Estimating Fitness Bias in Body Mass Index of Middle School Students

Stephen E. Erfle

International Business and Management, Dickinson College

P. O. Box 1773, Carlisle, PA, 17013, USA

E-mail: erfle@dickinson.edu

Phone: 1 (717) 245-1635

Fax: 1 (717) 245-1854

Abstract

Since muscle is more dense than fat, athletes tend to have greater mass and BMI than similarly sized non-athletes. Comparing direct adiposity measures and BMI confirms that BMI is a biased proxy for adiposity for elite athletes. A similar bias should exist for non-elite athletes as well as fit individuals.

This paper provides a methodology for indirectly estimating the size of the fitness bias in BMI using median physical activity performances. Approximately 30% of females and 33% of males are fit using this definition.

Using data from 9 062 students, regressions suggest 3.1, 95% CI [0.9, 5.3], of a female's BMI percentile of 85, and 3.6% of her weight, CI [1.6%, 5.6%], is due to being fit, but 5.6, CI [3.3, 7.9], of a male's BMI percentile of 85, and 5.9% of his weight, CI [3.9%, 7.9%], is due to being fit. These increases in weight are smaller than, but consistent with, the bias of more than 20% for elite athletes.

Strong performance on individual physical activities decreases BMI percentile and BMI, but doing well on multiple physical activities has the reverse effect. This provides evidence of a fitness bias. BMI report cards should include the caveat that BMI may overstate the adiposity status of fit children.

Keywords: Obesity, physical fitness, health report cards

1. Introduction

Body mass index (BMI = kg·m^{-2}) is a population measure used to define obesity in children and adults due to its ease of measurement, its inexpensiveness, and its relatively noninvasive nature. BMI does not measure body fat directly but it does correlate to direct adiposity measures (Mei et al., 2002; Sweeting, 2007). Given that BMI is an indirect measure, it is not surprising that it is an imperfect proxy for adiposity.

Athletes tend to have greater muscle mass than non-athletes and muscle mass is more dense than fat, hence there is a bias in BMI with athletes having a higher BMI compared to similarly sized non-athletes (Prentice & Jebb, 2001; Ode, Pivarnik, Reeves, & Knous, 2007). This has led coaches and trainers to question the validity of BMI as a measure of health risk among athletes (Riewald, 2008; Wein & Palmer, 2008).

Nevill et al. (2010) used skinfold thickness and BMI data from elite (Olympic) athletes in seven sports with age-matched controls to examine the adjustments required for elite athletes that would allow BMI for athletes to reflect the adiposity in nonathletic populations. They documented adjustments in the range of 21% to 39% that differed by sport with middle-distance runners requiring a greater adjustment in BMI than other sports studied (including lightweight- and heavyweight-rowers, long-distance runners and triathletes) (Nevill et al., 2010). A substantial athletic bias exists in interpreting BMI for elite athletes.

Few individuals are elite athletes. A similar, but less pronounced bias should exist for athletes of non-elite status but this has not been examined in the literature. More generally, BMI may overstate the adiposity status of physically fit individuals of all ages. The current analysis examines whether such a bias exists for young adolescent males and females.

One recent review of the literature suggests that this may be behind the lack of association between physical activity and BMI in adolescent males (Reichert, Baptista Menezes, Wells, Carvalho Dumith, & Hallal, 2009). Adolescence is noted as a time when youth typically become less physically active (Dumith, Gigante, Domingues, & Kohl, 2011) and males and females develop at different rates as they enter puberty (Centers for Disease Control and Prevention, 2010). Males increase muscle mass and reduce body fat, while females increase body fat due to hormonal changes in puberty (Knutson, 2005). These differences in maturation are considered when calculating BMI percentile among adolescents and thus it seems reasonable to hypothesize that differences may exist with regards to bias in BMI between adolescent males and females. The current study builds on the nascent literature regarding bias in BMI by providing a methodology for indirectly measuring the size of this bias. This study uses that methodology to provide evidence that a fitness bias exists among middle school students using individual student data from 30 schools in Pennsylvania.

