam, et al. (1981). "Transient VO₂ characteristics in children at the onset of steady-rate exercise." *Res Q Exerc Sport* **52**(2): 167-73.
- Gaultier, C., L. Perret, et al. (1981). "Occlusion pressure and breathing pattern in healthy children." *Respir Physiol* **46**(1): 71-80.
- Gill JS, Zezulka AV, et al. (1985). "Relationship between initial blood pressure and its fall with treatment." *Lancet*: i: 567-569.
- Gilliam, T., P. Freedson, et al. (1981). "Physical activity patterns determined by heart rate monitoring in 6-7 year-old children." *Med Sci Sports Exerc* **13**(1): 65-7.
- Grant, S., K. Corbett, et al. (1995). "A comparison of methods of predicting maximum oxygen uptake." *Br J Sports Med* **29**(3): 147-52.
- Gratas-Delamarche, A., J. Mercier, et al. (1993). "Ventilatory response of prepubertal boys and adults to carbon dioxide at rest and during exercise." *Eur J Appl Physiol Occup Physiol* **66**(1): 25-30.

- Green, H. J. and A. E. Patla (1992). "Maximal aerobic power: neuromuscular and metabolic considerations." Med Sci Sports Exerc **24**(1): 38-46.
- Hill, J. O. and E. L. Melanson (1999). "Overview of the determinants of overweight and obesity: current evidence and research issues." Med Sci Sports Exerc **31**(11 Suppl): S515-21.
- Hopkins, W. G. (2000). "Measures of reliability in sports medicine and science." Sports Med **30**(1): 1-15.
- Issekutz, B., N. Birkhead, et al. (1962). "Use of respiratory quotients in assessment of aerobic work capacity " J Appl Physiol **17**: 47-50.
- Kline, G. M., J. P. Porcari, et al. (1987). "Estimation of VO₂max from a one-mile track walk, gender, age, and body weight." Med Sci Sports Exerc **19**(3): 253-9.
- Kobayashi, K., K. Kitamura, et al. (1978). "Aerobic power as related to body growth and training in Japanese boys: a longitudinal study." J Appl Physiol **44**: 666-672.
- Krahenbuhl, G. S. and R. P. Pangrazi (1983). "Characteristics associated with running performance in young boys." Med Sci Sports Exerc **15**(6): 486-90.
- Krahenbuhl, G. S., J. Skinner, et al. (1989). "Fractional utilization of maximal aerobic capacity in children 6 to 8 years of age." Pediatr Exerc Sci **1**(3): 271-277.
- Krahenbuhl, G. S. and T. J. Williams (1992). "Running economy: changes with age during childhood and adolescence." Med Sci Sports Exerc **24**(4): 462-6.
- Lamb, K. L. and L. Rogers (2007). "A re-appraisal of the reliability of the 20 m multi-stage shuttle run test." Eur J Appl Physiol **100**(3): 287-92.
- Leger, L. and R. Boucher (1980). "An indirect continuous running multistage field test: the Universite de Montreal track test." Can J Appl Sport Sci **5**(2): 77-84.
- Leger, L. and C. Gadoury (1989). "Validity of the 20 m shuttle run test with 1 min stages to predict VO₂max in adults." Can J Sport Sci **14**(1): 21-6.
- Leger, L. and J. Lambert (1982). "A maximal multistage 20-m shuttle run test to predict VO₂ max." Eur J Appl Physiol Occup Physiol **49**(1): 1-12.
- Leger, L., D. Mercier, et al. (1988). "The multistage 20 metre shuttle run test for aerobic fitness." J Sports Sci **6**(2): 93-101.
- Lemmink, K. A., R. Verheijen, et al. (2004). "The discriminative power of the Interval Shuttle Run Test and the Maximal Multistage Shuttle Run Test for playing level of soccer." J Sports Med Phys Fitness **44**(3): 233-9.
- Lin, L. I. (1989). "A concordance correlation coefficient to evaluate reproducibility." Biometrics **45**(1): 255-68.
- Lind, A. R. (1970). "Cardiovascular responses to static exercise. (Isometrics, anyone?)." Circulation **41**(2): 173-6.
- Liu, N. Y., S. A. Plowman, et al. (1992). "The reliability and validity of the 20-meter shuttle test in American students 12 to 15 years old." Res Q Exerc Sport **63**(4): 360-5.
- Máček, M. and V. J. (1980). "The adjustment of oxygen uptake at the onset of exercise: a comparison between prepubertal boys and young adults." Int J Sports Med **01**: 70-72.
