Abstract

DEVELOPMENT AND VALIDATION OF A REGRESSION MODEL TO ESTIMATE VO_2 max FOR OLDER ADOLESCENTS FROM PACER 20-M SHUTTLE RUN PERFORMANCE

By

Matthew Hanna

09/2017

Director of Thesis: Dr. Matthew Mahar

Major Department: Department of Kinesiology

The PACER provides a convenient and relatively low cost assessment of aerobic fitness. The PACER is less expensive and more readily available than stationary or portable metabolic systems, can be completed in a condensed area of 20 meters, and can be administered to large groups. The PACER is the preferred (or default) test of aerobic fitness for the FITNESSGRAM® youth fitness program. The purpose of this study was to develop a regression model to estimate VO_2 max from PACER performance, age, gender, and body mass or body mass index in 17- to 18-year-old males and females. Additionally, several previously published equations were cross-validated to allow for comparison between prediction models. A secondary purpose was to examine the criterion-referenced validity of the models. Participants included 22 females and 26 males aged 17 to 18 years. Height and weight were assessed and percent body fat was measured with air displacement plethesmography via the BOD POD. Participants completed a maximal treadmill test following the Bruce protocol. The PACER 20-m multi-stage shuttle run was completed following FITNESSGRAM® procedures. Multiple linear regression analysis was used to predict VO_2 max from the number of laps completed on the PACER, age, gender, and body mass or body mass index. Estimates of VO_2 max from previously published prediction

models were cross-validated and the standard error of estimate (*SEE*) and total error (TE) were calculated. Comparison of the two error estimates identified the effect of systematic overestimation or underestimation on prediction accuracy. Mean scores for VO_2 max for males (41.9 ± 9.9) and females (33.1 ± 6.7) were lower than the 2010 FITNESSGRAM® Healthy Fitness Zone (HFZ) standards (males ≥ 44.2 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ and females ≥ 38.6 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$). Only 27% of female participants (*n* = 6) and 42% of male participants (*n* = 11) had a measured VO_2 in the HFZ. The only variable that significantly contributed to estimation of VO_2 max was PACER laps completed. The model developed on the current sample was: VO_2 max = 20.41012 + (PACER laps * 0.41304). The correlation between measured VO_2 max and VO_2 max predicted from this equation was high (R = .89). All cross-validated prediction equations produced high correlations between measured and estimated VO_2 max ($R \geq .81$). The prediction model developed in the current study and the Léger et al. (1988) model produced the highest correlations (r = .89) between measured and estimated VO_2 max, the lowest standard errors of estimate (4.36 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ and 4.43 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$), and the lowest *TE* (4.36 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ and 4.81 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$). Criterion-referenced analysis was used to examine the classification accuracy of the models for three categories (HFZ, Needs Improvement – Some Risk [NISR], and Needs Improvement – High Risk [NIHR]) and two categories (HFZ and Needs Improvement [NI]). Proportion of agreement for the three category analysis was moderate (*Pa* = .73) for the model developed in the current study. When classification was condensed into two classification zones (HFZ and NI), *Pa* increased (*Pa* = .88). The remaining cross-validated models had a low to moderate *Pa* (.52 - .66) under the three category format, and moderate *Pa* (.66 - .77) under the two category format. In conclusion, the prediction model developed in the current study developed on an older adolescent sample provides an accurate estimate of VO_2 max. The variable of laps completed on the PACER was the only significant contributor to the equation. Most of the previously published equations were notably less accurate than this new model. The current prediction model also produced accurate classification of fitness levels into the Healthy Fitness Zone or Needs Improvement categories.

DEVELOPMENT AND VALIDATION OF A REGRESSION MODEL TO ESTIMATE VO_2 MAX FOR OLDER ADOLESCENTS FROM PACER 20-M SHUTTLE RUN PERFORMACE.

A Thesis Presented to

The Faculty of the Department of Kinesiology

East Carolina University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Kinesiology/Physical Activity Promotion

by

Matthew Hanna

September 2017

DEVELOPMENT AND VALIDATION OF A REGRESSION MODEL TO ESTIMATE VO_2 MAX FOR OLDER ADOLESCENTS FROM PACER 20-M SHUTTLE RUN PERFORMACE.

by

Matthew Hanna

APPROVED BY:

DIRECTOR OF
THESIS: ____
Matthew Mahar, Ed.D.

COMMITTEE MEMBER: ____________________________________
Katrina DuBose, Ph.D.

COMMITTEE MEMBER: ____________________________________
C. David Kemble, M.A.

CHAIR OF THE DEPARTMENT
OF KINESIOLOGY: ______________________________________
Stacey Altman, J.D.

DEAN OF THE
GRADUATE SCHOOL: _____________________________________
Paul J. Gemperline, Ph.D.

DEDICATION

Dedicated to the memory of Dr. Nelson Cooper, who sat on this thesis committee and provided critical insight and support to this study. You were lost far too soon. You've left your indelible mark on East Carolina University and you will never be forgotten.

ACKNOWLEDGEMENT

Foremost, I would like to express my sincere gratitude to my advisor Dr. Matt Mahar for the continuous support of my graduate studies and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my graduate school studies.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Nelson Cooper, Dr. Katrina DuBose, and Mr. Dave Kemble, for their encouragement, insightful comments, and hard questions.

My sincere thanks also go to Dr. Kristina Karvinen and her research team for offering me an internship opportunity in her cancer research study.

I thank my fellow graduate assistants: Ashley Guerieri and Michelle Vuchenich for the stimulating discussions, for the long hours working together in the APL lab, and for all the fun we have had during our time at ECU.

Last but not the least, I would like to thank my dad, Ron Hanna, for supporting me throughout my time at ECU. No matter the personal struggles, you were an absolute inspiration. I wouldn’t have the work ethic I have today without your guidance and love. This accomplishment is just as much yours than mine.

TABLE OF CONTENTS

Introduction

Adolescent obesity has been rising at an alarming rate. Results from the National Health and Nutritional Examination Survey (NHANES) demonstrate that adolescent obesity has risen over 12% in the last three decades (1976-1980, 5% vs. 2003-2006 17.6%) [NHANES, 2006]. Physical activity drastically declines between the ages of 12-17 years. Caspersen et al. (1999) reported that the percentage of adolescent males that met physical activity recommendations decreased by 16% (40% to 24%) and the percentage of adolescent females that met physical activity recommendations decreased by 10% (30% to 20%) from young adolescents to older adolescents [Caspersen et al., 1999].

The obesity epidemic can be combated by improving all aspects of fitness. Blair et al. (1992) highlighted that improved aerobic fitness can benefit overall health amongst adolescents. Higher levels of aerobic fitness in older adolescents is associated with many health benefits, including: lower total cholesterol, higher systolic blood pressure, and lower levels of high-density lipoprotein cholesterol (Carnethon, Gulati, & Greenland, 2005).

Because of the importance of aerobic fitness for health, researchers have developed field tests to predict aerobic fitness. One available field test is the Progressive Aerobic Cardiovascular Endurance Run (PACER). The test is available through *FITNESSGRAM*® and is utilized by early education, secondary education, and college programs (Welk, Morrow, & Falls, 2002). The PACER is the recommended (or default) test of aerobic fitness in the *FITNESSGRAM*®. The PACER can be completed indoor, outdoor, without a track, and as an individual or group test. The protocol of the PACER resembles that of a maximal graded exercise test. Participants increase workload throughout the run and stop only when they have reached exhaustion or cannot maintain the cadence of the audio CD. A pre-recorded cadence is used to avoid pacing problems amongst participants.

In 2007, the *FITNESSGRAM®* physical fitness test was implemented into the Texas state curriculum. Texas legislated that grades 3 through 12 in every district must implement the *FITNESSGRAM®* test into every physical education class (Texas Department of Education, 2007). California implemented similar legislation within their school systems. Students in grades five, seven, and nine are required to take the *FITNESSGRAM®* fitness test whether or not they are enrolled in physical education class (California Department of Education, 2009). Over one million New York City students in grades K-12 completed the *FITNESSGRAM®* fitness test as part of city wide legislation (New York City Department of Education, 2009). Delaware has recently legislated the implementation of *FITNESSGRAM®* into their physical education classes. All Delaware students are required to complete the *FITNESSGRAM®* assessment with reporting of results for students in grades 4, 7, 9, and 10 (Delaware Department of Education, 2006). With several states beginning to implement the *FITNESSGRAM®* physical fitness test as a mandatory requirement, the need for a valid and reliable PACER prediction equation is of the utmost importance.

The PACER prediction equation for aerobic capacity (VO_2 max) used in the *FITNSSGRAM®* software was developed by Léger et al. (1988). The equation was developed on 188 males and females ages 8 to 19 years and had a moderately high correlation (r = .71) between estimated and measured VO_2 max. Léger et al. estimated VO_2 max by accounting for maximal speed attained during the 20-meter shuttle run, age, and the age-by-speed interaction. VO_2 max was measured via retroextrapolation after a test of maximal oxygen consumption.

The prediction equation established by Léger et al. (1988) has been tested on an adolescent sample by several researchers (Barnett, Chan, & Bruce, 1993; Liu, Plowman, & Looney, 1992, Mahar et al., 2006). However, little research has been conducted on the predictive power of the Léger et al. equation on an older adolescent sample. Mahar et al. (2006) indicated that close attention should be paid to the validity of the Léger et al. equation to predict VO_2 max amongst some populations. Mahar et al. noted that the distribution of ages for the Léger et al. sample was not published. Without knowing the age distribution of the Léger et al. sample it is difficult to determine how well different subgroups are presented. The need to re-examine the Léger et al. equation is strengthened by the absence of gender from the prediction model. Field tests like the one-mile run account for gender when estimating VO_2 max. The gender-by-age interaction is a significant determinant when estimating VO_2 max from the one-mile run test and may also influence performance on the PACER (Mahar et al., 2006).

Researchers have noted that accounting for age, gender, and body mass can improve the predictive power of the PACER equation (Mahar et al., 2006). Currently, limited research has investigated the predictive power of the PACER test for older adolescents. The development of a regression equation for predicting VO_2 max from the PACER could help practitioner's better assess the aerobic fitness of older adolescent students.

Purpose Statement

The purpose of this study was to develop a regression model to estimate VO_2 max from PACER performance, age, gender, and body mass or body mass index in 17- to 18-year-old males and females. Additionally, several previously published equations were cross-validated to allow for comparison between prediction models. A secondary purpose was to examine the criterion-referenced validity of the models.

Definitions

For the purposes of the study, the following terms are defined:

Air Displacement Plethysmography (ADP) **–** ADP is a device that estimates the amount of air displaced for body composition measurement. The BOD POD is a commercially available ADP system (Life Measurement Inc., Concord, CA).

Bioelectrical Impedance Analysis (BIA) **–** BIA is a method for estimating body composition. BIA works by measuring the resistance to flow of an electrical current as it moves through the body. By measuring this flow a practitioner can estimate body composition (Houtkooper, Going, Lohman, Rouche, & Van Loan, 1992).

Body composition - Body composition is the ratio of fat mass to fat free mass.

Bruce protocol - The Bruce protocol is a commonly used treadmill exercise stress test (Bruce, 1963).

Obesity **–** Obesity is the over accumulation of fat. Obesity is defined as a BMI $\geq$ 30.0 kg/m^2 for adults (NHLBI Obesity Education Initiative Expert Panel, 1998).

PACER **–** The PACER is a maximal effort test used to estimate aerobic fitness. The PACER test can be completed in two variations (15- and 20-meters) [Welk, Morrow, & Fall, 2002].

VO_2 max **-** Is the maximum capacity of an individual's body to transport and utilize oxygen during incremental exercise, which reflects the physical fitness of the individual (Bouchard et al., 1998).

Delimitations

The study will include the following delimitations:

1. All participants will be 17- or 18- years of age.

2. All participants will have their percent body fat assesses using ADP via the BODPOD.

3. All participants will perform the 20-m PACER (*FITNESSGRAM/ACTIVITYGRAM* Test Administration Manual Fourth Edition, 2007).

4. All participants will perform a maximal treadmill test following the Bruce protocol while VO_2 is assessed using the COSMED portable metabolic system.

Limitations

The study will be limited by the following:

1. The 20-m PACER and VO_2 max test will be completed in a laboratory setting.

2. Participants will complete the PACER individually.

Significance of the Study

Prediction of VO_2 max is essential for accurate representation of health amongst older adolescents. The current prediction equation provided by Léger et al. (1988) can produce spurious results for some populations (Mahar et al., 2002). The accuracy of predicted VO_2 max from the PACER for older adolescents has not been well established. Thus, research to examine the accuracy of a new regression equation from PACER performance, age, gender, and body composition for older adolescents, is needed.

Review of Literature

This chapter is devoted to a review of literature related to the importance of establishing a valid and reliable estimation of VO_2 max from the PACER 20-m multistage shuttle run. The review of literature contains the following sections: aerobic fitness standards for adolescents, benefits of aerobic fitness, physical activity levels among older adolescents, effects of physical activity levels on aerobic fitness, and reliability and validity of the 20-m multi-stage shuttle run as a predictor of VO_2 max; and summary.

Aerobic Fitness Standards for Adolescents

Pate, Wang, Dowda, Farrell, and O'Neil assessed the associations between aerobic fitness, age, sex, race/ethnicity, and self-reported physical activity in youth and adolescents aged 12 – 19 years. Pate et al. used data from 3,287 participants collected from 1999 – 2002 as part of the National Health and Nutrition Examination Survey (NHANES) study. Sex, race/ethnicity, age, body mass index (BMI), and activity levels were similar for the overall sample. Aerobic fitness was assessed using a submaximal treadmill exercise test. The test consisted of a 2-minute warm-up, followed by two 3-minute work stages, and a 2-minute cool down period. Aerobic fitness was estimated using extrapolation to an expected age-specific maximal heart rate using measured heart rate responses to the two 3-minute exercise stages. The participant's blood pressure, rate of perceived exertion, and heart rate were obtained during each stage of the test. The goal of the test was to elicit a heart rate that was with 80% of age-predicted maximum (220 – age) heart rate by the end of the exercise session. All statistical analyses were conducted separately for males and females. Means for estimated VO_2 max were calculated for four age groups (12 to 13, 14 to 15, 16 to 17, and 18 to 19 years of age).

Pate et al. (2006) found that males estimated VO_2 max was higher than the estimated VO_2 max of their female counterparts. Participants in the normal weight group had significantly higher estimated VO_2 max than those in the at-risk for overweight and overweight groups. In females, estimated VO_2 max was lower in the older group (aged 18 – 19 years) than in the younger group (12 – 13 years). Females in the normal weight group had higher estimated VO_2 max than females in the at-risk for overweight and overweight groups. Average estimated VO_2 max increased with age in males (12 – 13 years: 44.6 $ml \cdot kg^{-1} \cdot min^{-1}$ vs. 18 – 19 years: 47.6 $ml \cdot kg^{-1} \cdot min^{-1}$). However, the opposite was demonstrated amongst females. The highest values of estimated VO_2 max were recorded in the youngest age group (12 – 13 years: 39.7 $ml \cdot kg^{-1} \cdot min^{-1}$) with a steady decrease in aerobic capacity throughout adolescence. The lowest values were produced by the older adolescent age group (18 – 19 years: 37.5 $ml \cdot kg^{-1} \cdot min^{-1}$). Mean estimated VO_2 max was approximately 17% lower in females than in males in the overall sample, and the sex difference increased with age from 11% in 12- and 13-year-olds to 22% in 18- and 19-year-olds.

Estimated values of VO_2 max collected from the overall sample were then compared to the FITNESSGRAM® aerobic fitness standards. The criterion-referenced standards used by the FITNESSGRAM® for aerobic fitness were 42 $ml \cdot kg^{-1} \cdot min^{-1}$ for males and 35 to 37 $ml \cdot kg^{-1} \cdot min^{-1}$ (depending on age) for females. Findings from the study suggest that one-third of youth aged 12 to 19 years in the United States has low levels of aerobic fitness.