2. Methods

2.1 Participants

Pennsylvania Department of Health (PADoH) launched the Active Schools Program (ASP) to encourage daily physical activity in middle schools across the Commonwealth and to assess the change in physical activity performance due to daily PE across a school year (Erfle & Gamble, 2015). ASP schools agreed to institute a minimum of 30-minutes of daily physical education and administer a physical fitness assessment at the beginning and the end of the

2009-2010 academic year. The data gathered from the fall 2009 assessment were used for this analysis. PADoH received fall 2009 assessments from 11 932 students at 37 middle schools. The participants were evenly distributed across sex; 49.5% female and 50.5% male.

2.2 Instruments

PADoH provided assessment protocols to ASP instructors. This protocol has received Institutional Review Board approval from the Dickinson College IRB. School representatives were required to participate in a webinar on assessment protocols and use of the reporting template to ensure minimal bias in implementation. School nurses measured height and weight using established PADoH protocols (Department of Health, 2013). PADoH required ASP schools to administer physical activity tests on a variety of fitness dimensions at the start of year and end of year. The tests assessed included the mile run, curl-ups, and push-ups. Students had 60 seconds to perform as many curl-up repetitions as possible. For the push-up test, students were instructed to do push-ups until failure. Demographic and anthropometric data included, sex, age, grade, height, and weight. PADoH gathered data using a modified version of an Excel file created by the Centers for Disease Control for use in schools (Centers for Disease Control and Prevention, 2009).

2.3 Procedure

Removal of 1 914 students with missing or invalid data reduced the sample to 10 018 participants with full data. Seven schools were removed after preliminary analysis due to extreme values for those schools. Six of these schools reported excessive mile run times for students across all obesity status categories and one school reported implausibly large curl-ups performances. Both of these situations suggested to ASP administrators that the data from these schools was suspect and should be removed from further analysis. Removal of these schools decreased the sample by 956, with a final sample of 9 062 used in subsequent analyses. Excel was used for data cleaning and SPSS was used for statistical analysis. A 5% significance level was used for all tests.

Students were categorized as fit using performance on physical activities. Students were categorized in the top half on an activity if their performance exceeded the median performance for their Sex $\times$ Grade. A student who was in the top half on the mile run and push-ups was said to be in the Fit 4th and a student who was in the top half of all three activities was said to be in the Fit 8th.

2.4 Data Analysis

This paper uses a cross-sectional analysis of the correlates of physical activity and BMI instead of longitudinal analysis that focuses on the temporal relation between activity and adiposity (Reichert et al., 2009). Because BMI percentile is age- and sex-adjusted, it allows easier comparison across sex and age than BMI. Nonetheless, BMI regressions provide the ability to calculate expected weight change from altering performance on physical activities. Both indices are examined using ordinary least squares regression analysis.

It is appropriate to transform BMI percentile because it is a limited dependent variable. The logistic transform, $L = \ln(\text{BMI percentile}/[100 - \text{BMI percentile}])$, takes on values from $-\infty$ to $+\infty$ as BMI percentile ranges from 0 to 100. It is worth noting that this model differs from the logit model (which is based on a categorical dependent variable).

Estimates to how BMI and BMI percentile will change as a given physical activity outcome changes may depend on both age and sex. Separate regressions are provided for each sex, grade dummy variables are included to control for grade-to-grade cohort differences and age was included to control for within-grade age variation. Both indices should be a negative function of physical activity performance, all else held equal. However, it would not be surprising if there were nonlinearities involved with regard to each activity. As a result, each regression includes a quadratic term for each activity.

Fitness bias can be assessed by including a dummy variable using the definition of fitness described above. If strong performance on multiple dimensions occurs, is BMI percentile further reduced or does it increase? If BMI percentile increases, then a fitness bias has been established.

