Journal of Exercise Science & Fitness 18 (2020) 129–135

Contents lists available at ScienceDirect

Journal of Exercise Science & Fitness

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

Testing validity of FitnessGram in two samples of US adolescents (12–15 years)

Eun-Young Lee [a,**], Joel D. Barnes [b], Justin J. Lang [b,c], Diego A.S. Silva [d], Grant R. Tomkinson [e,f], Mark S. Tremblay [b]

[a] School of Kinesiology and Health Studies, Queen's University, Kingston, ON K7L2N9, Canada
[b] Healthy Active Living and Obesity (HALO) Research Group, CHEO Research Institute, Ottawa, ON K1H 8L1, Canada
[c] Centre for Surveillance and Applied Research, Public Health Agency of Canada, Ottawa, ON K1A 0K9, Canada
[d] Research Center in Kinanthropometry and Human Performance, Federal University of Santa Catarina, Florianopolis, 88040-900, Brazil
[e] Department of Education, Health and Behavior Studies, University of North Dakota, ND, 58202, USA
[f] Alliance for Research in Exercise, Nutrition and Activity, University of South Australia, SA, 5001, Australia

A R T I C L E I N F O

Article history:
Received 4 February 2020
Received in revised form
28 April 2020
Accepted 29 April 2020
Available online 10 May 2020

Keywords:
Aerobic fitness
Aerobic-capacity standards
Receiver operating characteristic curve analysis

A B S T R A C T

Background: This study examined the validity of the FitnessGram® criterion-reference cut-points for cardiorespiratory fitness (CRF) based on two samples of US adolescents (aged 12–15 years). This study also established the CRF cut-points for metabolically healthy weight status based on a recent national fitness survey for the purposes of cross-validating with pre-existing cut-points including FitnessGram.
Methods: Two cross-sectional data from the 2003–2004 National Health and Nutrition Examination Survey (NHANES) (n = 378) and 2012 NHANES National Youth Fitness Survey (NNYFS) (n = 451) were used. CRF (estimated $\dot{V}O_{2max}$ in mL/kg/min) was estimated from a submaximal exercise test. CRF categories based on FitnessGram cut-points, a clustered cardiometabolic risk factors score and weight status were used. A series of Receiver Operating Characteristic (ROC) curve analyses were conducted to identify age- and sex-specific CRF cut-points that were optimal for metabolically healthy weight status.
Results: Based on FitnessGram cut-points, having high risk CRF, but not low risk CRF, was associated with high cardiometabolic risk (OR = 3.17, 95% CI = 1.14–8.79) and unhealthy weight status (OR = 5.81, 95% CI = 3.49–9.68). The optimal CRF cut-points for 12-13-year-olds and 14-15-year-olds were 40 and 43 mL/kg/min in males and 39 and 34 mL/kg/min in females, respectively. Compared to meeting new CRF cut-points, not meeting new CRF cut-points was associated with higher odds of showing high cardiometabolic risk (OR = 2.91, 95% CI = 1.47–5.77) and metabolically unhealthy weight status (OR = 4.47, 95% CI = 2.83–7.05).
Conclusion: FitnessGram CRF cut-point itself has rarely been scrutinized in previous literature. Our findings provide partial support for FitnessGram based on two samples of US adolescents. CRF cut-points established in this study supports international criterion-referenced cut-points as well as FitnessGram cut-points only for males. FitnessGram should be continuously monitored and scrutinized using different samples.

Introduction

Cardiovascular disease is the leading cause of morbidity and mortality globally; approximately 17% of all deaths in high-income countries in 2016 was due to ischemic heart disease.[1] Cardiorespiratory fitness (CRF) is a strong predictor of various health outcomes[2] including cardiovascular and respiratory function in adolescents, independent of physical activity and obesity.[3] Given recent temporal declines in CRF[4] globally and US adolescents showing low levels of CRF relative to international norms,[5,6]

** Corresponding author. School of Kinesiology and Health Studies, Queen's University, KHS 307, 28 Division St, Kingston, ON K7L 3N6, Canada.
E-mail addresses: eunyoung.lee@queensu.ca, eunyoung.lee@queensu.ca (E.-Y. Lee), jbarnes@cheo.on.ca (J.D. Barnes), justin.lang@canada.ca (J.J. Lang), diegoaugustoss@yahoo.com.br (D.A.S. Silva), grant.tomkinson@und.edu (G.R. Tomkinson), mtremblay@cheo.on.ca (M.S. Tremblay).

https://doi.org/10.1016/j.jesf.2020.04.002
1728-869X/

monitoring CRF using standardized CRF cut-points among US adolescents is important for their current and future health.[7]

The FitnessGram® program (The Cooper Institute, Dallas, TX) is a fitness assessment that has been widely used in US schools to help teachers track and screen health-related fitness over time.[8,9] Owned by the Cooper Institute, FitnessGram-related products are being sold in a range of forms to US schools, such as flexibility measuring boxes ($165.99), pull-up bars ($599.99), and two forms of software ($599 hosted version and $799 enterprise version) with license renewals ranging between $99–149.[10]

CRF cut-points in FitnessGram were last updated in 2011.[9] Specifically, age- and sex-specific CRF cut-points (also called "standards") were developed based on the data from the National Health and Nutrition Examination Surveys (NHANES) 1999–2000 and 2001–2002 cycles.[11] Being mandated in several provinces in the US (https://www.cooperinstitute.org/pub/news.cfm?id=24), FitnessGram has been widely adopted by US schools for fitness testing with support from scholarly articles, the Cooper Institute and not-for-profit organizations such as Human Kinetics, SHAPE America, US games, and National Football League (via providing subsidies by having their logo on FitnessGram products).[10]