Lobelo, Pate, Dowda, Liese, and Ruiz (2009) investigated the ability of the criterion-referenced standards for aerobic fitness established by FITNESSGRAM® to discriminate between low and high cardiovascular disease risk in a sample of U.S. adolescents. A total of 4,902 adolescents aged 12 – 19 years were interviewed at home from 1999 through 2002 as part of the NHANES study. Of those individuals interviewed at home, 4,732 were examined at the mobile examination center. Only individuals who had complete data of systolic blood pressure (SBP), fasting glucose, insulin, high density lipoprotein cholesterol (HDLc), triglycerides (TG), total cholesterol (TC), and subscapular and triceps skinfolds (SKF), were included (*n* = 1,247). Aerobic fitness was assessed using a submaximal treadmill exercise test. The test consisted of a 2-minute warm-up, followed by two 3-minute work stages, and a 2-minute cool down period. The goal of the test was to elicit a heart rate that was with 80% of age-predicted maximum (220 – age) heart rate by the end of the exercise session. Aerobic fitness was estimated using extrapolation to an expected age-specific maximal heart rate using measured heart rate responses to the two 3-minute exercise stages.

Participants were classified into two levels (meeting/not meeting the cardiorespiratory fitness [CRF] standards) based on the FITNESSGRAM® standards. Individuals who met or exceeded 42 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ for males and 35 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ to 37 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ (depending on age) for females were classified as meeting CRF standards. Sex differences were assessed by one-way ANOVA. Mean differences for cardiovascular disease (CVD) risk factors and the CVD risk score by FITNESSGRAM® categories were analyzed using one-way ANOVA. Binary logistic regression was used to study the relationship between CVD risk factors and the CRF categories.

Lobelo et al. (2009) reported that 31% of young adolescent males (12 – 15 years) and 27% of older adolescent (16 – 18 years) males did not meet FITNESSGRAM® standards, while 30% of young adolescent females and 33% of older adolescent females did not meet the standards. Amongst males and females who did not meet FITNESSGRAM® aerobic fitness standards, CVD risk was significantly higher versus those who did meet the aerobic fitness standards. Regression results showed that males who met the aerobic fitness standard had significantly increased odds of not developing CVD when compared with males with aerobic fitness levels below the standard (12-15 years: OR = 5.17, 95% CI = 2.44 – 10.95; 16-19 years: OR = 3.78, 95% CI = 2.06 – 6.92). Females 12 – 15 years and 16 – 19 years who met aerobic fitness standards did not have a statistically significantly higher odds ratio of having low CVD risk than those who did not meet these aerobic fitness standards. Lobelo et al. concluded that adolescents who meet the aerobic fitness standards established by the FITNESSGRAM® have a significantly lower CVD risk score compared with those who do not meet aerobic fitness standard. However, Lobelo et al. also suggested that FITNESSGRAM® standards seem to discriminate adolescents with low risk for CVD from those with multiple risk factors, especially among males. But fitness cutoffs that better discriminate low from high CVD risk in youth that were calculated in their study were similar to those suggested by FITNESSGRAM®, which reinforces the value of the FITNESSGRAM® program for practitioners.

In the past year Welk, Cureton, Mahar, Zhu, Meredith, Laurson and Eisenmann (in press) reevaluated fitness standards that were previously published and those established by the FITNESSGRAM® group. Welk et al. developed new health related fitness standards using a nationally representative sample from the NHANES survey.

The new standards established by Welk et al. (in press) are age- and sex-specific and they also allow for normal changes in adolescent growth and maturation. The fitness standards for males increased (≥ 44.2 $ml \cdot kg^{-1} \cdot min^{-1}$), while standards decreased for females (≥ 38.8 $ml \cdot kg^{-1} \cdot min^{-1}$). The change speaks to the natural development trends of males and females during their adolescent years (males develop more muscle with age and become more fit while females have a tendency to gain body fat through adolescence and become less fit). Additionally, the new standards allow for three unique classifications in contrast to the standard two previously used in the FITNESSGRAM® program. Individuals who score above the new standard established for their gender would be classified in the “Healthy Fitness Zone”. Individuals who meet these criteria would be classified as having sufficient aerobic fitness for good health. Individuals who score below the standard would be categorized into the “Needs Improvement – SOME RISK” zone or “Needs Improvement – HIGH RISK” zone. Any individual who is categorized in the “Needs Improvement” zone would receive a message that informs them that improving aerobic fitness can reduce health risk.

Welk et al. (2010) concluded that new standards were needed to consider the natural physiological changes that occur to both males and females during adolescence. These changes often brought about misclassification of an individual’s aerobic fitness relative to the FITNESSGRAM® standards. The new age- and sex-specific standards categorize fitness levels into three zones instead of two. Differences in classifications will exist but these new standards are designed to decrease the misclassification of individuals into the Healthy Fitness Zone.

Benefits of Aerobic Fitness

Eisenmann, Wickel, Welk, and Blair (2005) investigated the relationship between aerobic fitness and body fatness during adolescence and the development of cardiovascular disease (CVD) risk factor in adulthood. The sample that was recruited for the study was comprised of 48 participants (36 men and 12 women). All participants received a health examination before the age of 18 (average age = 15.8 ± 2.1 years). Anthropometric measurements were taken on all participants (height, body mass, skinfolds, waist circumference, blood pressure (BP), fasting plasma lipids and glucose, and aerobic fitness determined by a maximal exercise test on a treadmill). Paired *t*-tests were used to identify differences between adolescent and adult variables for men, women, and the total sample. To account for the effect of sex and age on CVD risk factors, adolescent and adult variables were first regressed on the variables of age and sex and the residuals were saved. Adolescent residuals were subtracted from adult residuals and used to indicate differences for all relevant variables.

Eisenmann et al. (2004) reported that total treadmill time remained constant in male subjects, but significantly increased in female subjects from adolescence to adulthood. Additionally, they reported that body fatness during adolescence was inversely related to adult total treadmill time (r = -0.34 to -0.47). There was no significant association between adolescent total treadmill time and cardiovascular disease risk factors at follow-up. Adults' body fatness indicators (BMI, waist circumference, and percent body fat) were lower in participants classified in the adolescent high total treadmill time. Participants who were categorized in the high fat class in adolescence showed consistently lower total treadmill time and higher fatness in adulthood than those in the low fat class. Eisenmann et al. concluded that aerobic fitness and body fatness indicators crossed over from adolescence into adulthood - adolescent aerobic fitness is related to adult body fatness. Eisenmann et al. reported that high amounts of adult body fatness are moderately related to adult CVD risk factors.

Eisenmann, Katzmarzyk, Perusse, Tremblay, Després, and Bouchard (2005) investigated association of x (BMI) and aerobic fitness on (CVD) risk factors in adolescents. The sample was from the Quebec Family Study (QFS) and was comprised of 416 males and 345 females aged 9 – 18 years. Anthropometric measures were taken on all the participants and BMI was derived from these measures. Participants' aerobic capacities were measured using a cycle ergometer and the PWC_{150} protocol. The CVD risk factors Eisenmann et al. tested for were: triglyceride levels (TG), HDL levels, LDL levels, blood pressure, and fasting plasma glucose concentration.

Data collected for the PWC_{150} test was adjusted for body mass using regression. Furthermore, BMI, PWC_{150}, and the CVD risk factors were standardized for age by regressing them on age, age^2, and age^3 to account for any differences in the relationship with age. *Z*-scores were derived by summing the age-standardized residuals for TG, HDL, LDL, and blood pressure ratio. A lower score is indicative of a healthier profile. The sample was separated into four groups based on the median cut-points for age-adjusted BMI and PWC_{150}. These groups consisted of: low fit/low BMI, low fit/high BMI, high fit/low BMI, and high fit/high BMI. Differences between groups were assessed using a 2 (fitness) x 2 (BMI) ANOVA within each gender.

Eisenmann et al. (2005) reported that low fit males and females had higher blood lipids and glucose compared to their high fit counterparts. However, within their BMI categories none of the differences reached statistical significance. Low fit males in the high BMI category had higher systolic BP than their fit counterparts and low fit females in the low BMI category had higher systolic BP than their fit counterparts. Males that fell within the high fit/low BMI category had the lowest BP, TG, LDL, and the highest HDL. Amongst low fit/high BMI males and females had high BP, TG, GLU and the lowest HDL. CVD risk factors showed significant main effects with fitness and fatness. Within the high fit group, BMI was associated with blood pressure, and in girls fitness appears to influence TG and blood pressures as significant differences existed within fatness groups. Both males and females within the low fit group possessed a higher CVD risk score.

Eisenmann et al. (2005) concluded that aerobic fitness and BMI are associated with selected CVD risk factors in adolescents. Individuals who were classified in the low fit/high fatness group had the poorest CVD risk profile, while those classified within the high fit/low fatness group has the best CVD risk profile. Within fatness categories (i.e., low or high BMI), it suggests that a higher level of aerobic fitness is associated with a better CVD risk factor profile.

Eisenmann, Wickel, Welk and Blair (2004) examined adolescent fitness levels and the impact fitness has on CVD risk levels in adulthood. Eisenmann et al. selected participants from the Aerobic Center Longitudinal Study (ACLS) who were under the age of 18 years during their health examination. There was a total of 48 (36 men and 12 women) participants who met the criteria. Participants had a variety of medical and lifestyle assessments completed, the measures of most interest were height, body mass, skinfolds, waist circumferences, blood pressure (BP), fasting plasma lipids and glucose, and aerobic fitness determined by a maximal exercise test on a treadmill (Balke protocol).

Paired *t* – tests were used to examine differences between adolescent and adult variables for men, women, and the total sample. Regression was used to account for the effect of sex and age on CVD. Adolescent residuals were subtracted from adult residuals to obtain differences in scores. Correlations were computed to examine the associations between adolescent fitness and fatness and adult CVD risk factors. Adolescent aerobic fitness and body fatness variables were separated into two categories based on the median split: low (below median) and high (above median)

Eisenmann et al. (2004) reported that height, body mass, BMI, waist circumference, percent body fat, and blood pressure significantly increased from adolescence to adulthood. Total treadmill time stayed fairly constant in male subjects, but significantly increased in female subjects from adolescence to adulthood. The correlation between adolescent aerobic fitness and adult aerobic fitness was significant (r = .53) and there were moderate to moderately high correlations (r = .44 - .78) between adolescent fatness and adults fatness. Adolescent fatness had an inverse correlation (*r* = - 0.32 to – 0.44) with adult total treadmill time. Individuals who had low body fat during adolescence had higher total treadmill time during adulthood (*r* = -.34 to -.47). Additionally, adult body fatness indicators (BMI, waist circumference, and percent body fat) were significantly lower in subjects classified in the adolescent high total treadmill time category compared to those in the adolescent low total treadmill time category. Conversely, participants in the high fat categories during adolescence showed consistently lower total treadmill time and high fatness in adulthood when compared to low fat counterparts. Adults in the high fitness category had a lower BMI and percent fat during adolescents when compared to their low fitness adult counterparts. Eisenmann et al. concluded that there is a significant relationship between adolescent aerobic fitness and adult body fatness.

Carnethon, Gidding, Nehgme, Sidney, Jacobs, and Liu (2003) reported similar findings in their 15 year longitudinal study of aerobic fitness in 2,478 older adolescents and development of cardiovascular disease risk factors in adulthood. Participants were originally screened for incidence of hypertension, diabetes, the metabolic syndrome, and hypercholesterolemia. All participants had aerobic fitness measured via a treadmill test (Balke protocol) at baseline and then again 7 years later.

Carnethon et al. (2003) tested the association between fitness change over 7 years and development of CVD risk factors using proportional hazards regression. To evaluate if associations were similar across race, sex, and baseline obesity, researchers included multiplicative interactions between each covariate of interest and fitness and conduct stratified analyses. A significant interaction was determined by the significance of the interaction term and stratum-specific estimates that differed between each other.

Carnethon et al. (2003) reported that low and moderate fitness were strongly (3- to 6-fold risk elevation) associated with the development of CVD risk factors. They also reported a linear relationship between fitness and CVD risk factors that remained significant following adjustment for baseline BMI. There was a significant difference in treadmill time for individual non-obese (10.2 minutes) vs. obese individuals (6.7 minutes). Additionally, only 13% of non-obese participants were in the low fitness categories compared to 68% of obese participants. Being obese and having a low fitness level at baseline was associated with the development of diabetes and the metabolic syndrome in adulthood. Carnethon et al. concluded that low fitness during older adolescence has a crossover effect on developing CVD risk factors during adulthood. Interestingly, they also reported that if all unfit older adolescents had been fit, there may have been 21% to 28% fewer cases of hypertension, diabetes, and metabolic syndrome during adulthood.

Physical Activity Levels among Older Adolescents

In 2008, the U.S. Department of Health and Human (USDHSS) Services released their updated physical activity recommendation. These recommendations called for adolescents to engage in 60 minutes or more of physical activity daily (USDHSS, 2008). Participation in daily physical activity during adolescence can improve cardiovascular and muscular fitness, bone health, cardiovascular and metabolic health biomarkers, and overall body composition (USDHSS, 2008).

Troiano et al. (2008) examined the physical activity levels of children (6-11 years), adolescents (12-19 years), and adults (20+ years) using an objective measure (accelerometry) from a representative sample of the United States population. Troiano et al. examined survey data collected during the 2003-2004 National Health and Nutritional Examination Survey

(NHANES), where a total of 6,329 participants provided at least 1 day of accelerometer data and from 4,867 participants who provided four or more days of accelerometer data. Physical activity collected by accelerometers was presented in three ways: (a) mean counts per minute, (b) estimates of the time spent in physical activity according to count thresholds, (c) and an estimate of adherence to physical activity recommendations. Age, height, weight, and body mass index (BMI) were collected from the entire sample. For the purposes of this review the results for the older adolescent (16-19 years) subsample will be examined.

Troiano et al. (2008) reported that mean activity counts declined with age, with the most dramatic decline happening from childhood (646.5 $cnts \cdot min^{-1}$) to older adolescence (428.9 $cnts \cdot min^{-1}$). Activity duration also declined from childhood to older adolescence. Children (6 – 11 years) accumulated more than 1 $hour \cdot day^{-1}$ of physical activity above the moderate cut point for both males and females. By ages 16-19 years the average time dropped to 33 minutes for males and 20 minutes for females. When only bouts of activity were counted, boys and girls ages 6-11 years accumulated 45 and 26 minutes of moderate- to vigorous-intensity physical activity per day, respectively. For older adolescent participants, the corresponding time dropped to 11 and 6 minutes of moderate- to vigorous-intensity physical activity per day. Troiano et al. reported that only 6-8% of older adolescents met the physical activity recommendation of achieving 60 or more minutes of moderate- to vigorous-physical activity on 5 or more days out of 7 days. However, children met physical activity recommendations of 60 or more minutes of moderate- to vigorous-physical activity 35% of the time for females and 48% of time for males.

Troiano et al. concluded that when moderate- to vigorous-physical activity bouts of 8- to 10 minutes are considered, as appropriate for older adolescents, mean duration among those aged

16-19 years is less the 10 $min \cdot day^{-1}$ or less for both genders. This value is considerably less than the 60 minutes or more of physical activity recommended by USDHSS.

The findings reported by Troiano et al. are similar to findings published by Trost et al. (2001). Trost et al. examined physical activity differences amongst different ages and genders in a population-based sample comprised of students in grades 1 through 12 using the direct measure of accelerometry. Participants in the sample were divided into four grade groups: grades 1-3 (n = 261), grades 4-6 (n = 291), grades 7-9 (n = 332), and grades 10-12 (n = 226).

An accelerometer group was randomly selected from each group. This sample was comprised of 100 students (50 males and 50 females) from each of the grade level groups (N = 400). Middle school and high school students received their activity monitors during physical education class and were asked to wear them for 7 consecutive days, only taking them off to bath or go swimming. Time spent in moderate (3-5.9 METs) and vigorous ($\geq$ 6 METs) intensity physical activity was determined by minute-by-minute activity counts. Age specific cut points were determined using the Freedson et al. (1998) equation (METs = $2.27 + (0.0015 * counts \cdot min^{-1}) - (0.08957 * age[yr]) - (0.000038 * counts \cdot min^{-1} * age[yr])$.