2.4.1 Interpreting slope from quadratic coefficients

If the activity is x (curl-ups, mile, or push-ups) and $f(x) = bx + cx^2/100$ describes the effect of that activity on f (f is L or BMI), then the slope at x, $m(x)$, is the derivative of f with respect to x, $m(x) = \mathrm{d}f(x)/\mathrm{d}x = b + 0.02cx$.

2.4.2 Interpreting slope from logistically transformed models

The estimated slope coefficients in the L models describe how L changes as x changes, $\Delta L/\Delta x$. These slopes are constant in terms of L but L is a nonlinear function of BMI percentile, therefore, they are not constant in terms of BMI percentile. A scaling factor, S(BMI percentile), that connects L to BMI percentile is required to transform $\Delta L/\Delta x$ slopes into ΔBMI percentile/Δx slopes. The scaling factor for the logistic transform is given by:
S(BMI percentile) = BMI percentile·(100 - BMI percentile)/100.

2.4.3 Interpreting slope from BMI models as percentage change in weight

The estimated slope coefficients in the BMI models in Table 2 provide a best guess change in BMI associated with a 1 unit change in each independent variable x, ΔBMI/Δx. Change in BMI, ΔBMI, can be interpreted as a percentage change in weight, %ΔW, by dividing by the BMI of the student under consideration, %$\Delta W = \Delta$BMI/BMI.

3. Results

3.1 Defining Fitness

Table 1 presents median values (with standard deviation in parenthesis) for BMI, BMI percentile, and three physical activity performances organized by Sex × Grade. A student was categorized as being in the fit half of an activity if that student had above median performance for their Sex × Grade. This was done for each of the three activities.

Table 1. Performance by Sex × Grade

Variable	Grade			Total
	6	7	8	
		Female		
Body mass indices				
BMI percentile	76.1 (29.0)	74.4 (28.6)	75.1 (27.6)	75.1 (28.3)
BMI	20.1 (4.9)	20.5 (5.2)	21.4 (5.1)	20.8 (5.1)
Physical activities				
Curl-ups	30 (13.9)	31 (13.4)	32 (14.4)	31 (13.9)
Push-ups	7 (10.2)	8 (9.3)	8 (9.0)	8 (9.4)
Mile	12.12 (2.78)	11.50 (2.76)	11.25 (2.86)	11.58 (2.83)
n	1 026	1 768	1 701	4 495
		Male		
Body mass indices				
BMI percentile	77.5 (28.9)	71.8 (30.1)	74.1 (29.7)	73.9 (29.7)
BMI	19.7 (4.9)	19.9 (5.0)	20.8 (5.0)	20.2 (5.0)
Physical activities				
Curl-ups	34 (13.8)	38 (14.3)	38 (15.1)	37 (14.6)
Push-ups	14 (13.1)	16 (13.1)	17 (12.7)	16 (13.0)
Mile	11.14 (3.07)	9.95 (2.87)	9.78 (2.65)	10.10 (2.9)
n	1 100	1 819	1 648	4 566

Note. Median performance with standard deviation in parentheses.

Each student was placed into in one of eight cells of the resulting $2 \times 2 \times 2$ performance partition based on whether her or his performances placed that student into the fit or unfit half on each activity. Figure 1 describes two attributes of students within these two partitions. The upper panel depicts relative BMI percentile and the lower depicts cell frequency. In each panel, the right-most cell (column) is the Fit 8th and the two right-most cells combined are the Fit 4th.