In recent scholarship, it was noted that out of 300 peer-reviewed articles pertaining to FitnessGram, only a few scrutinized the shortcomings and challenges associated with its implementation or the methodology, while the limitations that FitnessGram itself has have been largely ignored.[10,12] FitnessGram CRF cut-points include three categories: Healthy Fitness Zone (HFZ) and two levels of Needs Improvement Zone (NIZ). The HFZ reflects the CRF values associated with adequate functional and/or health-related indicators while NIZ consists of two risk levels: some risk and high risk.[9] Given the temporal decline of CRF among adolescents worldwide, particularly among US adolescents,[4] it is important that we continue to regularly monitor and scrutinize FitnessGram cut-points to strive for better science and further our health surveillance efforts at a population level.[13]

The primary aim of this study was to examine the validity of the current FitnessGram standards based on two newer samples of US adolescents based on the national datasets. The secondary aim of this study was to generate new criterion-referenced age- and sex-specific CRF cut-points using the most recent dataset available in the US, the NHANES National Youth Fitness Survey (NNYFS) 2012. The secondary aim was to determine whether CRF cut-points established based on our sample align with FitnessGram[11] and two other interim international criterion-referenced cut-points.[6,14]

Methods

Data source

Cross-sectional data from the 2003-2004 NHANES and the 2012 NHANES NNYFS were used in the present study. The NHANES is a continually administered national survey conducted by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC). The NHANES is designed to assess the health and nutrition status of US citizens via in-person interviews and direct physical examinations in biannual cycles. Directly measured CRF data among adolescents since the most recent update of the FitnessGram which was based on the 1999–2002 NHANES datasets were only available up to the 2003–2004 cycle; thus, the 2003–2004 NHANES dataset was selected for analysis. Using stratified multistage clustered sampling, a total of 1131 adolescents aged 12–15 years were included. We only included this age group because fitness data were only available among adolescents aged 12–15 years in the other key dataset used in this study, NNYFS. The 2012 NNYFS was conducted by the CDC NCHS to obtain recent data on fitness and health indicators among a nationally representative sample of children and adolescents (aged 3–15 years). Measurement included a household interview and an assessment of anthropometric and fitness measures in a Mobile Examination Center. A total of 510 adolescents aged 12–15 years with information on body mass index (BMI) and CRF were included.

All publicly available files from NHANES and NNYFS were de-identified before the CDC released the data; therefore, Institutional Review Board approval was not obtained. Detailed data collection procedures have been described elsewhere (https://www.cdc.gov/nchs/nhanes/index.htm).[15]

Measures

Cardiorespiratory fitness

Predicted maximal oxygen uptake ($\dot{V}O_{2max}$) using a submaximal treadmill exercise test was used to estimate CRF. Trained health technicians conducted a submaximal treadmill exercise test with each participant assigned to one of eight test protocols varying in difficulty, based on their age, sex, BMI, and physical activity levels. Each protocol aimed to have participants reach ~75% of their age-predicted maximum heart rate (220 − age) by the end of the test. Those at high risk of adverse events during exercise testing (e.g., individuals with known disease or signs and symptoms of disease) were excluded based on pre-exercise screening using questionnaires and resting heart rate and blood pressure criteria. The testing protocol consisted of a 2-min warm-up and two 3-min stages, followed by a 2-min cool-down. Heart rate response was monitored at the end of each stage as well as after each minute of recovery using the Colin STBP-780 automated BP (San Antonio, TX, USA), and wires connected to four electrodes places on the thorax and abdomen (NHANES Cardiovascular Fitness Procedures Manual, 2005). $\dot{V}O_{2max}$ was estimated by extrapolation of heart rate from the two 3-min exercise stages to age-predicted maximum heart rate. Detailed description of the NHANES CRF testing procedures are available elsewhere (NHANES Cardiovascular Fitness Procedures Manual, 2005). Based on the age- and sex-specific $\dot{V}O_{2max}$ cut-points in FitnessGram (https://pftdata.org/files/hfz-standards.pdf), CRF scores were classified into three groups: "HFZ", "NI-some risk" and "NI-high risk".[11]

Cardiometabolic risk factors

Waist circumference, systolic and diastolic blood pressure, fasting triglycerides, high-density lipoprotein cholesterol (HDL-C), and fasting glucose were used to indicate clustered cardiometabolic risk among participants using the 2003–2004 NHANES data based on the criteria used in previous studies to determine metabolic syndrome in adults[17] and adolescents.[18,19] Waist circumference was measured to nearest 0.1 cm using a steel measuring tape just above the uppermost lateral border of the iliac crest. Systolic and diastolic blood pressure were measured at least three consecutive times manually using a mercury sphygmomanometer after a 5-min rest while sitting. The three consecutive measurements were then averaged to obtain the final blood pressure values. Triglycerides, HDL-C, and fasting glucose were determined from blood samples taken from participants. Triglyceride samples were only drawn from those who fasted at least 9 h before venipuncture. All data collection for cardiometabolic risk factors were done in the mobile examination center laboratory and the blood samples were processed, stored, and sent to the Johns Hopkins University Lipoprotein Analytical Laboratory (for HDL-C), Johns Hopkins Hospital (for triglycerides), and to the University of Missouri-Columbia (for glucose) for analysis. Age- and sex-specific cut-off values from the National Cholesterol Education Program/Adult

Treatment Panel III Adult values were used to determine cardiometabolic risk for each variable.[18] In addition, all cardiometabolic risk scores were standardized as z-scores, with HDL-C scores multiplied by -1 before standardization to be consistent with other risk factors (i.e., lower scores reflect favorable conditions while higher scores reflect unfavorable conditions). Clustered cardiometabolic risk was calculated using the following equation: ([standardized waist circumference + standardized mean arterial pressure + standardized triglycerides + standardized HDL cholesterol • -1 + standardized fasting glucose]/5). The clustered cardiometabolic risk values were then dichotomized into "high risk" and "low risk". High cardiometabolic risk was defined as >75th percentile of the total clustered cardiometabolic risk score (z-score of >0.67)[19].

