Evidence that Aerobic Fitness Is More Salient than Weight Status in Predicting Standardized Math and Reading Outcomes in Fourth- through Eighth-Grade Students

Robert R. Rauner, MD, MPH[1], Ryan W. Walters, MS[2], Marybell Avery, PhD[3], and Teresa J. Wanser, MA[3]

Objective To determine whether aerobic fitness is more salient than weight status in predicting performance on standardized math and reading tests in fourth- to eighth-grade students.

Study design A cross-sectional study of data abstracted from 11 743 students in 47 public schools. Aerobic fitness was defined by entering the healthy fitness zone of Fitnessgram's Progressive Aerobic Cardiovascular Endurance Run, which has been shown to correlate highly with maximum oxygen consumption. Mixed-effects logistic regression analyses were conducted to model the student-level effect of aerobic fitness status on passing the Nebraska State Accountability (NeSA) math and reading tests after adjusting for body mass index (BMI) percentile, free/reduced lunch status, sex, race, grade level, and school type.

Results After adjustment, aerobically fit students had greater odds of passing the NeSA math and reading tests compared with aerobically unfit students regardless of whether the students received free/reduced lunch; however, the effect of being aerobically fit on the standardized test scores was significantly greater for students not receiving free/reduced lunch. Weight status, as measured by BMI percentile, was not a significant predictor of passing the NeSA math or reading test after including free/reduced lunch status in the model.

Conclusions Aerobic fitness was a significant predictor of academic performance; weight status was not. Although decreasing BMI for an overweight or obese child undoubtedly improves overall health, results indicated all students benefit academically from being aerobically fit regardless of weight or free/reduced lunch status. Therefore, to improve academic performance, school systems should focus on the aerobic fitness of every student. *(J Pediatr 2013;163:344-8)*.

The long-term health consequences of obesity are well documented. Insufficient aerobic fitness has also been shown to be an independent risk factor for both long-term health problems and mortality.[1-3] Thus, monitoring aerobic fitness in addition to adiposity in children and adolescents is critically important. Despite the positive effects of improved aerobic fitness, school systems require empirical evidence showing an association between aerobic fitness and academic outcomes before prioritizing time and resources to this endeavor.

Several studies have shown significant associations between the Fitnessgram components and the standardized academic tests used for school accountability and funding.[4-7] Chomitz et al[4] indicated increased odds of passing math and reading tests with increases in the number of Fitnessgram components where the student entered the healthy fitness zone (HFZ); mixed associations were indicated between body mass index (BMI) and academics. Additional studies have shown aerobic fitness to predict academic outcomes similar to the multicomponent Fitnessgram.

The purpose of this study was 2-fold. First, we determined the unique student-level association between aerobic fitness and passing standardized math and reading tests. Second, we determined the student-level effect of BMI on passing standardized math and reading tests.

Methods

Math and reading test scores and Progressive Aerobic Cardiovascular Endurance Run (PACER), BMI, and demographic data, excluding unique individual identifiers, were abstracted from the school district's internal systems for all students enrolled in 37 elementary and 10 middle schools within a large Nebraska city. Abstracted data were collected between September 2010 and March 2011.

The primary outcome variables were the scaled scores from the Nebraska State Accountability (NeSA) tests for math and reading administered annually to all

BMI	Body mass index
HFZ	Healthy fitness zone
NeSA	Nebraska State Accountability
PACER	Progressive Aerobic Cardiovascular Endurance Run

From the [1]Partnership for a Healthy Lincoln, Lincoln, NE; [2]Division of Clinical Research and Evaluative Sciences, Department of Medicine, Creighton University, Omaha, NE; and [3]Lincoln Public Schools, Lincoln, NE

Funded by the Community Health Endowment of Lincoln and the US Department of Education Carol M. White Physical Education Program grant. The authors declare no conflicts of interest.