The frequency distributions in Figure 1 suggest that students are not evenly distributed within each partition. Students tend to demonstrate superior or inferior performance on multiple physical activities (especially mile and push-ups) and not one single activity. More than 20 percent are in the Fit 8th and approximately one third are in the Fit 4th. The upper panel of Figure 1 depicts systematic differences in average BMI percentile across cells. These average differences however, hide differences in the distribution of BMI percentile across fitness groups.
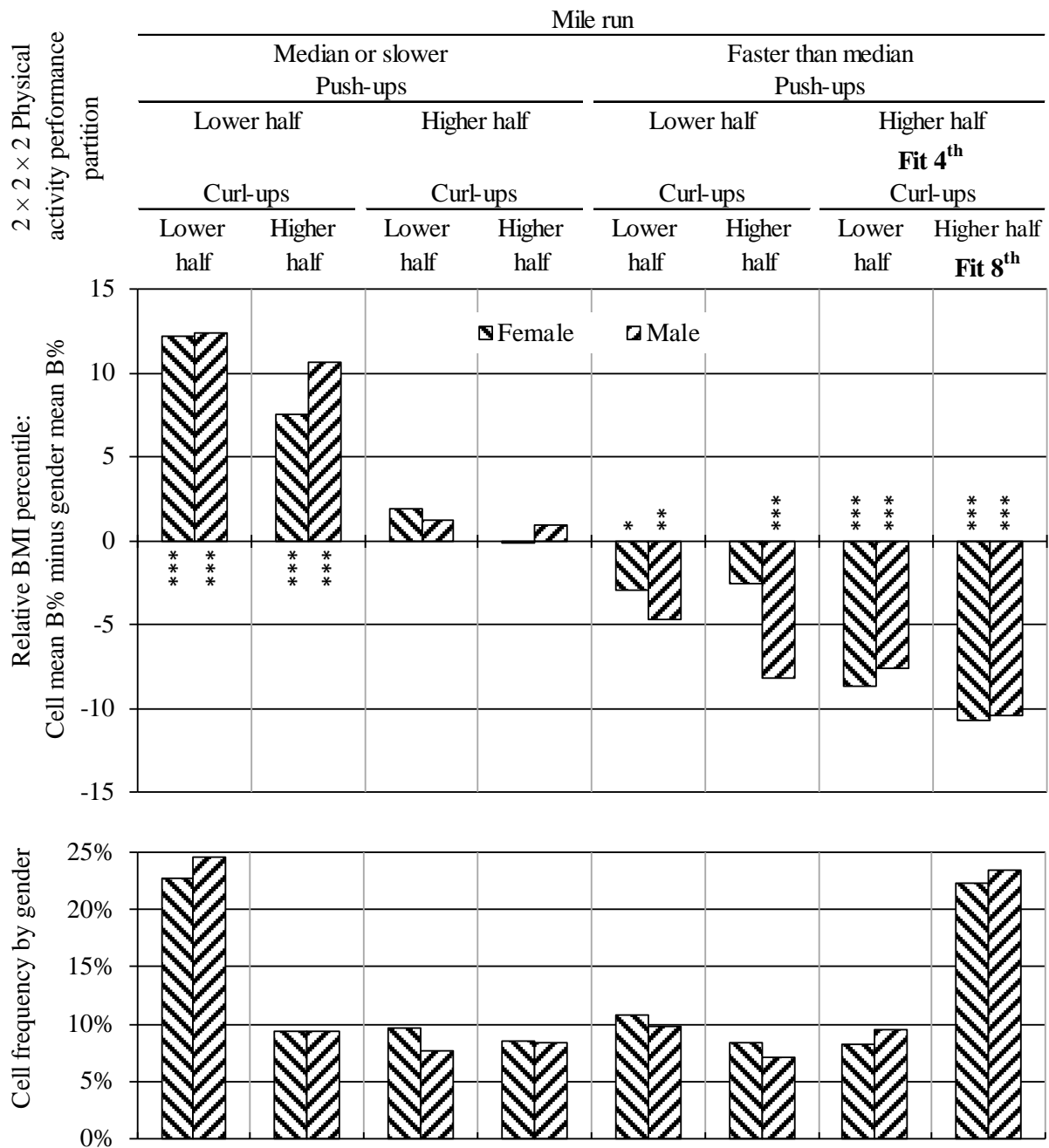
Mile run

	Median or slower				Faster than median			
	Push-ups				Push-ups			
	Lower half		Higher half		Lower half		Higher half **Fit 4th**	
	Curl-ups		Curl-ups		Curl-ups		Curl-ups	
	Lower half	Higher half	Lower half	Higher half	Lower half	Higher half	Lower half	Higher half **Fit 8th**