In the NNYFS, standing height was measured by a trained health technician using a portable stadiometer SECA 217 while body mass was measured using a portable digital weight scale SECA 869 (Chino, CA, USA). Body mass index (BMI) was calculated for each participant using the following equation: body mass (kg)/height (m).[2] Weight status was classified into metabolically healthy or unhealthy weight status based on the age-and sex-specific 5th and 85th percentiles using the 2000 CDC growth charts.[15]

Covariates

Age, sex, race/ethnicity, BMI, and annual household income were included as covariates for the primary aim of this study given their previously demonstrated relationships with cardiometabolic risk factors.[20] Age in years was reported as age at last birthday. Sex was self-reported as male or female. Race/ethnicity was derived by combining responses to questions on race and Hispanic origin. The categories included Mexican American, Other Hispanic, non-Hispanic Black, non-Hispanic White, and other race including non-Hispanic multiracial. Based on the recommendation from the NCHS, race/ethnicity was categorized as Hispanic, non-Hispanic Black, non-Hispanic White, and other race (including mixed race).[15] Annual household income was collected by asking individuals to report their income from a scale ranging from "1 ($0–4999)" to "11 ($75,000 and over)", "12 (Over $20,000)", "13 (Under $20,000)", "77 (Refused)", or "99 (Don't know)". The income variable was recoded into four categories: $0–19,999, $20,000–44,999, $45,000–64,999, ≥$65,000, and Refused/Don't know.

Statistical analyses

Descriptive statistics for the study variables were calculated and presented as means and standard deviations for continuous variables and as percentages for categorical variables. To achieve the primary aim, logistic regression analyses were performed using both the 2003–2004 NHANES and 2012 NNYFS datasets. Odds ratios (OR) and 95% confidence intervals (95% CI) was calculated to examine the associations between three risk levels of the FitnessGram CRF cut-points and a clustered cardiometabolic risk score in 2003–2004 NHANES, and metabolically healthy/unhealthy weight status in 2012 NNYFS. Exposure-outcome gradients were calculated using the analysis of linear trend (p-for-trend) to examine if associations between FitnessGram CRF cut-points and a clustered cardiometabolic risk score/weight status are in a linear fashion.

To explore whether the CRF cut-points established based on our sample support the interim international criterion-referenced and FitnessGram cut-points, Receiver Operating Characteristic (ROC) curves were used to identify age- and sex-specific CRF cut-points to discriminate metabolically healthy and unhealthy weight status, while maximizing both sensitivity and specificity using Youden's J statistic. Sensitivity is the proportion of individuals who are correctly classified as having healthy weight status according to the 2000 CDC growth chart, whereas specificity denotes the proportion of individuals who are correctly classified as having unhealthy weight status. FitnessGram cut-points categorize adolescents into three different groups (HFZ, NIZ-some risk, and NIZ-high risk) based on CRF scores. While having three categories can provide more specific information, it is also more susceptible to classification bias than having two categories. Therefore, CRF cut-points were only developed based on two categories (i.e., metabolically healthy vs. unhealthy weight status) in our study. Area under the curve (AUC) was used to determine the discriminating power of weight status with the AUC value of "1" indicating perfect discriminating power and "0.5" indicating no discriminating power. For sensitivity analysis purposes, logistic analyses were conducted to examine the association between CRF categories based on the newly generated age- and sex-specific CRF cut-points and a clustered cardiometabolic risk score in 2003–2004 NHANES and weight status in 2012 NNYFS. All analyses were conducted in R (version 3.5.1 for Windows) and statistical significance was set a priori at $p < 0.05$. The ROC curves were generated using the pROC package. All analyses incorporated sample weights, cluster, and strata variables provided with each respective dataset.

Results

Of the 1131 eligible adolescents who participated in 2003–2004 NHANES, only 362 adolescents had complete cases. This major loss was mainly due to missing cases in cardiometabolic risk factors (e.g., Triglycerides, missing n = 650) that are commonly observed with NHANES laboratory data. Regardless of large missing cases, no significant differences existed between the samples included and excluded in all key variables and covariates ($p > 0.05$). For the 2012 NNYFS data, 59 adolescents were excluded from the analysis due to missing data, leaving a total of 451 in the analytic sample. Between samples included and excluded, a significant difference was found in race/ethnicity ($p = 0.002$). Included sample had more Hispanic (26.8 vs 21.8%) and non-Hispanic Black (14.7 vs 17.4%) and less non-Hispanic White adolescents (56.0 vs 49.0%) than the excluded sample.

The sample characteristics are presented in Table 1. In the 2003–2004 NHANES dataset, the average estimated $\dot{V}O_{2max}$ for male and female adolescents were 44 mL/kg/min and 39 mL/kg/min, respectively. Percentage estimates for adolescents with unfavorable CRF based on the FitnessGram cut-points were 24% in males and 62% in females. The proportions of male and female adolescents showing clustered cardiometabolic risk (>75th percentile of the total clustered cardiometabolic risk score; z-score of >0.67) were 26% and 21%, respectively. In the 2012 NNYFS dataset, the average $\dot{V}O_{2max}$ for male and female adolescents were 43 mL/kg/min and 38 mL/kg/min, respectively. Percentage estimates for adolescents with unfavorable CRF based on the FitnessGram cut-points were 50% in males and 68% in females. The proportions of male and female adolescents showing metabolically unhealthy weight status were 49% and 42%, respectively.

Table 2 presents the associations between FitnessGram-derived age- and sex-specific CRF categories and clustered cardiometabolic risk factors after adjusting for covariates. Specifically, compared to those in HFZ, those in NI-high risk (OR = 3.17, 95%CI = 1.14–8.79) were more likely to have high cardiometabolic risk in 2003–2004 NHANES data. When the 2012 NNYFS data were used, NI-high risk was associated with metabolically unhealthy weight status (OR = 5.81, 95%CI = 3.49–9.68). NI-some risk was not associated with the outcomes in both datasets.