- Macfarlane, D. J. and G. R. Tomkinson (2007). "Evolution and variability in fitness test performance of Asian children and adolescents." Med Sport Sci **50**: 143-67.
- Mahon, A. D. and P. Vaccaro (1989). "Ventilatory threshold and VO₂max changes in children following endurance training." Med Sci Sports Exerc **21**(4): 425-31.
- Mahoney, C. (1992). "20-MST and PWC170 validity in non-Caucasian children in the UK." Br J Sports Med **26**(1): 45-7.
- Malina, R. M. (2001). "Physical activity and fitness: pathways from childhood to adulthood." Am J Hum Biol **13**(2): 162-72.
- Malina, R. M., C. Bouchard, et al. (1991). Growth, maturation, and physical activity. Champaign, Illinois, Human Kinetics.
- Massicotte, D. R., R. Gauthier, et al. (1985). "Prediction of VO₂max from the running performance in children aged 10-17 years." J Sports Med Phys Fitness **25**(1-2): 10-7.
- Matsuzaka, A., Y. Takahashi, et al. (2004). "Validity of the multistage 20-m shuttle-run test for Japanese children, adolescents, and adults." Pediatr Exerc Sci **16**(2).
- McLaughlin, J., G. King, et al. (2001). "Validation of the COSMED K4 b2 portable metabolic system." Int J Sports Med **22**(4): 280-4.
- McVeigh, S., A. Payne, et al. (1995). "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 **7**(1): 69-79.

- Mercier, D., L. L.A, et al. (1983). "Relative efficiency and predicted vo2 max in children." *Med Sci Sports Exerc* **15**(2): 143-143.
- Mota, J., S. Guerra, et al. (2002). "Association of maturation, sex, and body fat in cardiorespiratory fitness." *Am J Hum Biol* **14**(6): 707-12.
- Olds, T. and J. Dollman (2004). "Are changes in distance-run performance of australian children between 1985 and 1997 explained by changes in fatness?" *Pediatr Exerc Sci* **16**(3): 201-209.
- Olds, T., G. Tomkinson, et al. (2006). "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* **24**(10): 1025-38.
- Oliveira, J. (1998). *Validação directa do teste vaivém em 20 metros, de Luc-Léger em adolescentes portuguesas. Exercício e Saúde*. Lisboa, Universidade Técnica de Lisboa - Faculdade de Motricidade Humana. **Mestrado**.
- Paffenbarger, R. S., Jr., R. T. Hyde, et al. (1986). "Physical activity, all-cause mortality, and longevity of college alumni." *N Engl J Med* **314**(10): 605-13.
- Paliczka, V. J., A. K. Nichols, et al. (1987). "A multi-stage shuttle run as a predictor of running performance and maximal oxygen uptake in adults." *Br J Sports Med* **21**(4): 163-5.
- Paridon, S. M., B. S. Alpert, et al. (2006). "Clinical stress testing in the pediatric age group: a statement from the American Heart Association Council on Cardiovascular Disease in the Young, Committee on Atherosclerosis, Hypertension, and Obesity in Youth." *Circulation* **113**(15): 1905-20.
- Pate, R. R., T. Baranowski, et al. (1996). "Tracking of physical activity in young children." *Med Sci Sports Exerc* **28**(1): 92-6.
- Payne, V. G. and J. R. Morrow, Jr. (1993). "Exercise and VO2 max in children: a meta-analysis." *Res Q Exerc Sport* **64**(3): 305-13.
- Pitetti, K. H., B. Fernhall, et al. (2002). "Comparing two regression formulas that predict vo2peak using the 20-m shuttle run for children and adolescents." *Pediatr Exerc Sci* **14**(2): 125-134.
- Ramsbottom, R., J. Brewer, et al. (1988). "A progressive shuttle run test to estimate maximal oxygen uptake." *Br J Sports Med* **22**(4): 141-4.
- Ramsbottom, R., M. G. Nute, et al. (1987). "Determinants of five kilometre running performance in active men and women." *Br J Sports Med* **21**(2): 9-13.
- Rikli, R. E., C. Petray, et al. (1992). "The reliability of distance run tests for children in grades K-4." *Res Q Exerc Sport* **63**(3): 270-6.
- Rivera-Brown, A., M. Alvarez, et al. (2001). "Anaerobic power and achievement of VO2 plateau in pre-pubertal boys." *Int J Sports Med* **22**(2): 111-5.
- Rogers, D. M., B. L. Olson, et al. (1995). "Scaling for the VO2-to-body size relationship among children and adults." *J Appl Physiol* **79**(3): 958-67.
- Rosenthal, M. and A. Bush (1998). "Haemodynamics in children during rest and exercise: methods and normal values." *Eur Respir J* **11**(4): 854-65.