2 × 2 × 2 Physical activity performance partition (left vertical label)

Upper panel y-axis: Relative BMI percentile: Cell mean B% minus gender mean B% (15 to -15)

Legend: ◪ Female ▨ Male

Lower panel y-axis: Cell frequency by gender (0% to 25%)

Figure 1. Relative BMI percentile (B%) and frequency across sex-specific $2 \times 2 \times 2$ grade-adjusted physical activity performance partitions. Cell performance cut-points based on Table 1 with median performance included in each lower half. Relative B% is cell mean B% minus gender mean B% of 67.2 for females and 66.1 for males. Statistically significant differences are denoted via *s with $*P < .05$, $**P < .01$, $***P < .001$. Individual cell frequencies are relative to all females or all males. Students are labeled fit based performances on mile and push-ups (*Fit 4th*) or all three activities (*Fit 8th*).

3.2 Distribution of BMI Percentile across Fitness Groups

Figure 2 depicts three 100% basis BMI percentile histograms for each sex. The top panel shows the distribution of BMI percentile for the full sample and the lower shows two fractional fitness subsamples.
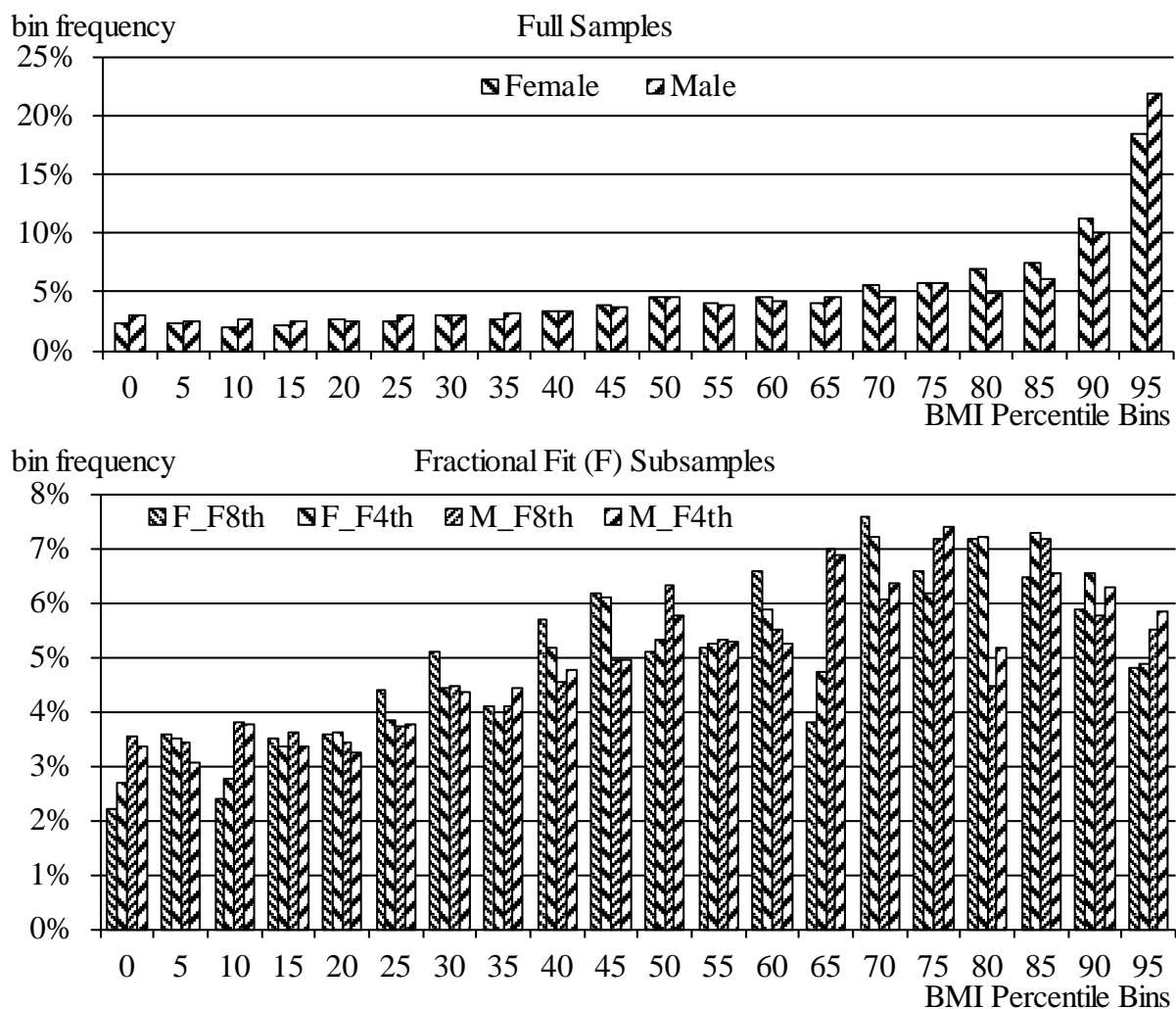
6