To generate age- and sex-specific CRF cut-points, ROC curve analyses were run using the 2012 NNYFS dataset. As shown in

Table 1
Sample-weighted characteristics.

	Before excluding missing cases (N = 1131)		After excluding missing cases		
			All	Males	Females
2003–2004 NHANES					
N			362	194	168
Age (years) (Mean, 95% CI)	1105	13.5 (13.4–13.5)	13.5 (13.4–13.6)	13.6 (13.5–13.8)	13.4 (13.1–13.6)
Height (cm) (Mean, 95% CI)	1096	162.6 (161.7–163.5)	163.5 (161.8–165.2)	166.9 (164.8–169.0)	159.6 (158.2–160.9)
Mass (kg) (Mean, 95% CI)	1096	59.8 (58.0–61.5)	59.2 (56.0–62.3)	62.1 (57.9–66.4)	55.8 (53.2–58.4)
Body mass index (kg/m^2) (Mean, 95% CI)	1096	22.4 (21.7–23.1)	21.9 (21.1–22.7)	22.0 (21.0–23.0)	21.8 (21.0–22.6)
Race/Ethnicity (%, 95%CI)	1105				
Hispanic		16.0 (10.2–24.3)	18.9 (11.5–29.4)	19.6 (11.2–32.0)	18.4 (11.2–28.7)
Non-Hispanic Black		15.5 (10.9–21.6)	18.5 (12.5–26.5)	18.0 (11.8–26.5)	19.0 (12.4–27.9)
Non-Hispanic White		63.0 (52.0–72.8)	56.2 (43.0–68.5)	55.8 (43.8–67.2)	56.4 (40.6–71.0)
Other		5.5 (3.3–8.9)	6.4 (3.4–12.0)	6.6 (2.3–17.6)	6.3 (2.6–14.1)
Annual household income (%, 95%CI)	1047				
<$20,000		16.6 (13.0–21.0)	17.3 (11.9–24.4)	14.9 (8.5–24.7)	20.0 (12.8–30.0)
$20,000–44,999		28.8 (25.8–32.1)	33.6 (27.1–40.8)	34.0 (23.4–46.6)	33.1 (24.3–43.2)
$45,000–64,999		18.7 (13.4–25.3)	15.9 (9.9–23.7)	15.1 (8.0–26.7)	16.2 (8.2–29.3)
≥$65,000		35.9 (30.2–42.1)	33.4 (25.0–43.1)	35.7 (24.9–48.2)	30.7 (21.2–42.2)
Don't know/refused		0.3 (0.1–1.4)	0.1 (0.0–1.0)	0.1 (0.0–1.0)	–
Estimated $\dot{V}O_{2max}$ (mL/kg/min) (Mean, 95% CI)	853	41.6 (40.1–43.1)	41.8 (40.2–43.4)	44.4 (42.5–46.3)	38.7 (36.8–40.6)
Cardiometabolic risk factors (Mean, 95% CI)					
Waist circumference (cm)	1080	78.5 (76.8–80.1)	77.0 (75.1–78.8)	77.8 (75.0–80.6)	76.1 (74.5–77.7)
NCEP-determined high risk (%, 95% CI)		19.9 (14.8–26.4)	16.8 (12.1–22.7)	8.5 (4.6–15.0)	26.2 (17.4–37.5)
Blood pressure (mmHg)	1075				
Systolic		107.1 (106.2–108.0)	107.4 (105.6–109.2)	111.9 (108.5–113.3)	103.3 (101.4–105.2)
NCEP-determined high risk (%, 95% CI)		6.3 (4.1–9.6)	5.0 (2.3–10.6)	9.2 (4.4–18.1)	0.3 (0.0–2.5)
Diastolic		57.2 (56.1–58.3)	58.4 (56.5–60.2)	56.3 (53.6–58.9)	60.9 (59.0–62.7)
NCEP-determined high risk (%, 95% CI)		0.9 (0.4–2.2)	1.5 (0.5–4.2)	2.8 (1.1–7.2)	0.0 (0.0–0.0)
Triglycerides (mmol/L)	481	1.0 (0.9–1.1)	1.0 (0.8–1.1)	1.0 (0.8–1.1)	0.9 (0.8–1.1)
NCEP-determined high risk (%, 95% CI)		12.4 (7.9–19.1)	12.9 (8.2–19.8)	16.5 (9.0–28.3)	8.8 (2.9–23.9)
HDL-C (mmol/L)	996	1.4 (1.3–1.4)	1.4 (1.4–1.5)	1.4 (1.3–1.5)	1.5 (1.4–1.5)
NCEP-determined high risk (%, 95% CI)		23.9 (18.9–29.7)	22.2 (16.7–28.9)	18.4 (9.1–33.6)	26.5 (20.4–33.6)
Fasting glucose (mmol/L)	487	5.1 (5.0–5.3)	5.0 (4.9–5.1)	5.1 (5.1–5.2)	4.9 (4.8–5.0)
NCEP-determined high risk (%, 95% CI)		0.5 (0.1–3.4)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
FITNESSGRAM standards (%, 95% CI)					
Healthy Fitness Zone		56.5 (48.8–63.8)	46.1 (37.8–54.6)	75.9 (63.0–85.3)	37.9 (29.2–47.6)
Needs Improvement Zone-some risk		12.1 (9.5–15.4)	17.2 (12.0–24.1)	10.8 (6.6–17.2)	16.3 (9.9–25.7)
Needs Improvement Zone-high risk		31.4 (24.2–39.7)	36.6 (27.4–46.9)	13.3 (6.5–25.2)	45.8 (32.3–59.9)
High clustered cardiometabolic risk score (>75th percentile) (%, 95% CI)		28.1 (19.6–38.5)	28.7 (21.8–36.7)	26.3 (16.3–39.6)	20.9 (14.6–29.1)
2012 NNYFS					
n		510	451	228	223
Age (years) (Mean, 95% CI)	510	13.4 (13.3–13.5)	13.3 (13.1–13.4)	13.5 (13.2–13.7)	13.0 (12.8–13.2)
Height (cm) (Mean, 95% CI)	510	162.4 (161.6–163.3)	162.6 (161.8–163.5)	165.5 (163.9–167.1)	159.3 (157.7–160.9)
Mass (kg) (Mean, 95% CI)	489	60.8 (59.1–62.4)	61.2 (58.9–63.5)	62.9 (59.7–66.1)	59.3 (56.3–62.2)
Body mass index (kg/m^2) (Mean, 95% CI)	489	22.8 (22.2–22.4)	22.9 (22.2–23.6)	22.7 (22.0–23.4)	23.2 (22.1–24.2)
Race/Ethnicity (%, 95% CI)	510	[a]			
Hispanic		21.8 (13.2–33.8)	26.8 (15.7–41.8)	27.9 (17.2–41.9)	25.4 (12.2–45.7)
Non-Hispanic Black		14.7 (7.0–28.3)	17.1 (8.2–32.2)	18.6 (8.4–36.3)	15.4 (6.9–30.7)
Non-Hispanic White		56.0 (38.8–71.9)	49.0 (30.9–67.3)	45.1 (29.4–61.9)	53.5 (30.0–75.6)
Other		7.5 (4.3, 12.7)	7.1 (4.3–11.6)	8.4 (4.9–14.0)	5.7 (2.2–13.8)
Annual household income (%, 95% CI)	493				
<$20,000		18.6 (14.6–23.2)	18.8 (14.8–23.5)	19.4 (12.9–28.0)	18.0 (12.8–24.9)
$20,000–44,999		40.5 (33.8–47.6)	37.6 (30.4–45.3)	34.2 (26.0–43.4)	41.6 (33.3–50.4)
$45,000–64,999		17.3 (13.9–21.3)	19.5 (15.2–24.7)	24.8 (18.0–33.0)	13.4 (10.4–17.2)
≥$65,000		21.3 (16.0–27.8)	21.3 (15.5–28.5)	19.1 (12.8–27.6)	23.8 (16.0–33.8)
Don't know/refused		2.3 (1.2–4.5)	2.8 (1.5–5.4)	2.6 (1.1–5.9)	3.1 (1.1–8.2)
Estimated $\dot{V}O_{2max}$ (mL/kg/min) (Mean, 95% CI)	451	40.9 (39.5–42.2)	40.5 (38.9–42.0)	43.1 (41.3–44.9)	37.5 (36.3–38.7)
FITNESSGRAM standards (%, 95% CI)	451				
Healthy Fitness Zone		43.1 (37.2–49.2)	41.6 (35.4–48.1)	49.9 (41.7–58.2)	31.9 (26.3–38.0)
Needs Improvement Zone-some risk		14.8 (10.1–21.1)	11.6 (8.2–16.3)	9.3 (5.3–15.9)	14.3 (9.0–22.2)
Needs Improvement Zone-high risk		42.1 (33.6–51.2)	46.8 (38.7–55.0)	40.8 (32.2–49.9)	53.8 (44.6–62.7)
Weight status (%, 95% CI)	489				
Metabolically unhealthy		44.4 (38.2–50.8)	45.5 (40.4–50.7)	48.7 (42.4–55.1)	41.7 (33.5–50.4)