 http://dx.doi.org/10.1016/j.jpeds.2013.01.006

third- through eighth-grade students as well as students in grade 11. The reading test is administered on computer and the math test is administered via paper-and-pencil. The NeSA tests were developed as a result of state legislation requiring statewide testing in math, reading, writing, and science.[8] NeSA scores are used for federal accountability reporting as required by the No Child Left Behind Act of 2001.[9] Raw test scores are converted to scaled scores ranging from 0 to 200. For state assessment purposes, scaled scores are categorized as below the standards (0-84), meets the standards (85-134), or exceeds the standards (135-200).[10] The tests have not been vertically scaled across grades; thus, identical scaled scores do not imply the same level of math or reading ability across grade levels, precluding the use of continuous scaled scores as outcomes. Therefore, students' scaled scores were modeled as dichotomous using the defined state-level cutoff between below the standards and meets the standards (ie, not pass: <85; pass: ≥85).

We focused on Fitnessgram's PACER as the measure of aerobic fitness, with the primary independent variable being aerobic fitness status (ie, fit vs unfit).[11] The PACER was chosen as the stand-alone measure of aerobic fitness based on results from Varness et al[12] that showed a strong correlation between the PACER and maximum oxygen consumption ($r_s = 0.83$) measured by maximal oxygen consumption treadmill testing, the gold standard for measuring aerobic fitness. The PACER consists of a multistage progressive 15- or 20-m shuttle run requiring students to run laps between 2 markers in time with prerecorded audible beeps. The time between beeps decreases each minute, requiring a progressive increase in pace, and students run laps until they are unable to finish before the beep on 2 separate occasions. The PACER determines aerobic fitness status based on the number of laps completed, with the minimum number of laps required to enter the HFZ varying based on age and sex.[13] Historically, the HFZ for 9- and 10-year-old girls has been either 7 or 15 laps. The schools' physical education staff had been critical of the 7-lap threshold, judging it to be excessively low and not an accurate reflection of aerobic fitness. Our data showed that 98.3% and 88.6% of fourth- and fifth-grade girls, respectively, entered the HFZ using the 7-lap threshold, whereas 71.5% and 77.1% entered the HFZ using the 15-lap threshold. Given that the latter results were consistent with HFZ rates from other age and sex combinations, we chose to use the 15-lap threshold. The number of PACER laps completed for each student was entered into the Fitnessgram software (http://www.fitnessgram.net/home/) by their physical education teacher. The majority of schools used the 20-m version; schools with insufficient gym space used the 15-m version. A conversion procedure in the Fitnessgram software ensures that entering the HFZ requires identical effort across versions.[14]

Remaining covariates were selected based on previous literature and theory. BMI, BMI z score, and weight status based on BMI percentile (ie, underweight to obese) have shown conflicting associations with academic outcomes. Trained health technicians and treatment nurses collected and entered height and weight data into Sapphire Suite (K12 Systems, Inc, Allentown, Pennsylvania), a computer program used to compute BMI percentile based on Centers for Disease Control and Prevention BMI-for-age growth charts. For analysis, BMI percentile was treated as a continuous variable to prevent loss of information due to classification.

Free/reduced lunch status was used as a proxy for student socioeconomic status. Families apply for the federally funded free/reduced lunch program with qualification determined by the US Department of Agriculture and the Nebraska Department of Education's Nutrition Services. Approval is based on several factors, including household income, number of persons living in the household, whether the family receives food stamps, and enrollment in the Temporary Assistance to Needy Families or Food Distribution Program on Indian Reservations programs.

Student ethnicity is collected via student or parent self-report and is allowed to change annually. Approximately 70.1% of the students identified as white, with the second highest proportion being Hispanic (11.6%). Proportions of remaining ethnicities ranged from 6.7% (black or African American) to 0.1% (Native Hawaiian). Before analysis, the decision was made to dichotomize ethnicity into white versus nonwhite.

Finally, students' sex, grade level, and school type (ie, elementary or middle) were included as covariates. For the school system in the current study, elementary schools included kindergarten through fifth-grade students and middle schools included sixth- through eighth-grade students.