Figure 2. Frequency histograms of BMI percentile by sex for full sample and two fitness subsamples. The alternative fractional definitions of fitness are based on the partitions in Figure 1. Each of the six frequency histograms is on its own 100% basis. The left portion of each bin is female (F) and the right is male (M). Bin labels are the bottom BMI percentile in the 5 percentile wide bin.

The female and male full sample histograms are strongly skewed and provide a visual representation of Pennsylvania's pediatric obesity crisis. There are more overweight females (bins 85 and 90) and there are more obese males (bin 95) (Barlow, 2007).

Fit fraction subsamples show more commonality than difference across fractions and sex with increasing representation across bins until about the 80th percentile. Fit students are distributed more than proportionately in the upper half of the BMI percentile distribution with approximately 60% of fit students having BMI percentile of 50 or higher, more than one sixth have BMI percentile of 85 or higher, and approximately 5% have BMI percentile of 95 or higher. Given that no substantive differences emerge between the two fractional fitness definitions in Figure 2, this study focuses on the Fit 4th for further analysis. In the absence of significant differences, the parsimonious solution is to use the definition that requires the least information.

3.3 Regression analysis

Table 2 reports four regression analyses, one for each sex using L and BMI as the dependent variable. Fitness bias is assessed by including a dummy variable equal to 1 when the student's performance on the mile run and push-ups places that student in the Fit 4th of that Sex $\times$ Grade as described in Figure 1.

Table 2. Logistic(BMI Percentile) and BMI as a Function of Performance on Three Activities by Sex

Variable	Logistic(BMI percentile), L		BMI	
	Female	Male	Female	Male
Constant	0.49 (-1.19, 2.18)	-1.28 (-2.86, 0.30)	12.1 (7.71, 16.5)	5.69 (1.94, 9.45)
Curl - ups	-0.04 (-0.05, -0.03)	-0.01 (-0.02, 0.002)	-0.14 (-0.18, -0.11)	-0.05 (-0.08, -0.01)
Curl-ups^2/100	0.05 (0.03, 0.07)	0.01 (-0.002, 0.03)	0.17 (0.12, 0.22)	0.05 (0.01, 0.08)
Push-ups	-0.07 (-0.09, -0.05)	-0.06 (-0.07, -0.04)	-0.21 (-0.25, -0.17)	-0.17 (-0.19, -0.14)
Push-ups^2/100	0.09 (0.05, 0.13)	0.05 (0.03, 0.07)	0.32 (0.22, 0.42)	0.17 (0.12, 0.21)
Mile	0.37 (0.25, 0.49)	0.57 (0.46, 0.68)	0.75 (0.43, 1.07)	1.24 (0.98, 1.50)
Mile^2/100	-0.66 (-1.11, -0.21)	-1.32 (-1.75, -0.89)	-0.57 (-1.75, 0.61)	-2.35 (-3.36, -1.33)
Age	-0.15 (-0.28, -0.03)	-0.11 (-0.23, 0.003)	0.37 (0.05, 0.70)	0.59 (0.31, 0.87)
Grade 7	0.15 (0.05, 0.24)	0.03 (0.0002, 0.06)	0.48 (0.23, 0.72)	0.10 (0.02, 0.17)
Grade 8	0.17 (0.07, 0.26)	0.07 (0.03, 0.11)	0.46 (0.21, 0.71)	0.17 (0.08, 0.26)
Fit 4th	0.24 (0.07, 0.41)	0.44 (0.26, 0.62)	0.80 (0.35, 1.25)	1.28 (0.85, 1.71)
R^2	.163	.181	.211	.234
F	87.5	101	120	139
n	4 495	4 567	4 495	4 567