Data source: 2003–3004 NHANES, 2012 NNYFS.
Note. HDL-C, High-Density Lipoprotein Cholesterol; National Cholesterol Education Program/Adult Treatment Panel III Adult (NCEP).
NHANES, National Health and Nutrition Examination Survey; NNYFS, NHANES National Youth Fitness Survey.
Values are described as mean or percentage (%) and 95% confidence intervals (95% CI).
[a] Significant difference between samples included vs excluded by $p < 0.05$.

Table 3, the overall CRF cut-points identified for adolescents (aged 12–15 years) were 41 mL/kg/min for males and 38 mL/kg/min for females. In addition, age- and sex-specific optimal CRF cut-points for 12- to 13-year-olds were 40 mL/kg/min for males and 39 mL/kg/min for females. The corresponding cut-points for 14- to 15-year-olds were 43 mL/kg/min for males and 34 mL/kg/min for females.

Table 4 shows the associations between newly generated CRF cut-points based on the 2012 NNYFS data and cardiometabolic score/weight status using 2003–2004 NHANES and 2012 NNYFS. Compared to meeting new CRF cut-points, not meeting new CRF cut-points was associated with higher odds of showing high cardiometabolic risk (OR = 2.91, 95% CI = 1.47–5.77) and metabolically unhealthy weight status (OR = 4.47, 95% CI = 2.83–7.05).

Discussion

This study provided only partial support for FitnessGram CRF cut-points based on 2003–2004 NHANES and 2012 NNFYS data. Though the health outcomes and samples used were different across the two datasets, our results suggest that it may be important to reconsider and potentially update the current FitnessGram CRF cut-points using a more robust methodology. This study established overall and age- and sex-specific CRF cut-points using 2012 NNYFS to cross-validate with FitnessGram and two interim international criterion-references cut-points.[6,14] The CRF cut-points established in this study for 12-15-year-olds are 41 mL/kg/min for males and 38 mL/kg/min for females, which align well with interim international cut-points but only partially for FitnessGram (only in males).