Statistical Analyses

Descriptive and demographic data are presented as mean and SD for continuous variables and as frequency and percent for categorical variables. An independent-samples *t* test was used to test for differences in BMI percentile, whereas the χ^2 test was used to test for differences in all remaining categorical variables.

For the primary analyses, mixed-effects logistic regression models were estimated to evaluate the effect of aerobic fitness status on passing the NeSA math and reading tests after adjusting for BMI percentile, free/reduced lunch status, sex, ethnicity, grade level, and school type. The mixed-effects analysis was used to account for dependence resulting from students being nested within schools.[15] Predictors were added sequentially to the model, including aerobic fitness status, to evaluate the unique effect of aerobic fitness after inclusion of the covariates, in which the fixed effects can be interpreted as conditional on schools, with all random effects fixed at zero (ie, unit-specific models).[16] At the student level, continuous predictors were centered near their grand mean; binary predictors remained uncentered. School-level means for all predictors, except school type, were included to represent contextual effects (ie, the incremental effects of school characteristics after controlling for student characteristics),[17] all of which were retained regardless of statistical significance for proper interpretation of student-level effects. Finally, the necessity for random slopes was evaluated separately for each student-level predictor; a random grade slope was observed for NeSA math and a random ethnicity slope was observed for NeSA reading.

All models were estimated via SAS PROC GLIMMIX version 9.2 (SAS Institute Inc, Cary, North Carolina) using numeric integration with 15 quadrature points, the logit link, and between–within denominator degrees of freedom. Improvement to model fit resulting from inclusion of random effects was evaluated via likelihood-ratio tests for nested models and Akaike information criterion for non-nested models. Statistical significance of fixed effects was evaluated using $P < .05$. Optimal parameter estimates for all fixed and random effects were indicated by gradient values (ie, partial first derivatives of the log-likelihood function with respect to each estimated parameter) approaching zero (ie, $\leq$0.0001). The McKelvey and Zavoina R^2 was used to estimate an R^2 similar to linear regression.[18,19]

Results

Data were abstracted for 12 678 students. Missing data were observed for ~7.5% of students resulting from parent refusal to provide data, disability, injury, or absence from school on testing days. One hundred students were missing both the NeSA math and reading test scores, with an additional 14 students missing only math scores and 1 student missing only the reading score. An additional 835 students were missing data for ≥1 model predictors. Final analyses were conducted on a sample of 11 729 students for math and 11 742 students for reading.

Descriptive and demographic data for all students included in analysis, stratified by aerobic fitness status, are presented in Table I. Approximately 69% of students entered the PACER's HFZ. This percentage was higher than that of Roberts et al,[6] who reported that 35.3% entered the 1-mile/run walk HFZ. Our percentage was more consistent with that of Welk et al,[7] who indicated that 76.9% of fourth-graders entered the HFZ for the PACER or 1-mile run/walk, a number that decreased steadily to include only 22.6% of 12th graders.

Bivariate correlations between all student-level variables included in the analysis are presented in Table II. Although the largest correlation was observed between aerobic fitness status and BMI percentile, this correlation was modest; all other correlations were small.

Intraclass correlations for all outcome and predictor variables were estimated using an empty model for each variable (ie, random intercept only, no predictors). Approximately 18% and 16% of the variance in the dichotomized NeSA math and reading test scores, respectively, were due to between-school differences. The predictor intraclass correlations were 12% for aerobic fitness status, 37% for free/reduced lunch status, 3% for BMI percentile, 0% for sex, 22% for ethnicity, and 68% for grade level.

Parameter estimates and implied simple effects from the final models are shown in Table III. The final models explained ~20% and ~18% of the variance in NeSA math and reading test scores, respectively.