Note. Abbreviation: Body mass index, BMI. Raw regression coefficients (with 95% confidence interval in parentheses). All regressions have significant F statistic at $P < .001$ level.

The performance on three physical activities together with age and fitness are strong consistent predictors of BMI and BMI percentile. The multiple coefficient of variation, R^2, shows that these models predict approximately one sixth of the variation in L and more than one fifth of the variation in BMI. The male models are modestly more predictive than the female models.

For curl-ups and push-ups, the negative linear and positive quadratic coefficients together with the location of the bottom of each estimated parabola in the upper level of that activity distribution implies that increasing curl-up and push-up performances have decreasing returns in terms of expected change in BMI percentile or BMI. Because increased mile run performance means lower mile run times, the positive linear and negative quadratic coefficients together with the location of the top of the estimated parabola in the lower level of mile run performance (slow mile run times) implies that increasing mile run performance has increasing returns in terms of expected change in BMI percentile or BMI.

To understand the relative importance of the various parts of these regression models, it is worthwhile to create a geometric interpretation of those parts. This will allow comparison of

individual physical fitness impacts to one another as well as to age within cohort (grade), cross-cohort variations, and fitness.

3.3.1 Physical Activity Performance, Gender, and Grade Differences

Figure 3 shows the estimated best guess impact and 95% CI on L (ΔL in the upper panel) and BMI (ΔBMI in the lower panel) by sex for seven items of interest. The first three ($\blacklozenge$) are physical activity impacts, the second three ($\lozenge$) examine age, and the final ($\blacksquare$) examines fitness.
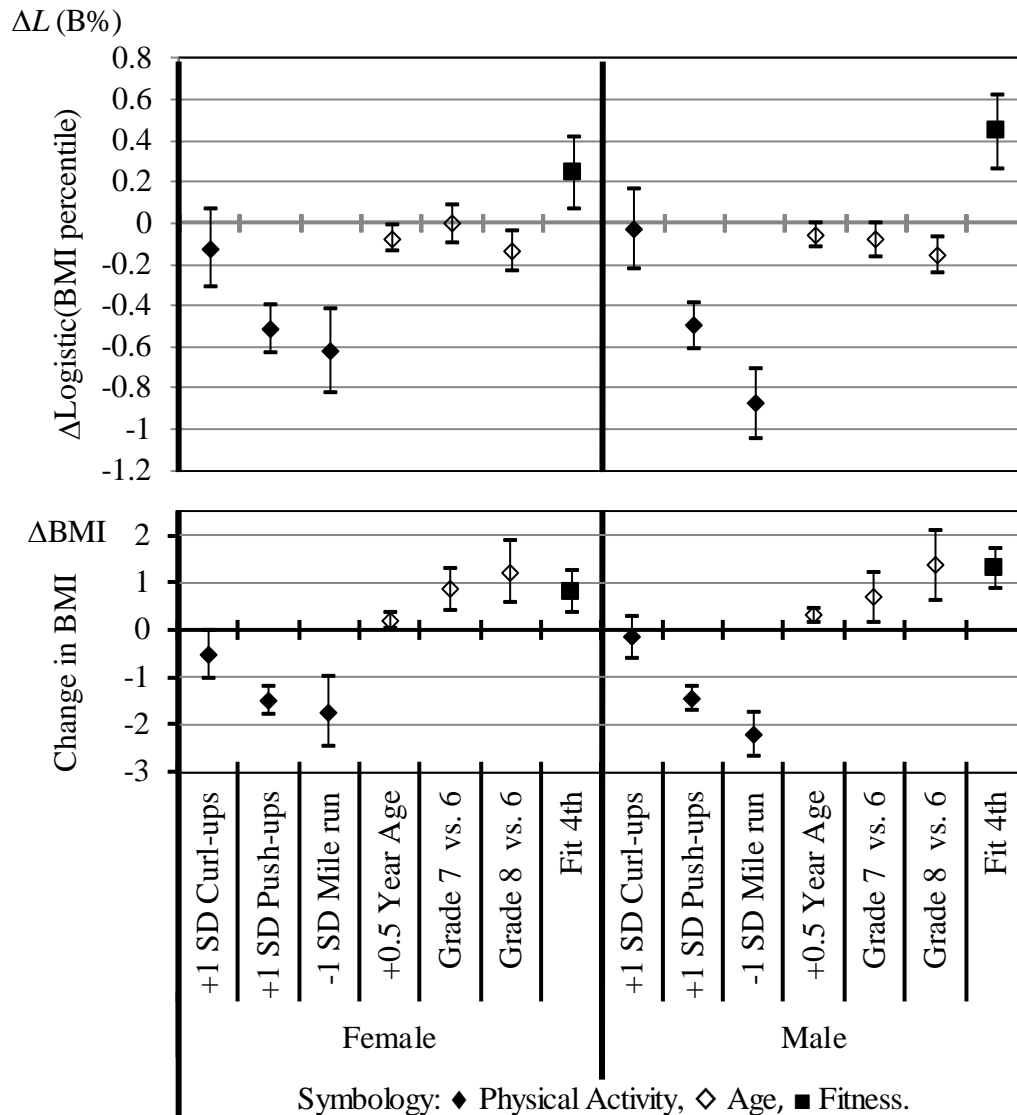