FitnessGram classifies adolescents into three groups based on their CRF performance: those who are “healthy”, and those with “some risk” and “high risk”.[21] FitnessGram has its merits for providing age- and sex-specific CRF cut-points and enabling large-scale school-based physical fitness testing to enhance physical education programming[22]; however, accumulating evidence also suggests that FitnessGram’s implementation[12] and interpretation of findings[10] are still in question and need to be considered. Perhaps this is due to having three unique categories to determined health risk based on CRF scores that largely vary by individuals based on many factors (e.g., biological, behavioural, environmental).[23,24] Also, FitnessGram CRF cut-point itself may need more thorough evaluation given that we fail to detect significantly different health outcomes among those at low risk compared to those who are healthy based on their CRF scores. Combined, FitnessGram cut-points should be continuously monitored and scrutinized using robust science that provides strong reproducibility and replicability.

Table 2
Associations between categories of cardiorespiratory fitness[a] and cardiometabolic risk factors among US adolescents aged 12–15 years (logistic regression results).

	Adjusted OR (95% CI)[b]
2003–2004 NHANES (n = 362)	
High clustered cardiometabolic score	
Healthy Fitness Zone	1.00 (reference)
Needs Improvement-some risk	2.63 (0.97–7.16)
Needs Improvement-high risk	3.17 (1.14–8.79)[d]
p-for-trend	$p < 0.001$
	Adjusted OR (95% CI)[c]
2012 NNYFS (n = 451)	
Metabolically unhealthy weight status	
Healthy Fitness Zone	1.00 (reference)
Needs Improvement-some risk	1.53 (0.68–3.46)
Needs Improvement-high risk	5.81 (3.49–9.68)[d]
p-for-trend	$p < 0.001$

Data source: 2003–3004 NHANES, 2012 NNYFS.
Note. CI, Confidence interval; NHANES, National Health and Nutrition Examination Survey; NNYFS, NHANES National Youth Fitness Survey; OR, Odd ratio.
[a] Defined based on FITNESSGRAM cutoff values.[11]
[b] Covariates included age, sex, race/ethnicity, household income, and body mass index.
[c] Covariates included age, sex, race/ethnicity, and household income.
[d] $p < 0.05$.

Overall, the new CRF cut-points generated in this study support the previously determined CRF cut-points suggested in the literature.[6,14] The utility of CRF as a predictor of cardiometabolic health, including metabolically healthy weight status, among adolescents has been well documented in previous reviews.[6,14] A recent systematic review on international interim criterion-referenced standards for healthy CRF is 42 mL/kg/min for males and 35 mL/kg/min for females.[6] The cut-points generated using the ROC curve technique in our study provides similar ranges of CRF cut-points (40–43 mL/kg/min for males and 34–39 mL/kg/min for females) in comparison to the interim international criterion-referenced cut-points.[6,14] Compared to our new CRF cut-points and the suggested interim international cut-points, FitnessGram (i.e., 42 mL/kg/min for males and 40 mL/kg/min for females) appear to be consistent for males but slightly over-estimated for females to accurately determine cardiometabolic risk.[11]

Other than our findings suggesting that the validity of FitnessGram cut-points may have been inflated in previous literature, several logistical challenges associated with FitnessGram were reported in previous literature, including children’s disliking of fitness testing,[25] parents’ resistance,[26,27] teachers’ unpreparedness to conduct fitness testing and manage data and lack of resources.[10,12,28] Here, it is important to highlight that this study did not intend to interfere with the implementation and advocacy for FitnessGram but to encourage the advancement of the science and implementation of future fitness testing through scholarly scrutiny.

This study is one of the first to scrutinize FitnessGram CRF cut-points using empirical data. Our findings evidently suggest that FitnessGram may should be continuously evaluated for its validity. There are several limitations to be acknowledged in this study. Though the sample for the NHANES is selected to represent the population, the findings are not likely generalizable to whole US adolescents due to the substantial sample loss in the 2003–2004 NHANES data and the robust sample exclusion applied for the treadmill test.[16] That said, the proportion of individuals who met the FitnessGram standards (those in the HFZ) (50–56% for males and 34–36% for females) observed in our study was consistent with previous findings.[22,29] It is indicated that to detect obesity, higher $\dot{V}O_{2max}$ values are required than to detect cardiometabolic risk factors.[6,30] Therefore, the inconsistency in the prognostic ability of FitnessGram standards between two datasets may be due to different outcome variables (clustered cardiometabolic score vs metabolically healthy/unhealthy weight status) used. Unfortunately, cardiometabolic risk measures were not available in the 2012 NNYFS data, thus, an alternative measure of cardiometabolic health, BMI-determined weight status, was used as the outcome. Additionally, ROC curve analyses were run using a relatively small sample size to obtain age- and sex-specific CRF cut-points. Given these limitations, the findings should be reproduced and replicated with larger samples and using more robust measures of CRF and health outcomes in different populations.

Monitoring population-level CRF has been suggested as an effective way to monitor and evaluate current and future health status among adolescents, and the first step towards CRF surveillance is to collate evidence from different populations using various health outcomes to establish standardized CRF cut-points that have improved utility globally.[13] The results of the present study demonstrate that FitnessGram partially discriminated those with and without cardiometabolic risk among US adolescents in two

Table 3
Receiver operation characteristic (ROC) curve for the optimal CRF cut-points for metabolically healthy weight status among US adolescents aged 12–15 years.