Trost et al. (2001) reported that weekly 5-, 10-, and 20-minute bouts of moderate- to vigorous-physical activity (MVPA) [$\geq$ 3 METs] differed by gender. Males averaged 22.0 minutes of MVPA for sustained 5 minute bouts and females averaged 15 minutes. For bouts of MVPA that were sustained for 10 minutes and 20 minutes, the average MVPA dropped drastically for males and females, respectively (10-minute bouts; males: 6 minutes vs. females: 5 minutes) (20-min bouts; 2 minutes vs. 1 minute). No bouts of continuous vigorous-physical activity lasted longer than 3 minutes for students in grades 10-12. Trost et al. concluded that males were consistently more active than females. A significant inverse relationship between physical activity participation and grade level was observed. The older adolescent sample participated in ~50 minutes of MVPA per week. This amount of physical activity would not meet the recommendation of 60 minutes of physical activity on most days of the week.

Leslie, Forthingham, Owen, and Bauman (2000) examined physical activity patterns of adolescents aged 18-29 years from three self-report questionnaires administered throughout Australia. The researchers investigated age specific physical activity patterns consisting of walking, moderate intensity activity, vigorous intensity activity, and participation in physical activity for long-term health benefits. Three surveys used were: The Pilot Survey of the Fitness of Australians (PSFA), The Active Australia (AA), and Active Recreation on Tertiary Education Campuses (ARTEC). The PSFA was completed by 2,729 college students (mean age 24 years). The AA survey was completed by 2,500 adults aged 18-75 years. The ARTEC survey was completed by 2,300 adults, aged 20-69 years. Although surveys were completed by individuals who were older than adolescents, the paper chose to only investigate individuals aged 18-29 years.

Leslie et al. (2000) estimated energy expenditure from the frequency and duration of walking, moderate exercise, and vigorous exercise reported in the previous two weeks. Energy expenditure was presented in metabolic equivalents (METS), then multiplied by the total amount of time spent participating in activity, from the previous two weeks. These values were then summed to produce a total energy expenditure expressed as $kcal \cdot week^{-1}$. These calculations were used to classify participants as either being sufficiently active ($\geq$ 800 $kcal \cdot week^{-1}$) or insufficiently active ($\leq$ 800 $kcal \cdot week^{-1}$). The sufficient classification was equated to meet to the Australian recommendation of 30 minutes of moderate intensity physical activity on most days of the week.

For participants aged 18-24 years, there were no statistically significant differences in moderate activity, vigorous activity, or walking among the three surveys. Leslie et al. (2000) noted that both males and females reported less physical activity as they aged. For example, males and females showed a significant decrease (18-19 years: 77% of 18-19 year old males met the recommendation of 30 minutes of physical activity per day, while only 59% of 25-29 year old males met this recommendation. For females, 60% of 18-19 year old participants and only 50% of 25-29 year old participants met this recommendation. Males were also more active than females over the 2 week time period. Leslie et al. (2000) concluded that the prevalence of males and females that met the physical activity recommendation of 30 minutes or more of moderate physical activity on most days of the week significantly decreased during older adolescence. Leslie et al. did not explore activities such as organized sports or leisure time activity in which an older adolescents can participate. Activities like organized sport and leisure time activity can give older adolescents an opportunity to participate in physical activity.

Like Leslie et al. (2001), Aarnio, Winter, Peltonen, Kujala, and Kaprio (2002) continued to look at the impact leisure time physical activity has on older adolescents. Aarnio et al. (2002) investigated the consistency of leisure time physical activity in Finnish males and females between ages 16-18 years. They also examined the impact of organized sport on the stability of physical activity during these years. The researchers built their sample from an early twin cohort study; which examined the influence environment has on teenage risk factors and risk of chronic disease. Respondents completed baseline assessment questionnaires within two months of their 16th birthday. Questionnaires asked participants about health habits and attitudes toward health-related issues, such as smoking, alcohol, and physical activity. Two new questionnaires were sent to all respondents who replied to the first questionnaire at the age of 16. One questionnaire was sent after their 17th birthday, and the other was sent six months after their 18th birthday. The overall response rate for those who answered all three questionnaires was 73.6% for males and 86.5% for females. The final sample was made up of 1,338 males and 1,596 females.

Two questions pertaining to physical activity were asked in all three questionnaires. One measured frequency of leisure time physical activity. Participants could pick from six alternatives: (a) not at all or less than once a month, (b) 1-2 times a month, (c) about once a week, (d) 2-3 times a week, (e) 4-5 times a week, and (f) every day. The second question consisted of the participant's self-perception of his or her own physical fitness. Participants could choose from five alternatives: (a) very good, (b) fairly good, (c) satisfactory, (d) rather poor, and (e) very poor. Additionally, participants were asked to identify if they participated in an organized sport. Participants were asked to choose from the following sports: biking, jogging, cross-country, skiing, and swimming. Other sports available to choose from: slalom, aerobics, gymnastics, tennis, football, volleyball, badminton, baseball, basketball, rinkball, ice hockey, or skating.

Among 16 year old males, 13% participated in physical activity less than once a month, 9% 1-2 times a month, 18% once a week, 26% 2-3 times a week, 16% 4-5 times a week, and 17% daily. For females, 11% participated in physical activity less than once a month, 10% 1-2 times a month, 22% once a week, 31% 2-3 times a week, 13% 4-5 times a week, and 13% daily. Of the males who participated in regular physical activity at age 16, 47% also participated at the age of 18. For females who reported participating in regular physical activity at the age of 16, 46% did so regularly at the age of 18. Only 19% of males and 11% of females were consistent exercisers, and only 16% and 5% were persistently fit (remained in the highest categories of exercise frequency or self-perceived fitness over the three measurement occasions). Amongst both groups the most reported organized sports were: basketball, tennis, gym-training, aerobics, cross-country skiing, and jogging.

Aarnio et al. (2002) reported that 20% of males and 10% of females were persistent exercisers (i.e., very active all three-years). Leisure time activities were more widely reported amongst those individuals who participated in several types of organized sport. Aarnio et al. reported that males were more active than females and participated in more physical activity over the three year period. Among males, 35% reported that they participated in organized sport compared to 21% of females. Participants who identified that they participated in organized sport were also more likely to be categorized as persistent exercisers (males: 80%; females: 61%). Of those who did not participate in organized sports, only 6% of males and 6% of females were classified as persistent exercisers. Only 7% of males and 2% of females were classified as persistently fit. Aarnio et al. concluded that organized sport and leisure time activities can promote more physical activity amongst older adolescence. Although, organized sport can promote more physical activity amongst older adolescents; the opportunity to participate in organized sport is sometimes limited. However, leisure time activity may yield a better opportunity for older adolescents to engage in physical activity than organized and team sports.

Caspersen, Pereira, and Curran (1999) investigated age related activity patterns among adolescent and young adults (12-21 years). Caspersen et al. (1999) examined four specific activity patterns. The areas in which they focused were: leisure time physical inactivity; regular, sustained, light to moderate physical activity; strengthening activities; and stretching activities. All of these activities are covered in Healthy People 2000 and in two nationally reported surveys (National Health Interview Survey-Youth Risk Behavior Survey [NHIS-YRBS] and National Health Interview Survey-Health Promotion/Disease Prevention [NHIS-HPDP]).

Physical activity data were collected from the NHIS-YRBS for 10,645 male and female participants who ranged in age from 12-21 years. The researchers collected physical activity data for individuals aged $\geq$ 18 years from the NHIS-HPDP. The researchers adapted questions about walking, bicycling for > 30 minutes, running, jogging, or swimming for exercise, exercise to strengthen or tone muscles, and stretching exercise to represent four of the five activity patterns chronicled in this study. Physical inactivity was defined as reporting no moderate activity, vigorous activity, walking, or bicycling.

Caspersen et al. (1999) reported that the prevalence of physical inactivity increased from 6% during young adolescence (~14 years) to 24% for individuals aged 20 years. Researchers were unable to find any statistically significant sex differences at any age (except at age 17 years). For male respondents, self-reported regular sustained physical activity dropped 16 percent (40% to 24%). For female respondents, self-reported regular sustained physical activity decreased by 10 percentage points (30% to 20%). These decreases were most prevalent in individuals aged 12 to 17 years. By age 14, regular vigorous physical activity begins to decrease amongst males respondents. In comparison, the decline in moderate physical activity began at age 12 and reached bottom at age 20 (> 30 min of physical activity on most days, age 12: 60%; age 20: 28%). A comparison of physical inactivity amongst males and females revealed a greater prevalence of inactivity amongst 17 year old females (17% for females vs. 11% for males). Amongst participants aged 12 to 21 years, differences in physical activity between the ages of 12 to 21 years (males vs. females) were: physical inactivity (+10% vs. +11%), regular sustained physical activity (-16% vs. -10%), regular vigorous physical activity (-27% vs. -36%), strengthening activity (-19% vs. -21%), and stretching activity (-17% vs. -28%). The most abrupt decrease in activity occurred during older adolescence. These decreases are steady and can contribute to sedentary behavior during adulthood.

Similar results were found by Aaron et al. (1993). Where they reported a median of 23 hours·week^{-1} of leisure time physical activity amongst their male sample. Females reported an average of 7 hours·week $^{-1}$ of leisure time physical activity. Males also reported more time spent in vigorous exercise and competitive sports compared to their female counterparts. White males reported a median of 23 hours·week $^{-1}$ of leisure time physical activity. Non-white males reported a median of 19 hours·week $^{-1}$ of leisure time physical activity. The same activity difference was found amongst the females. White females reported 7 hours·week $^{-1}$ of leisure time physical activity compared to 5 hours·week $^{-1}$ for non-white females. Researchers detailed that 16% of non-white females reported participating in < 1 hour·week $^{-1}$ of activity which was higher than white females (9%). There was little difference in the number of white (1.5%) and non-white males (1.3%) who spent < 1 hour·week $^{-1}$ in leisure time physical activity.

Telama and Yang (2000) investigated the decline in physical activity amongst Finnish young people in a three year longitudinal study. Data were collected in correspondence with the Cardiovascular Risk in Young Finns study. Researchers recruited 275 males and 297 females aged 15-18 years who were randomly selected out of the national population register. Physical activity was measured using a short self-report questionnaire. The questionnaire asked participants about frequency and intensity of leisure time physical activity, participation in sport clubs, and participation in sport competition, and regular activity during leisure time. Answers for each of the five variables were summed into a physical activity index (PAI). Telama and Yang used *t*-tests to examine genders differences amongst the different age groups.

Telema and Yang (2000) reported that the frequency of physical activity (at least twice a week) decreased 16% from ages 15-18 years amongst males and 10% amongst females. The number of individuals that participated in organized sports (club sports activities) decreased by 10% amongst males and by 7% amongst females during the same time span. When assessing moderate and vigorous physical activity (MVPA), the researchers reported that the MVPA of males decreased by 10% and the MVPA of females decreased by 9% from ages 15-18. Overall, physical activity participation decreased for both groups (males 7% vs. females 4%) during older adolescent years (15-18 years).

Telema and Yang (2000) concluded that there is a steady and rapid decline in physical activity participation during an individual's older adolescent years (15-18 years). Declines were seen in all types of physical activity (leisure time activity, MVPA, and organized sport). These declines indicate that the older adolescent years should be studied further for a better understanding of the physiological and environmental influences that may be producing these large declines.

Gordon-Larsen, McMurray, and Popkins (2000) investigated the environmental and sociodemographic determinants of physical activity and inactivity amongst Caucasians, non-black Hispanics, Hispanics, and Asians. Environmental determinants were defined as any factor that could manipulate an individual's opportunity to participate in physical activity. Two factors of most interest to the researchers were the influences physical education and recreation centers on overall physical activity participation. The sample in this study was comprised of 17,766 adolescents enrolled in U.S. middle and high schools. The sample was comprised of 9,348 Caucasians, 3,933 non-Hispanic blacks, 3,148 Hispanics, and 1,337 Asians. All students in this study completed questionnaires in compliance with the 1996 National Longitudinal Study of

Adolescent Health. Outcome variables measured were moderate to vigorous physical activity (MVPA) and inactivity. Sociodemographic and environmental correlates of physical activity and inactivity were identified as availability to in-school physical education and use of community recreation centers.

Gordon-Larsen et al. (2000) reported that males participated in more moderate and vigorous physical activity when compared to their female counterparts. There were no significant ethnic differences in activity levels amongst males who participated in MVPA. However, non-Hispanic whites and Asians females participated in the highest amounts of MVPA when compared to other female samples. Non-Hispanic black males and females and Hispanic males and females reported the highest levels of inactivity. Amongst females, Asians and non-Hispanic whites reported the lowest levels of inactivity.

Gordon-Larsen et al. (2000) reported that participation in physical education programs was associated with participation in MVPA. If a student had physical education 1 to 4 times a week, then his or her MVPA was increased by 44% when compared to those who did not participate in physical education class. The same effects were found amongst students who participated in physical education 5 days a week. At age 17, participation in regular physical education (> 1 day of physical education) fell to 8% amongst older adolescents. There was no difference in the percentage of Males (15%) or females (14%) who participated in physical education at least once a week. Use of a recreation center was associated with a 75% increase in MVPA for all demographics.

Gordon-Larson et al. (2000) concluded that physical education and the use of recreational facilities provide older adolescents an opportunity to increase MVPA. However, they remarked that fewer adolescents are participating in physical education classes. Furthermore, they remarked that recreational facilities are not always readily available to some individuals and use of these facilities may be limited in some communities.

However, recreational facilities and physical education classes are readily available for individuals who enroll in college. Butler, Black, Blue, and Gretebeck (2004) reported on physical activity levels of newly enrolled female college students. Butler et al. recruited a sample of 54 female students from a Midwestern university with a mean age of 18 years. Of those 54 students, 9% were 17 years of age, with the other 91% aged 18 years. The sample included approximately 93% white, 4% Asian, Pacific Islander, or Native American, and 4% African American. The mean body mass index (BMI) of the sample was 24 (SD = 4). Researchers assessed several different fitness/physical activity variables (estimated VO_2 max, recovery heart rate, occupational sports, non-sport leisure time activity, and total physical activity). Skinfolds were taken to estimate percent body fat. Physical activity was measured using the Baecke Questionnaire of Habitual Physical Activity. VO_2 max was estimated using the Queens College 3-minute Step test. Recovery heart rate was measured after the 3-minute step test was completed. All measurements were taken within the first three days of fall classes starting and then again when participants had completed their final class of the spring semester. Butler et al. (2004) reported an increase in body weight, BMI, percent body fat, fat mass, and a decrease in fat-free mass after the spring measurements. Additionally, spring data revealed that physical activity, sports participation, and VO_2 max decreased compared to pretest data. However, there was an increase in leisure time activities.

Kimm et al. (2002) published similar results amongst their older adolescent female sample. They reported that median MET scores decreased by 83% for their entire sample over their 10-year study (whites: 31 METs per week at age 9 or 10 vs. 11 METs per week at age 18 or

19; blacks: 27 METs per week at age 9 or 10 vs. 0 METs per week at age 18 or 19). For both races, mean MET scores decreased in the older adolescent years. On average, a 1 unit increase in BMI resulted in a subsequent decrease in total METs (0.2 METs per week for black and white females). Kimm et al. concluded that physical activity decreases at a steady and drastic rate during adolescence. By age 18 or 19 years, the majority of females participated in very little to no physical activity.

These decreases in physical activity amongst older adolescents were analyzed by Pate, Heath, Dowda, and Trost (1996). The sample data was collected from the 1990 Youth Risk Behavior Survey (YRBS). Physical activity was assessed in the YRBS by the following questions: (a) On how many of the past 14 days have you done at least 20 minutes of hard exercise that made you breathe heavily and made your heart beat fast; and (b) On how many of the past 14 days have you done at least 20 minutes of light exercise that made you breathe a little more than usual and made your heart beat a little faster than usual? Students in the sample were categorized as 1,646 low active (light exercise and no days of hard exercise in the past 14 days) and 2,661 high active (6 or more days of hard and 6 or more days of light exercise).