The unique effect of aerobic fitness remained significant for both math and reading after inclusion of all covariates; a significant interaction was also found between aerobic fitness and free/reduced lunch status. For both NeSA math and reading tests, aerobically fit students had greater odds of passing compared with aerobically unfit students regardless of whether the students received free/reduced lunch. However, as indicated by a statistically significant interaction, the unique effect of aerobic fitness differed significantly by a student's free/reduced lunch status. More specifically, aerobically fit students not receiving free/reduced lunch had 2.41 times greater odds of passing the NeSA math test and 2.23 times greater odds of passing the NeSA reading test compared with aerobically unfit students not receiving free/reduced lunch, holding all fixed effects constant and random effects at zero. By comparison, aerobically fit students receiving free/reduced lunch had 1.68 times greater odds of passing the NeSA math test and 1.56 times greater odds of passing the NeSA reading test compared with aerobically unfit students receiving free/reduced lunch, holding all fixed effects constant and random effects at zero. Thus, for both NeSA math and reading tests, the effect of being aerobically fit was significantly greater for students not receiving free/reduced lunch compared with students receiving free/reduced lunch.

BMI percentile was removed from the final models for both NeSA tests as it became nonsignificant after including free/reduced lunch status. This indicated that the unique association between BMI percentile and passing the NeSA tests

Table I. Demographic and descriptive statistics of students and schools

Student characteristics	Fit students (n = 8116)		Unfit students (n = 3627)		
	Mean (SD)	Range	Mean (SD)	Range	*P*
BMI percentile	60.05 (28.37)	0.01-99.96	78.28 (26.92)	0.01-99.86	<.05
	n (%)		n (%)		
Passed NeSA math	6532 (80.5)		2386 (65.8)		<.05
Passed NeSA reading	6845 (84.3)		2585 (71.3)		<.05
Free/reduced lunch	3039 (37.4)		2031 (56.0)		<.05
Boys	3646 (44.9)		2281 (62.9)		<.05
White	5846 (72.0)		2391 (65.9)		<.05
Grade					<.05
4th	1702 (21.0)		885 (24.4)		
5th	1829 (22.5)		725 (20.0)		
6th	1780 (21.9)		576 (15.9)		
7th	1424 (17.5)		703 (19.4)		
8th	1381 (17.0)		738 (20.3)		
Middle school	4585 (56.5)		2017 (55.6)		.37

School characteristics	Mean (SD)	Range
No. of students	454.38 (280.56)	36-844
BMI percentile	65.68 (5.27)	53.92-81.34
Proportion passing NeSA math	0.76 (0.14)	0.45-0.95
Proportion passing NeSA reading	0.80 (0.10)	0.50-1.00
Proportion fit	0.69 (0.13)	0.26-0.91
Proportion receiving free/reduced lunch	0.43 (0.26)	0.04-0.94
Proportion boys	0.50 (0.03)	0.42-0.65
Proportion white	0.70 (0.18)	0.22-0.93

Table II. Bivariate correlations between all variables included in analysis

	Passed math	Passed reading	Fitness	FRL	BMI percentile	Boys	White
Passed reading	0.77	-	-	-	-	-	-
Fitness	0.27	0.27	-	-	-	-	-
FRL	−0.26	−0.25	−0.17	-	-	-	-
BMI percentile	−0.08	−0.09	−0.38	0.14	-	-	-
Boys	0.01	−0.05	−0.17	−0.01	0.02	-	-
White	0.20	0.19	0.06	−0.41	−0.10	0.01	-
Grade level	−0.17	−0.08	−0.02	−0.05	−0.01	−0.01	0.03

FRL, free/reduced lunch.
Correlations conflate the student- and school-level effects; however, they can be considered to represent student-level correlations given that there were 11 743 students in the current study and only 47 schools. Pass math: 1 = yes, 0 = no; pass reading: 1 = yes, 0 = no; fitness: 1 = fit, 0 = unfit; FRL: 1 = received FRL, 0 = did not receive FRL.

was explained completely by free/reduced lunch status suggesting students from lower–socioeconomic status households, on average, had higher BMI percentiles. Indeed, mean BMI percentile for students receiving free/reduced lunch was 70.22 (SD 28.28) compared with 62.24 (SD 29.37) for students not receiving free/reduced lunch.