	CRF cut-point (mL/kg/min) (95% CI)	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Overall (n = 451)	39.6 (39.2–42.5)	0.66 (0.61–0.72)	0.72 (0.62–0.81)	0.60 (0.48–0.68)
Males (n = 228)	40.7 (39.5–47.2)	0.73 (0.66–0.80)	0.68 (0.55–0.88)	0.73 (0.49–0.83)
12–13 years (n = 116)	39.5 (37.4–43.8)	0.68 (0.59–0.78)	0.61 (0.46–0.81)	0.78 (0.58–0.90)
14–15 years (n = 112)	43.2 (40.5–48.9)	0.76 (0.67–0.85)	0.86 (0.67–1.00)	0.64 (0.43–0.83)
Females (n = 223)	38.3 (32.4–41.6)	0.64 (0.57–0.72)	0.79 (0.35–0.90)	0.51 (0.38–0.91)
12–13 years (n = 117)	38.7 (32.4–42.5)	0.67 (0.57–0.77)	0.83 (0.40–0.94)	0.54 (0.38–0.90)
14–15 years (n = 106)	34.1 (32.1–39.4)	0.63 (0.50–0.75)	0.55 (0.30–0.91)	0.79 (0.40–0.95)

Data source: 2012 NNYFS.
Note. AUC, Area under the curve; CI, Confidence interval; CRF, Cardiorespiratory fitness; NNYFS, National Health and Nutrition Examination Survey National Youth Fitness Survey; OR, Odd ratio.

Table 4
Associations between new CRF cut-points[a] and high clustered cardiometabolic score (2003–2004 NHANES)/weight status (2012 NNYFS) among US adolescents aged 12–15 years.

	%	Adjusted OR (95% CI)
2003–2004 NHANES (n = 362)		
High clustered cardiometabolic score		
Meeting new CRF cut-point	79.4	1.00 (reference)
Not meeting new CRF cut-point	20.6	2.91 (1.47–5.77)*
2012 NNYFS (n = 451)		
Metabolically unhealthy weight status		
Meeting new CRF cut-point	78.5	1.00 (reference)
Not meeting new CRF cut-point	21.5	4.47 (2.83–7.05)*

Data source: 2003–3004 NHANES, 2012 NNYFS.
Covariates included age, sex, race/ethnicity, household income, and body mass index for 2003–2004 NHANES and age, sex, race/ethnicity, and household income for 2012 NNYFS.
Note. CI, Confidence interval; CRF, Cardiorespiratory fitness; NHANES, National Health and Nutrition Examination Survey; NNYFS, NHANES National Youth Fitness Survey; OR, Odd ratio.
*$p < 0.05$.
[a] New CRF cut-points are derived from Table 3.

datasets. While the results suggest that it may be worthwhile to continue to monitor and scrutinize the validity of FitnessGram using more recent, larger nationally representative data, caution is needed when interpreting our findings due to the aforementioned limitations associated with the present study. Future research is needed to continue to replicate and reproduce CRF cut-points and test their utility in different populations using more robust methodology.

Funding/support statement

No financial or material support of any kind was received for the work described in this article.

Declaration of competing interest

The author(s) have no conflicts of interest relevant to this article.

CRediT authorship contribution statement

Eun-Young Lee: Conceptualization, Methodology, Validation, Investigation, Writing - original draft, Writing - review & editing, Formal analysis. **Joel D. Barnes:** Methodology, Formal analysis, Writing - review & editing, Visualization. **Justin J. Lang:** Methodology, Writing - review & editing. **Diego A.S. Silva:** Methodology, Writing - review & editing. **Grant R. Tomkinson:** Methodology, Writing - review & editing. **Mark S. Tremblay:** Conceptualization, Methodology, Writing - review & editing, Supervision.

Acknowledgements

The data that support the findings of this study are openly available in the US Centers for Disease Control and Prevention Webpage at https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2003 and at https://www.cdc.gov/nchs/nnyfs/index.htm.

References

1. World Health Organization. *Disease Burden and Mortality Estimates: Cause-specific Mortality*; 2000-2016. Published 2018 http://www.who.int/healthinfo/global_burden_disease/estimates/en/.
2. Lang JJ, Belanger K, Poitras V, Janssen I, Tomkinson GR, Tremblay MS. Systematic review of the relationship between 20 m shuttle run performance and health indicators among children and youth. *J Sci Med Sport*. 2018;21(4):383–397. https://doi.org/10.1016/j.jsams.2017.08.002.
3. Myers J, McAuley P, Lavie CJ, Despres JP, Arena R, Kokkinos P. Physical activity and cardiorespiratory fitness as major markers of cardiovascular risk: their independent and interwoven importance to health status. *Prog Cardiovasc Dis*. 2015;57(4):306–314. https://doi.org/10.1016/j.pcad.2014.09.011.
4. Tomkinson GR, Lang JJ, Tremblay MS. Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. *Br J Sports Med*. 2019;53(8):478–486. https://doi.org/10.1136/bjsports-2017-097982.
5. Lang JJ, Phillips EW, Orpana HM, et al. Field-based measurement of cardiorespiratory fitness to evaluate physical activity interventions. *Bull World Health Organ*. 2018;96(11):794–796. https://doi.org/10.2471/BLT.18.213728.
6. Lang JJ, Tremblay MS, Ortega FB, Ruiz JR, Tomkinson GR. Review of criterion-referenced standards for cardiorespiratory fitness: what percentage of 1,142,026 international children and youth are apparently healthy? *Br J Sports Med*. 2019;53(15):953–958.
7. Eisenmann JC, Laurson KR, Welk GJ. Aerobic fitness percentiles for U.S. adolescents. Suppl. 2 *Am J Prev Med*. 2011;41(4):S106–S110. https://doi.org/10.1016/j.amepre.2011.07.005.
8. Plowman S a, Sterling CL, Corbin CB, Meredith MD, Welk GJ, Morrow JR. The history of FITNESSGRAM. *J Phys Activ Health*. 2006;3:S5–S20. https://doi.org/10.1123/jpah.3.s2.s5.
9. Welk GJ, De Saint-Maurice Maduro PF, Laurson KR, Brown DD. Field evaluation of the new FITNESSGRAM®criterion-referenced standards. *Am J Prev Med*. 2011;41(4 Suppl. 2):S131–S142. https://doi.org/10.1016/j.amepre.2011.07.011.
10. Gard M, Pluim C. Why is there so little critical physical education scholarship in the United States? The case of Fitnessgram. *Sport Educ Soc*. 2017;22(5):602–617. https://doi.org/10.1080/13573322.2016.1213716.
11. Welk GJ, Laurson KR, Eisenmann JC, Cureton KJ. Development of youth aerobic-capacity standards using receiver operating characteristic curves. *Am J Prev Med*. 2011;41(4 Suppl. 2):S111–S116. https://doi.org/10.1016/j.amepre.2011.07.007.
12. Ferguson RH, Keating XD, Bridges DM, Guan J, Chen L. California secondary school physical education teachers' attitudes toward the mandated use of the fitnessgram. *J Teach Phys Educ*. 2007;26(2):161–176. https://doi.org/10.1123/jtpe.26.2.161.
13. Lang JJ, Tomkinson GR, Janssen I, et al. Making a case for cardiorespiratory fitness surveillance among children and youth. *Exerc Sport Sci Rev*. 2018;46(2):66–75. https://doi.org/10.1249/jes.0000000000000138.
14. Ruiz JR, Cavero-Redondo I, Ortega FB, Welk GJ, Andersen LB, Martinez-Vizcaino V. Cardiorespiratory fitness cut points to avoid cardiovascular disease risk in children and adolescents; What level of fitness should raise a red flag? A systematic review and meta-analysis. *Br J Sports Med*. 2016;50(23):1451–1458. https://doi.org/10.1136/bjsports-2015-095903.
15. Borrud L, Chiappa MM, Burt VL, et al. National Health and Nutrition Examination Survey: national youth fitness survey plan, operations, and analysis, 2014 *Vital Health Stat*. 2012;2(163):1–24. Accessed http://www.ncbi.nlm.nih.