Pate et al. (1996) reported that non-white females between the ages of 16 and 18 years were more likely to be categorized in the low active group when compared to their male counterparts. When compared to the high active group, students in the low active group reported limited participation in community-based activity (50% vs. 13%), physical education class (59% vs. 29%), and school sports (60% vs. 22%). These trends were most distinct among the female students. Students who participated in a school sport were three times more likely to be categorized in the high active group when compared to those who did not participate in a school sport (60% vs. 22%). Low physical activity was associated with many negative behavioral factors (cigarette smoking, marijuana use, poor dietary habits, television viewing, failure to wear seat belt, and perception of low academic performance). Students who reported these characteristics were 1.5 times more likely to be categorized as a low active individual compared to students who did not habitually partake in these actions.

It has been more than 10 years since the CDC/ACSM published their recommendation to accumulate at least 30 minutes of moderate intensity physical activity on most day of the week (Pate, et al. 1995). In 2008 the Department of Health and Human Services published updated recommendations for physical activity. It is recommended that adolescents accumulate 60 minutes or more of activity daily (USDHSS, 2008). Over the past three decades, the prevalence of overweight and obese adolescents has risen from 5% in 1976-1980 to 18% in 2007-2008 (USDHSS, 2008), and 65% of adolescents did not meet the 60 minutes of physical activity daily recommendation (National Youth Risk Behavior Survey Overview, 2007). With so few adolescents meeting physical activity recommendations, the need to appropriately measure physical activity and fitness is of the utmost importance. Continually, improving and creating physical activity and fitness measurement techniques can help educate our youth on the importance of being physically active and fit.

Effects of Physical Activity Levels on Aerobic Fitness

Talbot, Metter, and Fleg (2000) examined effects that leisure time physical activity (LTPA) has on VO_2 max in adults (18 – 95 years). Talbot et al. recruited their sample from the Baltimore Longitudinal Study on Aging. The sample consisted of 619 men and 497 women who completed an LTPA survey and maximal treadmill test. LTPA was self-reported by the participants using an instrument created for the study. The self-report instrument required the participants to report time spent performing 97 activities over the last 2 years. The 97 individual

LTPAs were categorized into nine major groupings for descriptive analysis. Additionally, household and childcare activities were grouped into home activities. Each reported activity was assigned a MET value based on the coding described by Ainsworth et al. (1993) and Jette et al. (1991). To counteract underestimation and overestimation of total daily time, the data for 97 activities was normalized to 1440 minutes (i.e., 24 hours). Low intensity LTPA was categorized as any activity requiring less than 4 METS, moderate-intensity LTPA included those activities requiring between 4.1 and 5.9 METS, and high-intensity LTPA included those activities requiring an energy expenditure of 6 METs or greater. Measurement of VO_2 max was completed using the Balke treadmill protocol. Additionally, weight and height were measured and used to calculate BMI.

Means, standard deviation, *t* – test, ANOVA, and regression analysis were used to statistically analyze the data. Gender differences were assessed by implementing unpaired *t* – tests. Univariate regression analysis was performed with VO_2 max as the dependent variable. The following independent variables were assessed for each gender: age, low-, moderate-, and high-intensity LTPA, total LTPA, and BMI. The contribution of LTPA, BMI, and age as they affect VO_2 max were examined by multiple regression analysis.

Talbot et al. (2000) reported a relationship between MET-minutes and VO_2 max in both sexes. However, the strength of the correlation for both men (r = .28) and women (r = .27) was less than the correlations observed for age (men: r = - .66and women: r = - .62) and BMI (men: r = - .40and women: r = - .45). There was a dose response observed, with the correlation between LTPA and VO_2 max increasing with the intensity of LTPA. High intensity LTPA was significantly correlated to VO_2 max in both men (r = .33) and women (r = .27). Moderate-intensity had a statistically significant correlation with VO_2 max in both men (r = .12) and women (r = .17). Although these correlations were statistically significant, the researchers commented that the correlations were weak in comparison to the other correlations. There was no significant relationship between VO_2 max and low-intensity activity in either sex.

For men, high-intensity LTPA accounted for 10.8% of the variance in VO_2 max, with moderate LTPA accounting for 0.9%, and low LTPA accounting for 1.5%. In women, high-intensity LTPA accounted for 6.9% of the variance in VO_2max, with moderate LTPA accounting for 2.6%, and low LTPA accounting for 1.5%. In both sexes age was the strongest independent predictor of VO_2 max, followed by BMI. Together these two variables explained 40% of the variance in VO_2 max.

Talbot et al. (2000) reported an inverse dose-response relationship between the three LTPA intensities and VO_2 max. The overall model (three intensity levels of LTPA) was predictive of VO_2 max for both men (r^2 = .42) and women (r^2 = .48). Talbot et al. (2000) concluded that age is associated with decreases in LTPA, particularly high-intensity activity. Positive correlations are observed between LTPA and VO_2 max, with high intensity (≥ 6 METS) LTPA correlating most strongly and low-intensity (< 4 METS) LTPA correlating minimally with VO_2 max. Overall, LTPA makes a modest contribution to VO_2 max after accounting for age and BMI.

Nokes (2009) recruited 275 adult females (mean age = 40 ± 3 years) and measured both physical activity and aerobic fitness. Aerobic fitness was measured using a graded treadmill test and physical activity was measured for 7 days with accelerometers. Percent body fat was measured via air displacement plethysmography (BODPOD®). Regression analysis was used to determine relationships between physical activity volume (PAv), physical activity intensity (PAi), and aerobic fitness (CRF).

Nokes (2009) reported that when PAv and CRF were treated as continuous variables, they were significantly correlated (r = .40). Additionally, when PAv levels were divided in quartiles based on lowest to highest levels of activity, the relationship between PAv and CRF remained significant (r^2 = .16). However, when age and percent body fat were accounted for, the relationship was weakened (age: r^2 = .11 and percent body fat: r^2 = .056). When PAi was controlled for a significant relationship was not seen (r^2 = .006). Furthermore, when controlling for age, percent body fat, and PAi simultaneously, PAv was not associated with CRF (r^2 = .006). Nokes also reported that CRF levels differed significantly across the three PAi groups (r^2 = .26). Participants who engaged in high intensity (physical activity at least 90 minutes per week) had a mean VO_2 max of 41 ± 7 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$. Those who participated in moderate intense PA 90 minutes or more each week had a mean VO_2 max of 34 ± 5 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$, and those in the low PAi category had a mean VO_2 max of 32 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$. Adjusting for age had little influence on the PAi and CRF association (r^2 = .26). Together PAv and PAi account for 27% of the variance in CRF. The overall model consisting of age, percent body fat, PAv, and PAi and explained 58% of the variance in CRF.

Nokes (2009) concluded that one-third of the variance of CRF is accounted for by physical activity volume and physical activity intensity and appears to be an important factor contributing to high levels of CRF. However, Nokes cautioned that physical activity only accounts for 30% of variance in CRF. Most of the differences in CRF can be accounted for by age and percent body fat and that both physical activity and aerobic fitness should be looked at as two different constructs. However, both physical activity and aerobic fitness can contribute to the prevention and treatment of many cardiovascular diseases.

Katzmarzyk, Malina, Song, Bouchard, and Claude (1998) evaluated the relationship between physical activity and health-related fitness in children and youth (9-18 years). Katzmarzyk et al. recruited participants from the greater Quebec City area as part of the Quebec Family Study. All subjects were non-obese and free of any cardiovascular problems. The sample was made up of 640 participants (356 males and 284 females). Physical activity was measured using a 3-day physical activity recall questionnaire. An estimate of moderate-to-vigorous physical activity (MVPA) was derived from the activity questionnaire. Those activities rated a 6 – 9 on the 9 point scale had a median MET value of ≥ 4.8 and were considered moderate-to-vigorous activity. Aerobic fitness was measured using a submaximal test on a cycle ergometer (PWC_{150}). The sample was separated into three age (8.50 – 12.49 years, 12.50 – 15.49 years, and 15.50 – 18.49 years) groups for statistical analysis. A canonical correlation analysis was used to assess the linear relationship between physical activity and aerobic fitness variables with each age and sex group.

Katzmarzyk et al. (1998) reported low but statistically significant correlations between MVPA and aerobic fitness (males 9-12 years: r = .28, males 16-18 years: r = .2; females 9-12 years: r = .25, females 13-15 years: r = .19, females 16-18 years: r = .27) with 13-15 year old males MVPA not correlating significantly with aerobic fitness (r = - .09). Additionally, only 11-21% of the variance in aerobic fitness was explained by physical activity. The relationship between physical activity and aerobic fitness was low to moderate for the entire sample and follows the same pattern as the aforementioned studies (Talbot et al., 2000; Nokes, 2009). Katzmarzyk et al. concluded that physical activity has a statistically significant influence on the health-related physical fitness of children and youth. However, when examining maturing youth and adolescents there is a possibility that an individual's level of physical fitness influences his or her level of physical activity rather than vice versa.

Twisk, Kemper, and Mechelen (2000) evaluated the relationship between physical activity and aerobic fitness over a 15 year period (13 yr. to 27 years). The sample was comprised of 307 subjects (148 males and 159 females) recruited for the Amsterdam Growth and Health Longitudinal Study (AGHLS). Physical activity was measured using the AGHLS interview administered questionnaire. The interview covered the previous 3 months and addressed the total time spent in all habitual activities in relation to school, work, sports, and other leisure time activities. Measured times were multiplied with the intensity of the different activities to calculate a total weighted activity score (METs). Aerobic fitness was assessed by a maximal running test on a treadmill.

Twisk et al. (2000) concluded that there was a low to moderate correlation between physical activity and aerobic fitness over the 15 year tracking period (r = .19 - .38) for both males and females alike. Twisk et al. also reported that tracking physically active people over a period of time is more difficult than tracking those who are inactive. Additionally, they reported that physical activity and VO_2 max patterns during adolescence rarely track into adulthood.

These previously reviewed publications come to a general consensus that physical activity has limited effect on aerobic fitness. This relationship is seen in youth, older adolescents, adults, and older adults. Because of the weak correlation between the two constructs a concerted effort should be made by practitioners to treat the two as separate constructs that both impact an individual's overall health.

Reliability and Validity of the 20-m Multistage Shuttle Run (PACER) as a Predictor VO_2 max

This section will summarize the literature regarding the 20-m multistage shuttle run test (MST). The initial study by Leger and Lambert (1982) reviewed is the original study that introduced the protocol for the 20-m MST. The 20-m MST is commonly called the PACER in the United States.

Léger and Lambert (1982) examined the reliability and validity of the 20-m MST as an appropriate estimation of VO_2 max. Léger and Lambert performed multiple validation studies to develop an appropriate protocol for the test. In the first series of experiments, they recruited 13 males and 12 females to complete three maximal multistage tests: (a) an inclined walking treadmill test, (b) the experimental 20-m MST test on rubber floor, and (c) 20-m MST test on vinyl-asbestos tiles (hard and low frictions types of surface). The participants completed an inclined treadmill walk test in accordance with the Balke protocol. In each of these tests VO_2 max was assessed via retroextrapolation. In the second series of test, the researchers had 66 additional subjects perform the experimental 20-m MST test on two different floor surfaces (vinyl-asbestos and rubber). In the third series of tests, 35 males and 35 females from the already established sample of 91 participants completed the final version of the 20-m MST test and a maximal multistage running track test. Predicted VO_2 max scores were compared between to the two tests. In the fourth series of tests, 27 males and 23 females completed the 20-m MST 1 week apart. The researchers found no significant sex or surface differences in the regression lines for VO_2 max achieved during the 20-m MST.

Léger and Lambert (1982) developed the following equation: VO_2 max ($ml \cdot kg^{-1} \cdot min^{-1}$) = 5.857(maximal speed attained in $km \cdot hr^{-1}$) – 19.458, for estimating VO_2 max from the 20-m shuttle run. A correlation coefficient of .84 between measured and estimated VO_2 max was reported for this equation. The equation yielded a standard error of estimate of 5.4 $ml \cdot kg^{-1} \cdot min^{-1}$. The researchers reported that VO_2 max measured directly during the last stage of the Balke protocol (42.7 $\pm$ 10.0 $ml \cdot kg^{-1} \cdot min^{-1}$) was similar ($r$ = .91 and SEE = 4.16 $ml \cdot kg^{-1} \cdot min^{-1}$) to VO_2 max measured directly at the end of the 20-m MST (43.9 $\pm$ 9.2 $ml \cdot kg^{-1} \cdot min^{-1}$). Means for the two tests (VO_2 max direct measure and 20-m MST) were not statistically different. The correlation between maximal multistage run test performance and the 20-m MST yielded high correlations (r = .92 and SEE 2.63 $ml \cdot kg^{-1} \cdot min^{-1}$). Léger and Lambert reported that the reproducibility of the 20-m MST test was high (r = .975 and SEM = 2.00 $ml \cdot kg^{-1} \cdot min^{-1}$).

Léger and Lambert (1982) concluded that the 20-m shuttle run could predict VO_2 max with an r of .84 and a standard error of 5.4 $ml \cdot kg^{-1} \cdot min^{-1}$. As well as being a reliable and valid test for predicting VO_2 max the 20-m MST can be completed at low cost and with little inconvenience to the participant completing the test.

Cooper, Baker, Tang, and Hansford (2005) investigated the reliability and validity of the 20-m multistage (20-m MST) fitness test to estimate VO_2 max. The sample was comprised of 30 male undergraduates (M ± SD: age = 21.8 ± 3.6 years, body mass = 76.9 ± 10.7 kg, and height = 1.76 ± 0.05 m). Researchers conducted the study in two phases. Phase 1 examined the test-retest reliability of the multistage fitness test. Researchers randomly selected 21 students from the original 30 to complete the multistage fitness test twice. In phase 2, all 30 subjects performed both the 20-m MST and a treadmill VO_2 max test. To avoid diurnal variation, participants completed each phase of the study during the same time of the day.

Participants completed the VO_2 max test on a treadmill and were asked to run at a speed of 3.06 $m \cdot s^{-1}$ and zero (0%) gradient to begin the test. Every 3 minutes the gradient of the treadmill was increased (2.5°). The increase in gradient continued until the participant was physically unable to continue and instructed the researcher to terminate the session. During the last minute of each 3 minute exercise period, expired gases were collected via a Douglas bag. Gases collected in each Douglas bag were analyzed to determine total oxygen uptake. Achieved VO_2 max was determined by the three criteria established by the British Association of Sport and Exercise Sciences: (a) subjective fatigue and volitional exhaustion, (b) a plateau in the oxygen uptake/exercise intensity relationship, and (c) a final respiratory exchange rate of 1.15 or above. Maximal oxygen uptake was expressed relative to body mass for each subject ($ml \cdot kg^{-1} \cdot min^{-1}$).

All 30 participants also completed the 20-m MST. The 20-m MST consisted of a shuttle run between two parallel lines set 20 meters apart. An audio recording determines the speed at which the participant is to complete each lap. The initial running speed is set at 2.36 $m \cdot s^{-1}$ and is increased by 0.14 $m \cdot s^{-1}$ every minute. The increase in running speed is comparable to the change in gradient during a VO_2 max test on a treadmill. Students ran in groups of five to add an element of competition and to ensure maximal effort. The 20-m MST results were presented as predicted VO_2 max ($ml \cdot kg^{-1} \cdot min^{-1}$).

In phase 1 of the study the agreement between repeat performances of the 20-m MST (test-retest) was determined by establishing 95% limits of agreement (LoA). Researchers transformed data into natural logarithms and recalculated the limits of agreement. This gave researchers the ability to express their results as mean bias ± the 95% agreement component (random error). In phase 2 the criterion-related validity of the multistage fitness test was investigated. Predicted VO_2 max from the multistage fitness test was compared to measured VO_2 max.