Discussion

Because entering the PACER's HFZ was associated with academic outcomes, we suggest that school systems use the PACER as a stand-alone measure of aerobic fitness. Use of the PACER has several advantages over standard physical fitness metrics. The PACER requires substantially less time compared with the complete multiple-component Fitnessgram. In addition, the PACER can be used in more diverse settings than the timed mile because it can be conducted indoors with less space, a feature attractive for school systems in northern climates as the assessment can be conducted any time during the school year.

There were several limitations to our study. First, the nonexperimental, cross-sectional design limited our ability to determine the causal direction between aerobic fitness and academic success. Second, performance on both the PACER

Table III. Final model results for NeSA math and reading outcomes

	NeSA math (n = 11 729)			NeSA reading (n = 11 742)		
	Estimate	SE	OR	Estimate	SE	OR
Intercept	−7.85	3.33		−3.08	3.11	
Student-level effects						
Fitness; no FRL	0.88	0.08	2.41*	0.80	0.09	2.23*
Fitness; FRL	0.52	0.07	1.68*	0.44	0.07	1.56*
Difference in fitness effects	−0.36	0.10	0.70*	−0.36	0.11	0.70*
FRL	−0.57	0.08	0.56*	−0.65	0.09	0.52*
Boys	−0.03	0.08	0.97	−0.22	0.05	0.80*
White	0.35	0.07	1.41*	0.48	0.07	1.61*
Grade level (0 = 6th)	−0.13	0.05	0.88*	−0.12	0.04	0.89*
Boys × white	0.27	0.10	1.31*	-	-	-
Boys × grade level	-	-	-	−0.09	0.04	0.92*
School-level effects (contextual)						
Fitness (0 = 0.70)	−0.77	0.82	0.46	0.04	0.68	1.04
FRL (0 = 0.45)	−1.05	0.84	0.35	0.15	0.71	1.17
Boys (0 = 0.50)	1.12	2.07	3.06	−1.46	6.06	0.23
White (0 = 0.70)	0.22	1.04	1.24	1.55	0.88	4.71
Grade level (0 = 6th)	−5.88	2.22	0.00*	−0.46	2.77	0.63
Middle school	14.05	5.46	1.26E+06*	3.26	6.06	26.16
Fitness × FRL	−0.61	2.44	0.54	0.08	2.04	1.09
Boys × white	−4.22	9.74	0.01	-	-	-
Boys × grade level	-	-	-	−2.40	4.08	0.09
Cross-level Interactions						
Student grade × school grade	-	-	-	0.16	0.04	1.18*
School grade × school grade	-	-	-	1.44	0.90	4.21

Random effects	Estimate	SE	Estimate	SE
Intercept variance	0.15	0.05	0.20	0.08
Grade slope variance	0.03	0.02	-	-
Intercept–grade slope covariance	−0.02	0.02	-	-
White slope variance	-	-	0.04	0.04
Intercept–white slope covariance	-	-	−0.06	0.05

Fitness: 1 = fit, 0 = unfit; free/reduced lunch: 1 = received free/reduced lunch; 0 = did not receive free/reduced lunch; middle school: 1 = middle school, 0 = elementary school.
*Statistically significant effects ($P < .05$). Model fit statistics for math: −2LL = 11 065.61; Akaike information criterion = 11 256.55; for reading: −2 log-likelihood = 10 244.70; Akaike information criterion = 10 405.45.

and standardized academic testing are in part dependent on the motivation and efforts of students. Third, the data were not gathered by trained researchers with standardized equipment. However, there was no reason to suspect variation in administration and assessment from school personnel would result in differential bias of aerobic fitness results as Morrow et al[20] indicated strong agreement between experts and trained teachers on PACER testing (κ = .84).

Our study indicated a significant association between aerobic fitness and academic performance that varied depending on whether the student received free/reduced lunch. Further, BMI percentile did not significantly predict passing either academic outcome after adjusting for free/reduced lunch status. Subsequent research should be longitudinal evaluating how changes in aerobic fitness and BMI affect academic performance. The cross-sectional data used in the current study were taken from the second year of our districtwide tracking. In addition, future research would also benefit from inclusion of both intervention and control arms, as piloting intervention efforts in select schools is the next step to determine whether efforts to improve aerobic fitness can narrow the achievement gap observed in these students. ■

The authors thank Dr Lesa Hoffman (Department of Psychology, University of Nebraska-Lincoln) for her helpful comments regarding the statistical analysis.