gov/pubmed/24709592. Accessed May 22, 2019.
16. Centers for Disease Control and Prevention. *National Health and Nutrition Examination Survey (NHANES) Anthropometry Procedures Manual*; 2007. Accessed https://www.cdc.gov/nchs/data/nhanes/nhanes_05_06/CV.pdf. Accessed August 2, 2019.
17. Alberti KGMM, Zimmet P, Shaw J. The metabolic syndrome - a new worldwide definition. *Lancet*. 2005;366(9491):1059–1062. https://doi.org/10.1016/S0140-6736(05)67402-8.
18. Jolliffe CJ, Janssen I. Development of age-specific adolescent metabolic syndrome criteria that are linked to the adult treatment Panel III and international diabetes federation criteria. *J Am Coll Cardiol*. 2007;49(8):891–898. https://doi.org/10.1016/j.jacc.2006.08.065.
19. Ruiz JR, Ortega FB, Rizzo NS, et al. High cardiovascular fitness is associated with low metabolic risk score in children: the European Youth Heart Study. *Pediatr Res*. 2007;61(3):350–355. https://doi.org/10.1203/pdr.0b013e318030d1bd.
20. Anderssen SA, Cooper AR, Riddoch C, et al. Low cardiorespiratory fitness is a strong predictor for clustering of cardiovascular disease risk factors in children independent of country, age and sex. *Eur J Cardiovasc Prev Rehabil*. 2007;14(4): 526–531. https://doi.org/10.1097/HJR.0b013e328011efc1.
21. Cureton KJ, Plowman SA, Mahar MT. *FITNESSGRAM/ACTIVITYGRAM Reference Guide*. fourth ed. 2013. https://doi.org/10.1055/s-0033-1334967. Dallas, TX.
22. Bai Y, Saint-Maurice PF, Welk GJ, Allums-Featherston K, Candelaria N, Anderson K. Prevalence of youth fitness in the United States: baseline results from the NFL PLAY 60 FITNESSGRAM partnership project. *J Pediatr*. 2015;167(3):662–668. https://doi.org/10.1016/j.jpeds.2015.05.035.
23. Ekelund U, Anderssen SA, Froberg K, Sardinha LB, Andersen LB, Brage S. Independent associations of physical activity and cardiorespiratory fitness with metabolic risk factors in children: the European youth heart study. *Diabetologia*. 2007;50(9):1832–1840. https://doi.org/10.1007/s00125-007-0762-5.
24. Mintjens S, Menting MD, Daams JG, van Poppel MNM, Roseboom TJ, Gemke RJBJ. Cardiorespiratory fitness in childhood and adolescence affects future cardiovascular risk factors: a systematic review of longitudinal studies. *Sports Med*. 2018;48(11):2577–2605. https://doi.org/10.1007/s40279-018-0974-5.
25. Naughton GA, Carlson JS, Greene DA. A challenge to fitness testing in primary schools. *J Sci Med Sport*. 2006;9(1-2):40–45. https://doi.org/10.1016/j.jsams.2006.01.002.
26. Madsen KA, Linchey J. School-based BMI and body composition screening and parent notification in California: methods and messages. *J Sch Health*. 2012;82(6):294–300. https://doi.org/10.1111/j.1746-1561.2012.00700.x.
27. Madsen KA. School-based body mass index screening and parent notification: a statewide natural experiment. *Arch Pediatr Adolesc Med*. 2011;165(11): 987–992. https://doi.org/10.1001/archpediatrics.2011.127.
28. Hartmeyer R, Mygind E. A retrospective study of social relations in a Danish primary school class taught in "udeskole. *J Adventure Educ Outdoor Learn*. 2016;16(1):78–89. https://doi.org/10.1080/14729679.2015.1086659.
29. Lobelo F, Pate RR, Dowda M, Liese AD, Ruiz JR. Validity of cardiorespiratory fitness criterion-referenced standards for adolescents. *Med Sci Sports Exerc*. 2009;41(6):1222–1229. https://doi.org/10.1249/MSS.0b013e318195d491.
30. Ruiz JR, Huybrechts I, Cuenca-García M, et al. Cardiorespiratory fitness and ideal cardiovascular health in European adolescents. *Heart*. 2015;101(10): 766–773. https://doi.org/10.1136/heartjnl-2014-306750.