Cooper et al. (2005) reported that mean multistage fitness test performance was 52.3 ± 8.8 $ml \cdot kg^{-1} \cdot min^{-1}$ and the retest mean multistage fitness test performance was 53.3 ± 8.9 $ml \cdot kg^{-}$

$^{1} \cdot min^{-1}$. A dependent group's *t*-test showed no significant difference between the two means. Test-retest residual error scores were normally distributed and the bias ± 95% LoA was -0.4 ± 2.7 $ml \cdot kg^{-1} \cdot min^{-1}$.

Mean measured VO_2 max was 57.5 ± 4.5 $ml \cdot kg^{-1} \cdot min^{-1}$ and the mean predicted VO_2 max was 55.7 ± $ml \cdot kg^{-1} \cdot min^{-1}$. Residual errors were normally distributed and the mean bias ± LoA (1.8 $ml \cdot kg^{-1} \cdot min^{-1}$ ± 3.2 $ml \cdot kg^{-1} \cdot min^{-1}$) was statistically significant. Researchers computed a zero order correlation between test and retest results from phase 1 and found a high correlation (r = .99). For phase 2, researchers computed the zero order correlation between predicted VO_2 max and measured VO_2 max. The correlation was moderately high (r = .79) and statistically significant. The effect size for phase 2, revealed a small to medium size difference between means (d = 0.40). Cooper et al. (2005) reported that mean MFT predicted VO_2 max was 3.1% lower than mean measured VO_2 max.

Ransbottom, Brewer, and Williams (1988) examined the validity of the 20-m MST to estimate VO_2 max amongst an adult population. Participants 20-m MST results were compared to their results from a 5 km performance run (PR). Ransbottom et al. (1988) studied a sample of 74 volunteers (36 men, 38 women; aged 19-36 years). Physical characteristics of the sample were as follows: (mean ± SD) height 176.0 ± 8.1 cm (men) and 168.1 ± 6.2 (women); weight 70.9 ± 7.7 (men) and 64.1 ± 9.5 kg (women).

All participants completed three different runs in a random order (VO_2 max performance test, a 20-m MST, and a 5 km time trial). VO_2 and heart rate were collected every minute. Heart rate was collected using an oscilloscope. The 20-m MST was performed in accordance with the Eurofit Provisional Handbook. The 5 km run took place on a collegiate track. Participants were instructed to run their fastest 5 km time. Time was assessed using a digital stop watch. The difference in means was compared using a *t*-test and the correlation was determined using the Pearson product moment correlation coefficient.

VO_2 max was predicted from the 20-m MST by using the regression equation provided by Léger and Lambert (1982). Final running speed was directly correlated with VO_2 max values [r = .82 (men), r = .85 (women), and r = .90 (men and women)]. Ransbottom et al. (1988) a high correlation between estimated VO_2 max from 20-m MST and directly measured VO_2 max (r = .92). Ransbottom et al. reported that VO_2 max estimates produced by the 20-m MST and by the 5 km performance run correlated highly (20-m MST, r = .93; 5 km PR, r = -.94) with the VO_2 max criterion. Both men and women were able to attain 97% of their laboratory defined maximal heart rate during the 20-m MST. Ransbottom et al. concluded that the 20-m MST was a valid test for estimating VO_2 max amongst adults. On average, VO_2 max estimated by the 20-m MST deviated from measured VO_2 max by 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$.

Paliczka, Nichols, and Boreham (1987) investigated the validity of the 20-m MST as a predictor of VO_2 max and competitive performance in an endurance event. Paliczka et al. collected data on nine male endurance runners (mean age = 35.4 ± 5.8 years). The first test completed was the 20-m MST, followed 2 days later by the 10 km run, and eventually each participant completed a VO_2 max test. The VO_2 max test was completed within 2 weeks of the 10 km run. The 20-m MST was completed using the protocol established by Léger and Lambert (1982). Velocity was set at 8.5 $km \cdot hr^{-1}$ for the first minute and increased by 0.5 $km \cdot hr^{-1}$ every minute thereafter. Participants were asked to run for as long as possible. Participants stopped when they were physically exhausted or the test was ended when they failed to be within 3 meters of the end line on two consecutive tones. VO_2 max was assessed via indirect calorimetry during a treadmill run. Each participant was given a 3 minute warm-up. The speed was set to the participant's average speed from his 10 km run. Correlations were calculated using the Pearson product moment coefficient.

Average VO_2 max for the sample was 59 ± 10 $ml \cdot kg^{-1} \cdot min^{-1}$ and the average number of laps (paliers) completed by the sample during the 20-m MST was 105 ± 24 (11 ± 3 paliers). The average time it took the sample to complete the 10 km run was 42 minutes. Paliczka et al. (1987) reported that correlations between measured VO_2 max and 20-m MST performance (r = .93) and the 10 km run (r = -.90) were high. Paliczka et al. concluded that the 20-m MST was a valid means of predicting VO_2 max when direct measurement of VO_2 max is not readily available.

Barnett, Chan, and Bruce (1993) examined the validity of the 20-m MST as a predictor of VO_2 max in 12- to 17-year-old Hong Kong Chinese secondary students. They also compared estimates produced by their prediction equations to values established by an accepted prediction equation. Barnett et al. recruited 55 secondary school students (27 males and 28 females) to perform the 20-m MST during their physical education classes. Running speed was 8 $km \cdot hr^{-1}$ for the first minute and 9 $km \cdot hr^{-1}$ for the 2nd minute. For every minute after the initial two minutes, speed increased by 0.5 $km \cdot hr^{-1}$. VO_2 max was measured on a treadmill using a continuous protocol. Anthropometric measures were taken on each student [age, height, weight, and skinfolds (triceps and calf)]. The treadmill test protocol consisted of 2-minute stages at 9, 9.5, and 10 $km \cdot hr^{-1}$ and every 2 minutes the grade was increased 2.5% to 3%. Expired air was collected via an on-line computerized system with data recorded in 20-s intervals. The researchers made sure that participants achieved at least two of the following criteria for a maximal exertion test: (a) voluntary exhaustion, (b) heart rate greater than 90% of predicted maximum (220 – age in years), and (c) a respiratory exchange ratio (R) greater than 1.0.

The initial analysis was conducted to examine the correlation between measured VO_2 max and VO_2 max predicted from the formula provided by Léger et al. (1982) and root mean square error was calculated. Variables plugged into the regression analysis for the prediction equations were: age, height, weight, triceps skinfold, calf skinfold, maximal speed, total shuttles completed, and gender.

Barnett et al. (1993) reported a significant correlation (r = .72) between measured VO_2 max and VO_2 max predicted from the 20-m MST. The mean difference $\underline{+}$ SD of VO_2 max using the Léger et al. equation was 2.0 ± 4.9 $ml \cdot kg^{-1} \cdot min^{-1}$. The equation created from maximal shuttle run speed produced a moderate correlation (r = .74, SEE = 4.6 $ml \cdot kg^{-1} \cdot min^{-1}$) for estimation of VO_2 max for this sample. The equations that best predicted VO_2 max in this sample were as follows: (a) VO_2 max = 28.3 – 2.1(sex) – 0.7(triceps skinfold) + 2.6(maximum speed) [n = 55, r = .85, SEE = 3.7.]; (b) VO_2 max = 25.8 – 6.6(sex) – 0.2(weight) + 3.2(maximal speed) [n = 55, r = 0.84, SEE = 3.7,]; and (c) VO_2 max = 24.2 – 5.0(sex) – 0.8(age) + 3.4(maximal speed) [n = 55, r = 0.82, SEE = 4.0]. Values for each variable were as follows: sex (0 = male, 1= female), age (years), weight (kg), triceps skinfold (mm), and maximal speed in $km \cdot hr^{-1}$ (the speed of the last stage completed during the MST).

Barnett et al. (1993) concluded that the inclusion of gender into the regression equation to predict VO_2 max from 20-m MST improved the accuracy of the prediction. They also concluded that a good prediction of VO_2 max can be obtained by the 20-m MST for 12-17 year old Chinese secondary students. The regression equation that took triceps and calf skinfolds into account produced a similar standard error of estimate to the one produced by the Leger and Lambert (1982) equation. Barnett et al. suggested that the following equation may be preferred by practitioners because it does not require any measurement in addition to the recording of maximum speed during the MST [VO_2 max = 24.2 – 5.0(sex) – 0.8(age) + 3.4(maximal speed)].

Liu, Plowman, and Looney (1992) investigated the test-retest reliability of the 20-m MST, the concurrent validity of the 20-m MST, and the validity of the Léger et al. (1982) 20-m MST equation to predict VO_2 max for their young adolescent sample. Liu et al. recruited 62 students between the ages of 12 and 15 years. Twenty participants (12 males and 8 females) participated in the test-retest reliability study. Forty-eight of the participants participated in the validity study and six participants completed both the reliability and validity study.

Participants completed the 20-m MST as designed by Léger et al. (1982). The starting speed was 8.5 $km{\cdot}hr^{-1}$ and increased by 0.5$\cdot km{\cdot}hr^{-1}$ every successive minute after the first. Participants were instructed to run in between two lines 20-m apart until they were physically exhausted or could no longer continue. The number of laps completed was used for statistical analysis. Participants in the validity study ran the 20-m MST when they arrived at the laboratory and were given as much rest time as needed before the treadmill test. Treadmill speed was set at 5 mph (8 $km{\cdot}hr^{-1}$) for females and 6 mph (9.6 $km{\cdot}hr^{-1}$) for males, with a constant grade (5%) maintained for both males and females throughout the session. Oxygen consumption was collected using a metabolic measurement cart. Metabolic data were analyzed in 20 second intervals. Heart rate was recorded via a heart rate monitor and recorded each minute.

Test-retest reliability was determined using an intraclass correlation (one-way ANOVA model for a single trial). Independent *t*-tests were used to identify any significant sex differences on all variables. Concurrent validity was determined with a correlation coefficient between the number of laps and measured VO_2 max. The error of predicting VO_2 max from the Léger et al. (1982) equation was determined using the standard error of estimate and total error. A paired *t*-

test was used to identify any significant differences between predicted VO_2 max and directly measured VO_2 max.

The intraclass correlation coefficient for the number of laps completed during the first shuttle run and the second shuttle run was high (R = .93). Reliability by sex was also high (males R = .91; females R = .89). The number of laps completed during the 20-m MST was moderately and significantly correlated with directly measured VO_2 max (males n = 22; r = .65; females n = 26; r = .51). The correlation for males and females together was .69. The correlation between measured and estimated VO_2 max from 20-m MST was .72 (SEE = 5.27 $ml \cdot kg^{-1} \cdot min^{-1}$). The Total Error reported was 5.40 $ml \cdot kg^{-1} \cdot min^{-1}$. These findings are similar to results reported by Léger et al. (1982) (r = .71). Liu et al. (1992) concluded that the 20-m MST provided valid and reliable estimates of VO_2 max in 12-15 year old adolescents.

Mahoney (1992) investigated the validity of the 20-m MST in a non-Caucasian sample of 103 students (53 boys and 50 girls) from an inner-city school in the United Kingdom. A few white children were included, but accounted for less than 5% of the total sample. The sample was mainly comprised of Indian, African American, Caribbean, and mixed races. Height, weight, and skinfolds (biceps, triceps, subscapular, and suprailiac) were measured during the initial meeting.

Participants completed the 20-m MST in accordance with the Eurofit protocol. Boys and girls were tested separately in groups of 10-15. Gender separation was chosen because traditional physical education classes in the UK were gender specific. Results from the 20-m MST were compared to directly measured VO_2 max. Female participants completed the Balke treadmill walking test protocol and male participants completed a jogging protocol where participants started at 8.0 $km\cdot hr^{-1}$ and increased to 9.4 $km\cdot hr^{-1}$ followed by an increase of 5% gradient every 2 minutes. A subsample of 20 students completed the 20-m MST twice to assess reliability.

Mahoney (1992) reported that the test-retest reliability of the 20-m MST for his sample was high (females: $r = .89$; males: $r = .73$). Correlations between VO_2 max estimated from 20-m MST and directly measured VO_2 max were moderately high to high (females: $r = .76$; males: $r = .83$). Mahoney concluded that the 20-m MST was a reliable and valid field test to estimate VO_2 max in non-Caucasian young adolescent students.

Lamb and Rogers (2007) investigated how test-retest reliability of the 20-m MST is affected when participants are given multiple attempts to complete the test. Lamb and Rogers recruited 22 males and 13 female athletes from a university population. The collaborating sports that were represented were: rugby ($n = 14$), football ($n = 13$), and tennis ($n = 8$). The study incorporated a cross-sectional design with a single cohort engaged in repeated measures. All participants were assigned to one of three groups depending on the type of sport they played. Two of the three groups completed the 20-m MST in an indoor gymnasium. A third group completed the 20-m MST in an indoor dance studio. The 20-m MST was administered in accordance with the protocol established by Brewer et al. (1988). The total number of shuttles completed by each participant was recorded and implemented into the equation established by Léger et al. (1982) to predict VO_2 max. Valid stages and shuttles were also recorded and incorporated into a table established by Brewer et al. to predict VO_2 max.

Means and standard deviations were calculated for the number of laps completed and predicted VO_2 max scores over the three trials. Variability between VO_2 max and total number of shuttles completed was assessed with two-way repeated measures ANOVA. Reliability between repeated trials (trial 1 vs. trial 2, trial 2 vs. trial 3, and trial 1 vs. trial 3) was conducted by establishing 95% limits of agreement and Bland-Altman analysis. To compare results to previous studies, Pearson correlations and intraclass correlations were computed.

Lamb and Rogers (2007) concluded that the test-retest intraclass correlation of the 20-m MST was high (*R* = .94 to .96). Lamb and Rogers also concluded that participants may benefit from a practice run before their actual test scores are recorded.

Sproule, Kunalan, McNeil, and Wright (1993) investigated the validity of the 20-MST in 20 undergraduate students from Singapore (males *n* = 16; female *n* = 4). Students were 20-23 years old. Height, weight, and percent fat were measured. The researchers had participants complete the 20-MST outdoors on a flat, unshaded, plexipave surface. The 20-MST was administered following the protocol established by Léger et al. (1988). Along with running two 20-MST, participants completed two VO_2 max tests following the Taylor (1963) protocol. Pearson product moment correlations were used to estimate test-retest reliability. A *t*-test was used to examine mean differences between measured VO_2 max and predicted VO_2 max from the 20-MST.

Sproule et al. (1993) reported that test-rest reliability for the 20-MST was high (*r* = 91). *T-test* results indicated a significant difference between means for measured VO_2 max and estimated VO_2 max. Estimated VO_2 max from the 20-MST was lower for the whole sample (49 ± 7 $ml \cdot kg^{-1} \cdot min^{-1}$ vs. 52 ± 6 $ml \cdot kg^{-1} \cdot min^{-1}$) when compared to measured VO_2 max.

McVeigh, Payne, and Scott (1995) investigated the reliability of the 20-m MST in 13- and 14-year-old Scottish children on multiple days. Furthermore, McVeigh et al. examined the validity of the 20-m MST as a predictor of VO_2 max in 13- and 14-year old students. Finally, the researchers examined the different anthropometric measures (height, weight, and skinfold thickness measures) that could improve the prediction of VO_2 max from the 20-MST. The researchers recruited 33 children (15 males, 18 females) from a state comprehensive high school in Edinburgh, Scotland. McVeigh et al. used a repeated measures design to assess the reliability of VO_2 max and the 20-m MST test. On each of the three testing days, height, weight, and skinfold thicknesses were recorded. Reliability of these measures was assessed. Each child performed three laboratory tests of VO_2 max and three 20-m MST.