Submitted for publication Jun 14, 2012; last revision received Dec 17, 2012; accepted Jan 3, 2013.

Reprint requests: Robert R. Rauner, MD, MPH, Partnership for a Healthy Lincoln, 4600 Valley Rd, Suite 250, Lincoln, NE 68510. E-mail: brauner@healthylincoln.org

References

1. Pedersen BK. Body mass index-independent effect of fitness and physical activity for all-cause mortality. Scand J Med Sci Sports 2007;17:196-204.
2. Gill JMR, Malkova D. Physical activity, fitness and cardiovascular disease risk in adults: interactions with insulin resistance and obesity. Clin Sci 2006;110:409-25.
3. Fogelholm M. Physical activity, fitness and fatness: relations to mortality, morbidity and disease risk factors. A systematic review. Obes Rev 2010;11:202-21.
4. Chomitz VR, Slining MM, McGowan RJ, Mitchell SE, Dawson GF, Hacker KA. Is there a relationship between physical fitness and academic achievement? Positive results from public school children in the northeastern United States. J Sch Health 2009;79:30-7.
5. Castelli DM, Hillman CH, Buck SM, Erwin HE. Physical fitness and academic achievement in third- and fifth-grade students. J Sport Exerc Psychol 2007;29:239-52.
6. Roberts CK, Freed B, McCarthy WJ. Low aerobic fitness and obesity are associated with lower standardized test scores in children. J Pediatr 2010;156:711-8.
7. Welk GJ, Jackson AW, Morrow JR, Haskell WH, Meredith MD, Cooper KH. The association of health-related fitness with indicators of academic performance in Texas schools. Res Q Exerc Sport 2010;81:S16-23.
8. Nebraska Legislature. Quality education accountability act. 79 USC § 757-762. http://uniweb.legislature.ne.gov/laws/browse-chapters.php?chapter=79. Accessed May 18, 2012.
9. No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425.
10. Data Recognition Corporation. Nebraska state accountability (NeSA) technical report. http://www.education.ne.gov/assessment/pdfs/Complete_NeSA_2011_Technical_Report.pdf. Accessed May 17, 2012.
11. The Cooper Institute. Fitnessgram & activitygram test administration manual–updated. 4th ed. Champaign, IL: Human Kinetics; 2010.
12. Varness T, Carrel AL, Eickhoff JC, Allen DB. Reliable prediction of insulin resistance by a school-based fitness test in middle-school children. Int J Pediatr Endocrinol 2009;487804.
13. Fitnessgram Healthy Fitness Zones. http://www.cde.ca.gov/ta/tg/pf/documents/healthfitzone09.pdf. Accessed September 10, 2012.
14. PACER Conversion Chart. http://www.fitnessgram.net/PACER_conversion.pdf. Accessed September 10, 2012.
15. Snijders TAB, Bosker RJ. Multilevel analysis: an introduction to basic & advanced multilevel modeling. 2nd ed. Thousand Oaks, CA: Sage; 2011.
16. Hox J. Multilevel analysis. Techniques and applications. 2nd ed. New York: Routledge; 2010.
17. Raudenbush SW, Bryk AS. Hierarchical linear models: Applications and data analysis methods. 2nd ed. Newbury Park, CA: Sage; 2002.
18. McKelvey R, Zavoina W. A statistical model for the analysis of ordinal level dependent variables. J Math Sociol 1975;4:103-20.
19. Long JS. Regression models for categorical and limited dependent variables. Thousand Oaks, CA: Sage; 1997.
20. Morrow JR, Martin SB, Jackson AW. Reliability and validity of the FITNESSGRAM: Quality of teacher-collected health-related fitness surveillance data. Res Q Exerc Sport 2010;81:S24-30.