- Rowland, T., D. Goff, et al. (2000). Influence of cardiac functional capacity on gender differences in maximal oxygen uptake in children. *Chest*. **117**: 629-635.
- Rowland, T. W. (1993). "Does peak VO2 reflect VO2max in children?: evidence from supramaximal testing." *Med Sci Sports Exerc* **25**(6): 689-93.
- Rowland, T. W. (2002). "Declining cardiorespiratory fitness in youth: fact or supposition?" *Pediatr Exerc Sci* **14**(1): 1-8.
- Rowland, T. W. and L. N. Cunningham (1992). "Oxygen uptake plateau during maximal treadmill exercise in children." *Chest* **101**(2): 485-9.
- Rowland, T. W. and L. N. Cunningham (1997). "Development of ventilatory responses to exercise in normal white children. A longitudinal study." *Chest* **111**(2): 327-32.
- Rutenfranz, J., K. Lange Andersen, et al. (1982). "Maximal aerobic power affected by maturation and body growth during childhood and adolescence." *Eur J Pediatr* **139**(2): 106-12.
- Safrit, M. (1990). "The validity and reliability of fitness tests for children: a review." *Pediatr Exerc Sci* **2**(1): 9-28.
- Sallis, J. and K. Patrick (1994). "Physical Activity Guidelines for Adolescents: Consensus Statement " *Pediatr Exerc Sci* **6**(4): 302-314.
- Sallis, J. F. and T. L. McKenzie (1991). "Physical education's role in public health." *Res Q Exerc Sport* **62**(2): 124-37.
- Saltin, B. and S. Strange (1992). "Maximal oxygen uptake: "old" and "new" arguments for a cardiovascular limitation." *Med Sci Sports Exerc* **24**(1): 30-7.

- Slaughter, M. H., T. G. Lohman, et al. (1988). "Skinfold equations for estimation of body fatness in children and youth." Hum Biol **60**(5): 709-23.
- Sproule, J., C. Kunalan, et al. (1993). "Validity of 20-MST for predicting VO₂max of adult Singaporean athletes." Br J Sports Med **27**(3): 202-4.
- St Clair Gibson, A., S. Broomhead, et al. (1998). "Prediction of maximal oxygen uptake from a 20-m shuttle run as measured directly in runners and squash players." J Sports Sci **16**(4): 331-5.
- Sutton, J. R. (1992). "VO₂max--new concepts on an old theme." Med Sci Sports Exerc **24**(1): 26-9.
- The Cooper Institute for Aerobics Research (1994). FITNESSGRAM - the Prudential FITNESSGRAM test administration manual. Dallas, Texas, USA.
- The Cooper Institute for Aerobics Research (2002). FITNESSGRAM - manual de aplicação de testes. Lisboa, Faculdade de Motricidade Humana.
- Tomkinson, G. R., L. A. Leger, et al. (2003). "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 **33**(4): 285-300.
- Trowbridge, C. A., B. A. Gower, et al. (1997). "Maximal aerobic capacity in African-American and Caucasian prepubertal children." Am J Physiol **273**(4 Pt 1): E809-14.
- van Mechelen, W., H. Hlobil, et al. (1986). "Validation of two running tests as estimates of maximal aerobic power in children." Eur J Appl Physiol Occup Physiol **55**(5): 503-6.
- Vincent, S. D., R. Barker, et al. (1999). "A comparison of peak heart rates elicited by the 1-mile run/walk and the progressive aerobic cardiovascular endurance run." Res Q Exerc Sport **70**(1): 75-8.
- Vincent, S. D., R. P. Pangrazi, et al. (2003). "Activity levels and body mass index of children in the United States, Sweden, and Australia." Med Sci Sports Exerc **35**(8): 1367-73.
- Washington, R. L., J. T. Bricker, et al. (1994). "Guidelines for exercise testing in the pediatric age group. From the Committee on Atherosclerosis and Hypertension in Children, Council on Cardiovascular Disease in the Young, the American Heart Association." Circulation **90**(4): 2166-79.
- Whitaker, R. C., J. A. Wright, et al. (1997). "Predicting obesity in young adulthood from childhood and parental obesity." N Engl J Med **337**(13): 869-73.
- Wilkinson, D. M., J. L. Fallowfield, et al. (1999). "A modified incremental shuttle run test for the determination of peak shuttle running speed and the prediction of maximal oxygen uptake." J Sports Sci **17**(5): 413-9.
- Zar, J. H. (1999). Biostatistical Analysis. Upper Saddle River, NJ, Prentice-Hall, Inc.