VO_2 max was assessed using a continuous incremental test protocol completed on a treadmill. Oxygen uptake was measured every 30 seconds using an online-gas analyzer. To determine if students had reached VO_2 max, two of three physiological criteria needed to be met: (a) maximal heart rate greater than 95% of predicted maximal heart rate (220 – age), (b) a respiratory exchange rate greater than unity, and (c) signs of fatigue (e.g., inability to keep pace with the ergometer, excessive sweating, facial flushing, dyspnea, and unsteady gait). The 20-m MST was completed according to the Léger et al. (1982) protocol. Intraclass correlations were calculated to assess reliability of all tests taken over the 3 day testing period. *t*-tests were used to determine any significant sex differences. Multivariate regression analysis was used to create a prediction equation to estimate VO_2 max. The researchers incorporated height, weight, sex, and skinfold thickness as predictors of VO_2 max within their regression equation.

Reliability estimates of all measures for the 3 days were calculated. McVeigh et al. (1995) reported that the intraclass correlation coefficients for height, weight, and skinfold thickness were high (> .94). Intraclass correlations for VO_2 max and 20-m MST were lower .72 and .75, respectively. McVeigh et al. reported that a good predictor of VO_2 max was maximal running speed attained during the three trials (males R^2= .60, SEE = 3.46 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$; females R^2 = .62, SEE = 3.66 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$). When the 20-m MST was performed once the maximal shuttle run speed accounted for 42% (SEE = 4.57) and 36% (SEE = 4.42) of the variance in VO_2

max in females and males, respectively. Prediction of VO_2 max for males and females was improved by the addition of skinfold thickness measures into the regression equation. Amongst the girls, the most significant increase in prediction ability was achieved when incorporating maximal recorded speed and triceps skinfolds thickness (R^2 = .85, *SEE* = 2.40 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$). The best correlation for boys was produced when maximal shuttle speed and sum of triceps and subscapular skinfolds (R^2 = .68, *SEE* = 3.23 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$) were included.

McVeigh et al. (1995) concluded that the inclusion of skinfold measurement into the 20-m MST prediction equation can strengthen the estimation of VO_2 max for both males and females. They also concluded that 20-m MST was a reliable field test for participants aged 13 to 14 years. Intraclass correlations between directly measured VO_2 max and the 20-m MST were moderate (r = .72 – .75).

Pitetti, Fernhall, and Figoni (2002) investigated whether the 20-m MST regression formula developed for students with mental retardation was valid for students without any mental disabilities. The secondary purpose of the study was to compare the equation created for students with mental retardation by Fernhall et al. (2000) to the Léger et al. (1982) equation established for children and adolescents. The researchers recruited 51 healthy, non-disabled children and adolescents (females, n = 38, males, n = 13) ranging in age from 8 to 15 years. Students completed two tests for VO_2 max and the 20-m MST. A modified Bruce treadmill protocol was used for the VO_2 max test, with the starting speed determined by the participant's fitness level (2 to 3.5 mph or 3.3 to 5.8 $km{\cdot}hr^{-1}$). The 20-m MST was run following the protocol established by the FITNESSGRAM® (Welk, Morrow, & Fall, 2002).

The equations examined by Pitetti et al. (2002) were: (a) VO_2 max [$ml{\cdot}kg^{-1}{\cdot}min^{-1}$] = 0.35 x [# of laps] – [0.59 x BMI] – [4.61 x sex (male = 1, female = 2) + 50.6 (Fernhall et al., 2000);

and (b) VO_2 max [ml·kg^{-1}·min^{-1}] = 31.025 + [3.238 x [speed of level stopped in km·hr^{-1}] – [3.248 x [top speed km·hr^{-1}] – 3.248 x age [yrs] + 0.1536 x [speed x age] (Leger et al., 1988). Test-retest reliability estimates for the maximal treadmill test and 20-m MST were both high (*r* = .89 and .86). Pitetti et al. reported a correlation between measured VO_2 max and estimated VO_2 max (from laps completed, age, weight, height, and BMI) was moderately high (*r* = .53). When comparing the estimates of VO_2 max from the regression equations established by Fernhall et al. and Léger et al. to measure VO_2 max the researchers reported that both estimates were moderately correlated (Fernhall et al., *r* = .66) (Léger et al., *r* = .57). However, when the estimates of VO_2 max were correlated with each other they produced a higher correlation of .78.

Mahar, Welk, Rowe, Crotts, and McIver (2006) developed an equation to estimate VO_2 max from the 20-m MST (i.e., PACER) when considering laps completed, gender, and body mass or body mass index (BMI) in 12- to 14-year old males and females. Additionally, Mahar et al. cross-validated the Léger et al. (1988) model to compare its accuracy with their newly developed model.

Mahar et al. (2006) recruited 74 females and 61 males aged 12 to 14 years from two different sites. Age, gender, and body composition (height, weight, BMI, and percent body fat) were collected. All participants underwent a graded treadmill test to volitional exhaustion to determine VO_2 max. Two different treadmill speeds were held constant for females and males, respectively. Females performed the test at a speed of 5.0 mph, males completed the test at a speed of 5.5 mph. At the beginning of the second minute, the treadmill grade was increased to 2%. Every subsequent minute thereafter the grade of the treadmill was increased by an additional 2%. VO_2 was assessed with a metabolic measurement system. VO_2 max was accepted if the participants met three conditions: (a) showed signs of intense effort (e.g., hyperpnea, facial flushing, grimacing, unsteady gait, sweating); (b) peak heart rate reached a value at least 95% of maximal heart rate as predicted by age; and (c) respiratory exchange rate (RER) was at least 1.0. The PACER was administered following standards published by FITNESSGRAM® (Welk, Morrow, & Fall, 2002).

Multiple linear regression was used in the validation sample to estimate VO_2 max from laps completed on the PACER, gender, and body mass or BMI. The equation developed on the validation sample was applied to the cross-validation sample. Correlations between measured VO_2 max and estimated VO_2 max from the validation sample were calculated. Standard errors of estimate and total errors were calculated for the cross-validation sample. Residual error was correlated with estimated VO_2 max to check for accuracy.

Mahar et al. (2006) reported that when performance on the PACER, gender, and BMI or percent body fat are accounted for or the multiple correlation (*R*) are .65 - .66 with an SEE of 6.4 $ml \cdot kg^{-1} \cdot min^{-1}$. These correlations were found on both the cross-validation and total sample. The mean differences between measured and estimated VO_2 max were less than 2.5 $ml \cdot kg^{-1} \cdot min^{-1}$. The two regression equations that took body mass and BMI into account had a higher correlation (r = .62 - .64) and lower standard errors of estimate than the Léger et al. (1982) equation (r = .53 - .54). All predictor variables applied in the regression equation made statistically significant contributions to the prediction of VO_2 max. Mahar et al. concluded that the regression equation (VO_2 max = 47.438 + [PACER laps*0.142] + (Gender [m=1 , f=0]*5.134) – (body mass [kg]*0.197) developed for the 12- to 14-year-old sample was more accurate than the Léger et al. (1988) model with easier conversion of PACER laps completed to VO_2 max.

Summary

The PACER is a valid and reliable test for predicting VO_2 max for an adolescent population (Léger & Lambert, 1982); Barnett et al., 1993); Liu et al., 1992). However, recent research has suggested that the commonly used Léger et al. (1988) prediction equation may produce spurious results for some populations (Mahar et al., 2006). Mahar et al. concluded that a regression model that considers PACER performance, gender, and body mass can produce a higher correlation and lower SEE for 12- to 14-year-old adolescents when compared to the Léger et al. (1988) equation. Therefore, the creation of a regression model for older adolescents may or may not produce a better prediction of aerobic fitness in an older adolescent population when compared to previous published models.

Methods

Participants

Forty eight 17-18 year old participants (26 males, 22 females) were recruited from the Pitt County, North Carolina area. Participants were recruited via class presentations, study flyers, e-mail, and word of mouth. Prior to the first visit participants were given written instructions not to eat or drink up to 3 three hours before their visit. At the first visit, all participants were provided informed consent and/or assent. Parental consent was obtained for all participants under 18 years of age. The informed consent described the purpose, procedures, risks, and benefits of the study. All methods and procedures were approved by the University Medical Center Institutional Review Board of East Carolina University.

Protocol

Testing took place in two 1-hour sessions. During the first visit, all participants were habituated to treadmill exercise and the PACER test as needed, completed body composition measures, and completed one of the two aerobic fitness measures. The second aerobic fitness measure was completed during the second visit. The order of the aerobic fitness measures was counterbalanced.

Participants were instructed to meet at the East Carolina University Activity Promotion Laboratory, where all measurements were taken. The following body composition measurements were taken: height, weight, skinfolds, bioelectrical impedance analysis (BIA), and air displacement plethesmography via the BOD POD®. Participants were instructed to wear minimal clothing while body composition measures are assessed (swim suit or compression shorts, swim cap, and sports bra for females). All measurements were conducted by a trained research assistant.

Body Composition Measurement

Air displacement plethesmography (ADP). The BOD POD was used to measure body composition within this study. Two hours prior to the assessment, the BOD POD was calibrated according to the procedures recommended by the BOD POD manufacturer. Body weight was assessed using a calibrated scale attached to the BOD POD. Raw body volume was measured as participants sat in the BOD POD while breathing normally following a bar graph on a computer screen. This measurement was repeated twice. A third test was required if two consecutive trials are not within 200 ml. Thoracic gas volume (TGV) was measured to ensure a more accurate prediction of percent body fat. TGV was measured by a breathing circuit system on the BOD POD. Participants were asked to breathe normally for about 20 seconds. Participants continually breathed through a tube connected to the wall of the BOD POD and were asked to exhale with three smooth breaths when the screen instructs them to ‘huff, huff, huff’. TGV measurements were repeated until the BOD POD accepts the breath sequence yielded by the participant. Three developed equations for predicting percent body fat were used. The Siri (1961) equation was used for Caucasians, the Wagner and Heyward (2001) equation was used for African Americans males, and the Ortiz et al. (1992) equation was used for African American females.

Skinfold measurements. Body composition was assessed using Lange skinfold calipers (Cambridge, MD). Five sites were assessed for males (triceps, calf, chest, abdomen, thigh) and four sites were assessed for females (triceps, calf, suprailium, thigh). Percent fat was estimated using the equation established by Slaughter et al. (1988).

Aerobic Fitness Measurement

Progressive Aerobic Cardiovascular Endurance Run (PACER). The PACER 20-m multi-stage shuttle run was completed following standard procedures (Welk, Morrow, & Falls, 2002). Participants ran from one starting line to a subsequent line 20-m apart, while keeping pace with a prerecorded audio cadence. The cadence was set to music and started slowly and progressively got faster throughout the run. Participants were asked to keep up with the cadence for as long as they physically could. The test was terminated when a participant was unable to complete two consecutive laps in the allotted time or could not physically continue. The number of laps completed was recorded.

Maximal Treadmill Test. Prior to the treadmill test, all participants were familiarized with the treadmill exercise. All participants underwent a graded exercise test to exhaustion on a Trackmaster (model TMX425C, Carrollton, TX) treadmill to determine maximal oxygen consumption (VO_2 max). Participants completed the maximal treadmill test following the Bruce protocol (1963).

VO_2 was assessed using a COSMED portable metabolic system (model K4b^2, Rome, Italy). Prior to the test, the COSMED was calibrated with known sample gases. Maximal effort was accepted if the participant met two of three criteria (Armstrong & Welsman 1994; Rowland, 1993): (a) shows signs of intense effort (e.g., hyperpnea, facial flushing and grimacing, unsteady gait, sweating), (b) peak heart rate reaching a value at least 90% of maximal heart rate as predicted by age, and (c) a respiratory exchange ratio (RER) of at least 1.0.

Statistical Analysis.

Multiple linear regression analysis was used to predict VO_2 max from the number of laps completed on the PACER, age, gender, and body mass. For comparison purposes, cross-validation was conducted on several published prediction models. Prediction error was assessed with two equations. The standard error of estimate (*SEE*) was calculated as: $SEE = S_Y\sqrt{1 - R^2_{YY'}/N}$. Total Error was calculated as: $TE = \sqrt{\sum (Y - Y')^2/N}$. For these equations, Y is measured

VO_2 max and Y' is VO_2 max estimated from the prediction equation. Comparisons of these two error estimates help establish the overestimation and underestimation of each prediction model. The final analyses were conducted in a criterion-referenced framework. Estimated VO_2 max for each participant from each of the previously developed models was classified according to the new FITNESSGRAM® standards (Welk et al. 2010).

Results

Participants included in the analyses were 22 females and 26 males aged 17 to 18 years who were recruited from the Pitt County, North Carolina area. Physical characteristics of the participants are presented in Table 1. Visual inspection of the relationship between VO_2 max and PACER laps identified no outliers in this sample and all participants' (*N* = 48) data were used in the final analyses. Mean scores for VO_2 max for both males (41.9 ± 9.9) and females (33.1 ± 6.7) were lower than the 2010 FITNESSGRAM® Healthy Fitness Zone standards (Males: ≥ 44.2 $ml \cdot kg^{-1} \cdot min^{-1}$ and Females: ≥ 38.6 $ml \cdot kg^{-1} \cdot min^{-1}$). Significant differences were found between males and females in height, body mass, percent fat measured by BODPOD, PACER, VO_2 max, and maximal treadmill time. Mean BMI and percent body fat measurements for both males and females fell within the Healthy Fitness Zone (HFZ - BMI: 16.8 to 25.2 $kg \cdot m^{-2}$; HFZ - percent fat: males 18.9 - 22.3% and females 20.9 - 31.4%).

Table 1. *Means* ± *SD* of Physical Characteristics and PACER 20-m Shuttle Run Performance

Variable	**Total (*n* = 48)**	**Males (*n* = 26)**	**Females (*n* = 22)**
Age (years)	17.6 ± 0.5	17.5 ± 0.5	17.7 ± 0.5
Height (cm)	172.4 ± 8.2	178.3 ± 5.4	165.5 ± 4.7*
Body mass (kg)	70.3 ± 13.5	76.5 ± 11.2	62.8 ± 12.3*
BMI ($kg \cdot m^{-2}$)	23.6 ± 4.2	24.1 ± 3.3	23.1 ± 5.0
Percent fat (%)	19.9 ± 8.9	14.8 ± 6.4	25.8 ± 7.8*
PACER (# laps)	42.4 ± 20.6	52.4 ± 20.2	30.5 ± 13.9*
VO_2 max ($ml \cdot kg^{-1} \cdot min^{-1}$)	37.9 ± 9.6	41.9 ± 9.9	33.1 ± 6.7*
Maximal heart rate ($b \cdot min^{-1}$)	193.5 ± 10.0	190.7 ± 11.9	196.8 ± 6.1
Maximal RER	1.26 ± 0.15	1.26 ± 0.15	1.26 ± 0.14
Maximal Treadmill time (min)	10.8 ± 2.2	11.8 ± 2.1	9.5 ± 1.4*

Note: * Statistically significant ($p < .05$) difference between male and female samples.

Estimation of VO_2 max

Multiple regression was used to predict VO_2 max from the number of laps completed on the PACER, age, gender, body mass, and BMI or percent fat. The VO_2 max-PACER scatterplot was examined (see Figure 1) for linear and quadratic trends. The quadratic term did not add significantly to the estimation of VO_2 max; therefore, only linear models were further examined. In addition, two interaction terms (i.e., age x gender; gender x PACER laps) were entered into the model to determine if they significantly contributed to the prediction. These two interactions did not contribute significantly to the prediction of VO_2 max and were disregarded. Correlations between VO_2 max and the predictor variables are presented in Table 2. The correlation between

VO_2 max and PACER laps completed was high and accounted for 78% of the variance in VO_2 max. Gender (26% of the variance in VO_2 max), BMI (12% of the variance in VO_2 max, and percent fat (41% of the variance in VO_2 max) were also significantly ($p < .05$) correlated with VO_2 max. However, age, and body mass were not significantly correlated with VO_2 max.

Table 2. Zero-Order Correlations between Measured VO_2 max and Predictor Variables for Total Sample ($N = 48$)

Variable	Correlation
PACER (# laps)	.89*
Age (years)	.06
Gender (1=male, 0=female)	.46*
Body mass (kg)	-.07*
BMI ($kg \cdot m^{-2}$)	-.34*
Percent fat (%)	-.65*

* $p < .05$

Results of the multiple regression analysis revealed that PACER laps contributed significantly to the estimation of VO_2 max. After PACER laps was entered into the model, no other variable contributed significantly to the estimation of VO_2 max. The model developed on the current sample was: VO_2 max = 20.41012 + (PACER laps * 0.41304).

For comparison purposes, cross-validation was conducted on several published prediction models (see Figure 2). Prediction error was assessed with two equations. The standard error of estimate (*SEE*) was calculated as: $SEE = S_Y\sqrt{1 - R^2_{YY'}}/N$. Total Error was calculated as: $TE = \sqrt{\sum (Y - Y')^2/N}$. For these equations, Y is measured VO_2 max and Y' is VO_2 max estimated from the prediction equation. Comparisons of these two error estimates help establish the overestimation and underestimation of each prediction model.

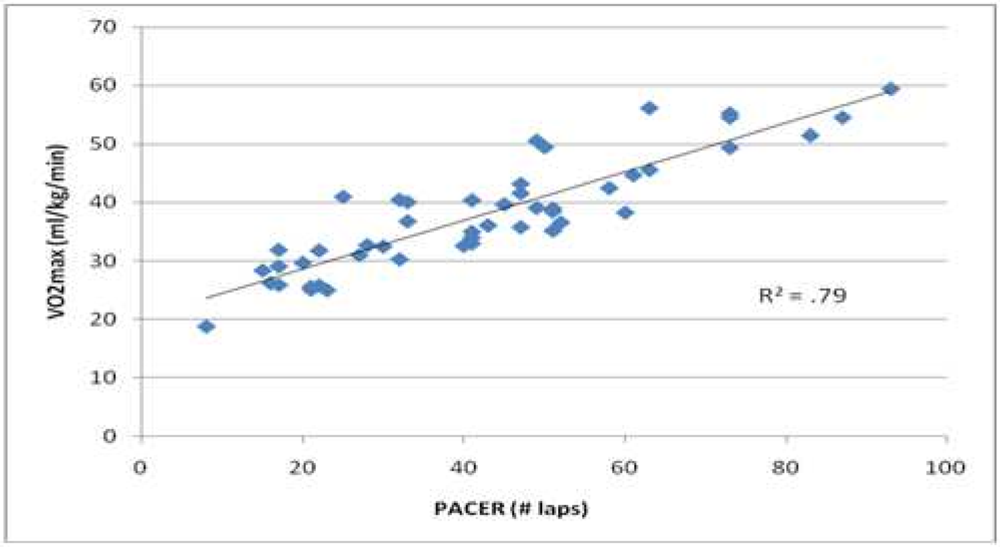

Figure 1. Scatterplot of measured VO_2 max vs. PACER performance in total sample (*N* = 48)

Reference	Prediction Equation
Leger & Lambert (1982)	VO_2 max = 5.857 * (speed in km·h^{-1}) – 19.458
Leger et al. (1988)	VO_2 max = 31.025 + (3.238 * PACER speed) – (3.248 * age) + (0.1536 * PACER speed * age)
Mahar et al. (2010)	VO_2 max = 41.76799 + (0.49261 * PACER laps) – (0.00290 * PACER laps^2) – (0.61613 * BMI) + (0.34787 * sex [M=1, F=0] * age)
Barnett et al. (1993) A	VO_2 max = 25.8 – (6.6 * sex [M=0, F=1]) – (0.2 * mass [kg]) + 3.2 * PACER speed)
Barnett et al. (1993) B	VO_2 max = 24.2 – (5.0 * sex [M=0, F=1]) – (0.449 * age) – (0.831 * BMI) + (4.12 * PACER speed)
Matzusaka et al. (2004) A	VO_2 max = 25.9 – (2.21 * sex [M=0, F=1]) – (0.449 * age) – (0.831 * BMI) + (4.12 * PACER speed)
Matzusaka et al. (2004) B	VO_2 max = 61.1 – (2.20 * sex [M=0, F=1]) – (0.462 * age) – (0.862 * BMI) + (0.192 * PACER laps)

Note: PACER speed is maximal speed attained on the PACER test and age is in years.

Figure 2. Previously published prediction equations

Table 3. Correlations, Standard Errors of Estimate (*SEE,)* and Total Errors (*TE*) between Measured VO_2 max and VO_2 max Predicted from PACER Equations

Prediction Equation	**Sample**	***R***	***SEE***	***TE***
Current study	17 – 18 yrs (*N* = 48)	.89	4.36	-
Léger & Lambert (1982)	18 – 36 yrs (*N* = 91)	.90	4.19	7.11
Léger et al. (1988)	8 – 19 yrs (*N* = 188)	.89	4.43	4.81
Mahar et al. (2010)	10 – 16 yrs (*N* = 244)	.82	5.47	8.88
Barnett et al. (1993) A	12 – 17 yrs (*N* = 55)	.85	5.02	7.50
Barnett et al. (1993) B	12 – 17 yrs (*N* = 55)	.81	5.62	8.70
Matsuzaka et al. (2004) A	8 – 23 yrs (*N* = 132)	.84	5.13	6.33
Matsuzaka et al. (2004) B	8 – 23 yrs (*N* = 132)	.81	5.58	5.96

Note: *R* = Correlation between measured and estimated VO_2 max, *SEE* = Standard Error of Estimate, *TE* = Total Error.

All prediction equations produced high correlations with VO_2 max (≥ .81). The current prediction model and the Léger and Lambert (1982; 1988) models produced the highest correlations with VO_2 max and the lowest standard errors of estimate. In general, all equations produced relatively similar correlations (*R* = .81 – 90) and *SEEs* (4.19 – 5.62 $ml \cdot kg^{-1} \cdot min^{-1}$). The *TE* provides a more complete estimate of error in cross-validation situations because *TE* considers mean differences, as well as correlations. Apart from the model developed on the current sample and the Léger et al. (1988) model, the other prediction models produced *TEs* that were ≥ 5.96 $ml \cdot kg^{-1} \cdot min^{-1}$ (See Figure 3).

Criterion-referenced Analysis

The final analyses were conducted in a criterion-referenced framework. Estimated VO_2 max for each participant from each of the previously developed models was classified according to the new FITNESSGRAM® standards (Welk et al., 2010). For the first analysis, estimated VO_2 max was classified into one of three zones. These zones included: Healthy Fitness Zone (HFZ), Needs Improvement – Some Risk (NISR), and Needs Improvement – High Risk (NIHR). For the second analysis, only two classifications were used: Healthy Fitness Zone and Needs Improvement (NI). Table 4 illustrates the proportion of agreement (*Pa*), modified kappa (*Kq*), and phi coefficient (*Phi*) statistics between measured and estimated VO_2 max.

Table 4. Classification Agreement with Measured VO_2 max ($ml \cdot kg^{-1} \cdot min^{-1}$) (*N* = 48)

	3 categories (HFZ, NISR, NIHR)			2 categories (HFZ, NI)		
Model	*Pa*	*Kq*	*Phi*	*Pa*	*Kq*	*Phi*
Current Study	.73	.60	.68	.88	.76	.73
Léger & Lambert (1982)	.54	.31	.76	.75	.50	.60
Léger et al. (1988)	.75	.63	.76	.85	.70	.69
Mahar et al. (2010)	.52	.28	.57	.66	.32	.50
Barnett et al. (1993) A	.52	.28	.57	.68	.36	.48
Barnett et al. (1993) B	.66	.32	.56	.66	.32	.50
Matsuzaka et al. (2004) A	.63	.45	.58	.70	.40	.51
Matsuzaka et al. (2004) B	.63	.45	.57	.77	.54	.53

Note: HFZ is Healthy Fitness Zone, NISR is Needs Improvement – Some Risk, NIHR is Needs Improvement – High Risk, NI is Needs Improvement and includes both the Some Risk and High Risk categories. *Pa* is Proportion of Agreement, *Kq* is Modified Kappa and *Phi* is Phi Coefficient.

For the three category analysis, proportion of agreement was low to moderate (.52 - .73) for all prediction equations. The Léger et al. (1988) model and the prediction model developed in the current study produced the highest *Pa* (.75 and .73). These models also had the highest *Kq* and *Phi* statistics. The remaining prediction models produced *Pa* statistics that were low. For the Léger et al. (1982), Mahar et al. (2010), Barnett et al. (1993) A, and Barnett et al. (1993) B models, the proportion of agreement ranged from .52 - .66.

When classification was condensed into two classification zones (HFZ and NI), *Pa* was increased on all prediction models. Furthermore, the prediction model developed in the current study produced the highest *Pa* amongst all the models (.88). The Léger et al. (1988) model also produced a *Pa* (.85) similar to that of the prediction model from the current study. Both the prediction model from the current study and Léger et al. (1988) model appear to provide a reasonably valid categorization of aerobic fitness.

Discussion

The PACER provides a convenient and relatively low cost assessment of aerobic fitness, with acceptable evidence of accuracy and reliability (Léger & Lambert, 1982; Léger et al., 1988; Mahar et al., 2010). The PACER is less expensive and more readily available than stationary or portable metabolic systems, can be completed in most open spaces or gymnasiums, and administered to a large group. The PACER does not require a track or walking path and can be completed in a large hallway or multi-purpose room. The PACER is the preferred (or default) test of aerobic fitness for the FITNESSGRAM® youth fitness program. The ability of the PACER to accurately predict VO_2 max has been examined in children and young adolescents (Barnett, Chan, & Bruce, 1993; Léger et al., 1988; Liu, Plowman, & Looney, 1992; Mahar et al., 2006). Currently, limited research has investigated the predictive power of the PACER test to accurately estimate VO_2 max in older adolescents (Mahar et al., 2006).

The purpose of the present study was to examine the validity of the PACER test. Validity was examined by comparing VO_2 max predicted from the PACER test to measured VO_2 max. In addition, previously published PACER prediction models were cross-validated. Furthermore, criterion-referenced aerobic fitness standards (Welk et al., in press) were used to determine if VO_2 max estimated from the prediction model developed in the current study and VO_2 max predicted from previously published prediction models correctly categorized individuals into the appropriate fitness zone.

Prediction of VO_2 max from the PACER Test

Estimates of VO_2 max from the PACER prediction models developed in the current study and in previously published studies were significantly correlated with measured VO_2 max. The prediction model developed in the current study and the Léger and Lambert (1982) and Léger et al. (1988) models had the highest correlations (.89 - .90) amongst all the prediction models.

Correlations for all the models examined were greater (.81 - .90) than correlations presented by previous researchers (Mahar et al., in press) [.58 - .73]. Total error accounts for both correlations and absolute differences between measured and predicted values. Examination of total error distinguishes the more and less accurate models. The prediction model developed in the current study and the Léger et al. (1988) model produced relatively low standard errors (4.36 and 4.81 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively). The *TEs* produced by the other models were relatively high (5.96 $ml \cdot kg^{-1} \cdot min^{-1}$ to 8.88 $ml \cdot kg^{-1} \cdot min^{-1}$) and resembled the *TEs* reported by Mahar et al. (6.37 $ml \cdot kg^{-1} \cdot min^{-1}$ to 7.61 $ml \cdot kg^{-1} \cdot min^{-1}$). Although, the Mahar et al. data were collected on a slightly younger sample (10 to 16 years of age), the analyses of the data mimicked that of the current study.

Comparison of *SEEs* from previous studies and the current study can be challenging because the standard deviation of the criterion variable (VO_2 max, in this case) may differ from study to study, and the standard deviation is used to calculate the *SEE*. However, within the same study the examination of standard errors of estimate provides a good indication of the models' accuracy. Examination of the *SEE*s and *TE*s can help identify the more accurate equations. Table 3 demonstrates the comparative accuracy of the current and previously published prediction models. The correlations between measured and predicted VO_2 max for the model developed in the current study were similar (r = .89) to the correlations for the Léger and Lambert (1982) [r = .90] and Léger et al. (1988) [r = .89] models. The prediction model developed in the current study was more accurate than the other prediction models examined (i.e., lower *SEE*, with the only exception being the Léger and Lambert (1982) model, which had a similar *SEE*). The high *TE* for the Leger et al. (1982) model indicates the prediction accuracy of that model is questionable for 17 and 18 year old participants.

Léger and Lambert (1982) developed a regression model to estimate VO_2 max from the PACER test in adults aged 18 – 30 years. They reported a significant correlation (r = .84) between VO_2 max estimated from their model and measured VO_2 max, with an *SEE* of 5.4 $ml \cdot kg^{-1} \cdot min^{-1}$. When examining the cross-validation evidence, correlations between measured and estimated VO_2 max for the model developed in the current study were similar to the correlation reported by Léger and Lambert. The model developed in the current study produced a similar *SEE* (4.36 $ml \cdot kg^{-1} \cdot min^{-1}$) to the Léger and Lambert model (4.19 $ml \cdot kg^{-1} \cdot min^{-1}$). The difference of 0.17 $ml \cdot kg^{-1} \cdot min^{-1}$ is minimal. One difference between the two models is the predictor variable of maximal speed attained in the Léger and Lambert model and the variable of laps completed in the model developed in the current study. Maximal speed attained for the Léger and Lambert model categorizes all participants who reach a particular level as attaining the same speed. However, a person who only completes the first lap of a level and a person who may complete all but one lap in a level (e.g., 7 more laps), may have different fitness levels. In the Léger and Lambert model two participants who complete a different number of laps on the same level will be categorized as having similar fitness levels. Although, both models produced high correlations with measured VO_2 max, use of laps completed rather than maximal speed attained may differentiate better among people with different levels of fitness.

Léger et al. (1988) developed a regression model to estimate VO_2 max from the PACER test in children, young adolescents, and older adolescents (8 – 19 years). The Léger et al. model accounted for the variables of maximal speed attained, sex, and the speed by age interaction. Léger et al. reported a significant correlation (r = .71) between measured and estimated VO_2 max. The correlation between measured and estimated VO_2 max was stronger for the prediction model developed in the current study (r = .89) compared to the Léger et al. model.

Léger et al. recruited a sample of 8 – 19 year old participants, whereas in the current study only 17 – 18 year old participants were recruited. Interestingly, when the Léger et al. model was cross-validated in the current study, it produced a similar correlation between measured and estimated VO_2 max and similar *SEEs* and *TEs* as the model developed in the current study. The model developed in the current study and the Léger et al. model produced the most accurate prediction of VO_2 max for the current study sample.

Mahar et al. (in press) developed a regression model to estimate VO_2 max from the PACER test in youth and young adolescents (aged 10 – 16 years). Mahar et al. reported a significant correlation (r = .75) between measured and estimated VO_2 max. Additionally, Mahar et al. cross-validated their prediction model with the same models that were cross-validated in the current study. When the previous prediction models were cross-validated with the Mahar et al. sample, moderate correlations (r = .58 - .69) were produced. Furthermore, when *SEE*s and *TE*s were compared between the Mahar et al. model and the previously published models, the Mahar et al. model produced the lowest *SEE* (6.39 $ml \cdot kg^{-1} \cdot min^{-1}$) and *TE* (6.37 $ml \cdot kg^{-1} \cdot min^{-1}$).

The PACER model developed in the current study produced similar results to other prediction models like the 1 mile walk/run, which is often used by practitioners to estimate VO_2 max amongst adolescents. The overall correlation between measured and estimated VO_2 max in the present study (r = .89) was higher than the correlation reported by Cureton et al. (1995) for the 1 mile run/walk test (r = .71). Cureton et al. reported a significant sex x age interaction associated with estimating VO_2 max from the 1 mile run/walk. Mahar et al. (in press) also reported a significant sex x age interaction. However, the sex x age interaction was not a significant contributor to the estimation of VO_2 max in the current model. The lack of a sex x age interaction in the current model can be attributed to the narrow age range and the fact that the number of laps completed on the PACER accounted for a large portion of variance in estimating VO_2 max in this sample. The 1 mile run/walk aerobic field test has been used as a primary field test for aerobic capacity in the past. However, the need for a large area or track to perform the 1 mile run/walk can hinder a practitioner's ability to administer the test. The PACER test requires limited space and has been shown to provide as accurate an estimate of VO_2 max as the 1 mile run/walk for youth and adolescents.

Classification of Fitness Levels

The ability of the prediction model developed in the current study to classify participants into an appropriate fitness zone (i.e., Healthy Fitness Zone or Needs Improvement Zone) was also explored. The FITNESSGRAM® report contains an individual's estimated VO_2 max, and a participant's fitness zone. These fitness zones help educate a participant on whether he/she is categorized in the Healthy Fitness Zone or in the Needs Improvement Zone. Participants are categorized into the Healthy Fitness Zone when their estimated VO_2 max falls above the criterion-referenced standard. Participants are categorized into the Needs Improvement Zone when their estimated VO_2 max falls below the criterion-referenced standard. The Needs Improvement Zone is further categorized as 'Needs Improvement – Some Risk' and 'Needs Improvement – High Risk' (Welk et al., in press). Criterion-referenced validity can be examined by comparing a participant's classification based on estimated VO_2 max from a prediction model to classification based on measured VO_2 max.

Criterion-referenced validity for the prediction model developed in the current study was moderate (Pa = .73) and higher than the other models with the only exception being the Léger et al. (1988) model (Pa = .75). The model developed in the current study and the Léger et al. (1988) model misclassified a lower percentage of participants (25 – 27% misclassified) into the HFZ than the other models examined (34% - 48% misclassified). The remaining prediction models inaccurately categorized participants nearly 50% of the time. If a practitioner was trying to reliably categorize participants into the Healthy Fitness Zone or the Needs Improvement Zone, he/she would need to consider the capability of the equation to accurately categorize participants. However, when the Needs Improvement Zones were condensed from two (Need Improvement - Some Risk and Needs Improvement - High Risk) into one category (Needs Improvement), the classification accuracy improved. Criterion-referenced validity for the newly developed model under the two category framework was high (*Pa* = .88) and outperformed all other predictions models. Furthermore, the Léger et al. (1988) model had similar criterion-referenced validity (*Pa* = .85) when compared to the current model. This is encouraging for practitioners as the need to accurately classify participants into a Healthy Fitness Zone or not is one of the main concerns for the FITNESSGRAM® program.

Limitations of the Current Study

Results of the current study should be interpreted in light of the study limitations. Based on the criterion-referenced standards established for aerobic fitness, the current prediction model and previously published models had moderate to low agreement when trying to correctly categorize individuals into one of the three fitness zones (Healthy Fitness Zone, Needs Improvement – Some Risk, Needs Improvement – High Risk). However, the overall capabilities of the current model to accurately categorize participants increased when the Needs Improvement category was reduced from two categories to one category. The current model was developed on a homogeneous age group (17 – 18 years). If a practitioner attempted to implement this model to estimate VO_2 max in a sample of participants outside the age range of the current model sample, the results may not accurately predict aerobic fitness.

Suggestions for Future Research

In the present study, laps completed on the PACER were the only variable that significantly contributed to the estimation of VO_2 max in this sample. Previously published models (Barnett et al., 1993; Léger & Lambert, 1982; Léger et al., 1988; Matsuzaka et al., 2004) used maximal speed attained on the PACER as a predictor variable. Further research is needed regarding this issue of whether to use laps completed on the PACER instead of maximal speed attained.

Conclusion

In conclusion, the prediction model generated in the current study developed on an older adolescent sample provides an accurate estimate of VO_2 max. The variable of laps completed on the PACER was the only significant contributor to the equation. Age, sex, BMI, and percent body fat did not significantly contribute to the estimation of VO_2 max for the current model. The current prediction model also produced accurate classification of fitness levels into the Healthy Fitness Zone or Needs Improvement Zone.

References

Aarnio, M., Winter, T., Peltonen, J., Kujala, U., & Kaprio, J. (2002). Stability of leisure-time physical activity during adolescence a longitudinal study among 16- 17-and 18 year-old Finnish youth. *Scandinavian Journal of Medicine and Science in Sports*, 12, 179-185.

Aaron et al. (1993). The epidemiology of leisure physical activity in an adolescent population. *Medicine and Science in Sports and Exercise*, 25, 847-853.

Armstrong, N. & Welsman, J. R. (1994). Assessment and interpretation of aerobic fitness in children and adolescents. *Exercise and Sports Science Reviews*, 22, 435-476.

Barnett, A., Chan, L. S., & Bruce, I. C. (1993). A preliminary study of the 20-m multistage shuttle run as a predictor of peak VO_2 in Hong Kong Chinese students. *Pediatric Exercise Science*, 5, 42-50.

Blair, S. N., Kohl, H. W., Gordon, N. F., Paffenbarger, R. S. (1992). How much physical activity is good for health? *Annual Review of Public Health*, 13, 99-126.

Bouchard et al. (1998). Familial resemblance for VO_2 max in the sedentary state: the HERITAGE family study. *Medicine and Science in Sports and Exercise*, 30, 252-258.

Bruce, R. A., Blackman, J. R., & Jones, J. W. (1963). Exercise testing in adult normal subjects and cardiac patients. *Pediatrics*, 32, 742-755.

Butler, S. M., Black, D. R., Blue, C. L., & Gretchbeck, R. J. (2004). Change in diet, physical activity, and body weight in female college freshman. *American Journal of Health*, 28, 24-33.

California Senate Report No. 265 (2008), the State Board of Education approved amendments to the Education Code.

Carnethon, M. R., Gidding, S. S., Nehgme, R., Sidney, S., Jacobs, D. R., & Liu, K. (2003). Cardiorespiratory fitness in young adulthood and the development of cardiovascular disease risk factors. *The Journal of the American Medical Association*, 290, 3092-3100.

Carnethon, M. R., Gulati, M., & Greenland, P. (2005). Prevalence and cardiovascular disease correlates of low cardiorespiratory fitness in adolescent and adults. *The Journal of the American Medical Association*, 294, 2981-2988.

Caspersen, C. J., Pereira, M. A., & Curran, K. M. (2000). Changes in physical activity patterns in the United States, by sex and cross-sectional age. *Medicine and Science in Sports and Exercise*, 32, 1601-1609.

Center for Disease Control and Prevention National Health and Nutritional Examination Survey. http://www.cdc.gov/nchs/about/major/nhanes/datalink.htm. Accessed: February 25, 2010.

Center for Disease Control and Prevention National Youth Risk Behavior Survey. http://www.cdc.gov/healthyyouth/yrbs/pdf/yrbs07_us_overview.pdf. Accessed: March 1, 2010.

Cooper, S. M., Baker, J. S., Tong, R. J., Roberts, E., & Hanford, M. (2005). The repeatability and criterion related validity of the 20-m multistage fitness test as a predictor of maximal oxygen uptake in active young men. *British Journal of Sports Medicine*, 39, e19-26.

Cureton, K. J., Sloniger, M. A., O'Bannon, J. P., Black, D. N., & McCormack, W. P. (1995). A generalized equation for prediction of VO_2peak from 1-mile run/walk performance. *Medicine and Science in Sports and Exercise*, 27, 445-451.

Delaware Senate Report No. 503 (2009), the Delaware Board of Education approved amendments to the Education Code.

Eisenmann, J. C., Katzmarzyk, P. T., Perusse, L., Tremblay, A., Després, J. P., & Bouchard, C. (2005). Aerobic fitness, body mass index, and CVD risk factors among adolescents: the Québec family study. *International Journal of Obesity*, 29, 1077-1083.

Eisenmann, J. C., Wickel, E. E., Welk, G. J., & Blair, S. N. (2005). Relationship between adolescent fitness and fatness and cardiovascular disease risk factors in adulthood: The Aerobic Center Longitudinal Study (ACLS). *American Heart Journal*, 149, 46-53.

Fernhall, B., Pitetti, K. H., Millar, L., Hensen, T., & Vukovich, M. (2000). Cross validation of the 20-meter shuttle run test for children and adolescents with mental retardation. *Adaptive Physical Activity Quarterly*, 17, 402-412.

Fotheringham, E. L., Neville, M. J., & Bauman, A. (2001). Age-related differences in physical activity levels of young adults. *Medicine and Science in Sports and Exercise*, 33, 255-258.

Gordan-Larsen, P., McMurray, R. G., Popkins, B. M. (2000). Determinants of adolescent physical activity and inactivity patterns. *Pediatric Exercise Science*, 105, e83.

Heyward, V. H., & Wagner, D. R., (2004). *Applied body composition assessment* (pp. 87-98) (2nd Ed.). Champaign, IL: Human Kinetics.

Houtkooper, L. B., Lohman, T. G., Going, S. B., & Howell, W. (1996). Why bioelectrical impedance analysis should be used for estimating adiposity. *American Journal of Clinical Nutrition*,64, 436-448.

Jackson, A. S., & Pollock, M. L. (1985). Practical assessment of body composition. *Physician and Sports Medicine*, 13, 76-90.

Janz, K. F., Dawson, J. D., & Mahoney, L. T. (1999). Tracking physical fitness and physical activity from childhood to adolescence: the Muscatine study. *Medicine and Science in Sports and Exercise*, 32(7), 1250-1257.

Katzmarzyk, P. T., Malina, R. M., Song, T. M., & Bouchard, C. (1998). Physical activity and health-related fitness in youth: amultivariate analysis. *Medicine and Science in Sports and Exercise*, 30(5), 709-714.

Kimm et al. (2002). Decline in physical activity in black girls and white girls during adolescence. *New England Journal of Medicine*, 347, 709-715.

Lamb, K. L., & Rogers, L. (2007). A re-appraisal of the reliability of the 20-m multi-stage shuttle run test. *European Journal of Applied Physiology*, 100, 287-292.

Léger, L. A., & Lambert, J. (1982). A maximal multistage 20-m shuttle run test to predict VO_2 max. *European Journal of Applied Physiology*, 49, 1-12.

Léger, L. A., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 meter shuttle run test for aerobic fitness. *Journal of Sports Sciences*, 6, 93-101.

Liu, N. Y., Plowman, S. A., & Looney, M. A. (1992). The reliability and validity of the 20-meter shuttle test in American students 12 to 15 years old. *Research Quarterly for Exercise and Sport*, 63, 360-365.

Lobelo, F., Pate, R. R., Dowda, M., Liese, A. D., & Ruiz, J. R. (2009). Validity of cardiorespiratory fitness criterion-referenced standards for adolescents. *Medicine and Science in Sports and Exercise*, 41, 1222-1229.

Mahar, M. T., Guerieri, A. M., Hanna, M. S., & Kemble, C. D. (in-press). Estimation of aerobic fitness from PACER 20-m shuttle run test performance. *American Journal of Preventive Medicine*.

Mahar, M. T., Welk, G. J., Rowe, D. A., Crotts, D. J., & McIver, K. L. (2006). Development and validation of a regression model to estimate VO_2peak from PACER 20-m shuttle run performance. *Journal of Physical Activity and Health*, 3, S34-46.

Mahoney, C. (1992). 20-MST and PWC170 validity in non-Caucasian children in the UK. *British Journal of Sports Medicine*, 26, 45-47.

Matsuzaka, A., Takahashi, Y., Yamazoe, M., Kumakura, N., Ikeda, A., Wilk, B., & Bar-Or, O. (2004). Validity of the multistage 20-m shuttle run test for Japanese children, adolescents, and adults. *Pediatric Exercise Science*, 16, 113-125.

McVeigh, S. K., Payne, A. C., & Scott, S. (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. *Pediatric Exercise Science*, 7, 69-79.

Meredith, M. D. & Welk, G. J. (Eds.). (2010). *FitnessGram and ActivityGram test and administration manual* (Updated 4th Ed.). Champaign, IL: Human Kinetics.

New York Senate Report (2009), The New York City Board of Education approved amendments to the Education Code.

NHLBI Obesity Education Initiative Expert Panel on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. (1998). Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: The Evidence Report. *Obesity Research*, 6, S51-S209.

Nokes, N. (2009). Relationship between physical activity and aerobic fitness. *Journal of Sports Medicine and Physical Fitness*, 49, 136-141.

Ortiz et al. (1992). Differences in skeletal muscle and bone mineral mass between black and white females and their relevance to estimates of body composition. *American Journal of Clinical Nutrition*, 55, 8-13.

Paliczka, V. J., Nichols, A. K., & Boreham, C. A., (1987). A multi-stage shuttle run as a predictor of running performance and maximal oxygen uptake in adults. *British Journal of Sports Medicine*, 21, 163-165.

Pate, R. R., Heath, G. W., Dowda, M., & Trost, S. G. (1996). Association between physical activity and other health behaviors in a representative sample of US adolescents. *American Journal of Public Health*, 86, 1577-1581.

Pate, R. R., Pratt, M., Blair, S. N., (1995). Physical activity and public health: a recommendation from the Centers for Disease Control and the American College of Sports Medicine. *The Journal of the American Medical Association*, 273, 402-407

Pate, R. R., Wang, C. Y., Dowda, M., Farrell, S. W., & O'Neil, J. R. (2006). Cardiorespiratory fitness levels among US youth 12 to 19 years of age. *Archives of Pediatric and Adolescent Medicine*, 160, 1005-1012.

Pitetti, K. H., Fernhall, B., & Figoni, S. (2002). Comparing two regression formulas that predict VO_2peak using the 20-m shuttle for children and adolescents. *Pediatric Exercise Science*, 14, 125-134.

Ransbottom, R., Brewer, J., & Williams, C. (1988). A progressive shuttle run test to estimate maximal oxygen uptake. *British Journal of Sports Medicine*, 22, 141-144.

Rowland, T. W. (1993). Does peak VO_2 reflect VO2 max in children? Evidence from supramaximal testing. *Medicine and Science in Sports and Exercise*, 25, 689-693.

Siri, W. E., (1961). Body composition from fluid spaces and density: Analysis of methods. In J. Brozlek and A. Henschel (Eds.), *Techniques for measuring body composition* (pp. 223-224). Washington, DC: National Academy of Sciences.

Slaughter et al. (1988). Skinfold equation for estimation of body fatness in children and youth. *Human Biology*, 60, 709-723.

Sproule, J., Kunalan, C., McNeil, M., & Wright, H. (1993). Validity of 20-MST for predicting VO_2 max of adult Singaporean athletes. *British Journal of Sports Medicine*, 27, 202-204.

Talbot, L. A., Metter, E. J., & Fleg, J. L. (2000). Leisure-time physical activities and their relationship to cardiorespiratory fitness in health men and women 18-95 years old. *Medicine and Science in Sports and Exercise*, 32, 417-425.

Telama, R., & Yang, X. (2000). Decline in physical activity from youth to young adulthood in Finland. *Medicine and Science in Sports and Exercise*, 32, 1617-1622.

Texas Senate Report No. 530 (2007), the State Board of Education approved amendments to the Education Code.

Troiano, R. P., Berrigan, D., Dodd, K. W., Másse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometers. *Medicine and Science in Sports and Exercise*, 40, 181-188.

Trost et al. (2002). Age and gender differences in objectively measured physical activity in youth. *Medicine and Science in Sports and Exercise*, 34, 350-355.

Twisk, J. W., Kemper, H. C., & van Mechelen, W. (2000). Tracking of activity and fitness and the relationship with cardiovascular disease risk factors. *Medicine and Science in Sports and Exercise*, 32, 1455-1461.

U.S. Department of Health and Human Services. (1996) *Physical activity and health: A report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Center for Disease Control and Prevention, National Centers for Chronic Disease Prevention and Health Promotion, 1996, pp. 1-278.

U.S. Department of Health and Human Services. (2008). *2008 physical activity guidelines for the American [electronic resource]: be active, healthy, and happy!* Washington, DC: U.S. Department of Health and Human Services.

Welk, G. J., & Meredith, M. D. (Eds.). (2007). *Fitnessgram/Activitygram test administration manual* (4^{th} ed.). Champaign, IL: Human Kinetics.

Welk, G. J., Morrow, J. R., & Falls, H. B. (Eds.). (2002), *Reference Guide*. Dallas, TX: The Cooper Institute.