Figure 3. Comparative analysis of the effect of changes in activity, age and fitness on logistic of BMI percentile (L(B%)) and BMI by sex. Each activity effect calculation is based on linear and quadratic activity coefficients from Table 2, evaluated at median activity and standard deviation levels by sex from Table 1. Grade-to-grade comparisons incorporate age adjustments using the coefficients in Table 2. Fit 4th coefficients from Table 2 are based on the definition created using the physical activity performance partitions by sex in Figure 1. Each 95% confidence interval is based on the significance level of the linear activity coefficient in Table 2.

The raw regression coefficients for physical fitness activities in Table 2 are difficult to compare across measures for two reasons. First, these activities are differentially difficult to achieve; a one minute decrease in mile run time is more difficult to achieve than one more push-up or one more curl-up. Second, given the specification of each activity as a quadratic function, it also depends on the initial level of that activity. A balanced method to compare across physical fitness activities is to start at the median level of each measure and to consider a 1 standard deviation (SD) change in that measure.

One SD physical activity impacts are calculated for each activity for both sexes and both dependent variables. For example, the best guess impact on L of a 1 SD increase in mile run performance from median mile run time for males is approximately ΔL = -0.9 in the upper panel of Figure 3. This value is calculated as -0.88 = 0.30·-2.9 based on mile run slope at the male median mile run time of 10.10 minutes of 0.30 = 0.57 + (0.02·-1.32·10.10) using the quadratic slope calculation in section 2.4.1 and using a 1 SD decrease in mile run time of 2.9 minutes from Table 1.

To examine how within-grade age differences affect L and BMI, the Age coefficient is multiplied by 0.5 (within-grade SD of Age is 0.44). The Grade 7 and Grade 8 comparisons with Grade 6 include increases in Age of 1 and 2 years respectively. Finally, both panels include the Fit 4th coefficient from each model in Table 2.

Both panels in Figure 3 depict similar activity effect patterns across sex and grade. The rank order of effects on BMI and BMI percentile across activities is mile run, push-ups, curl-ups. Males exhibit greater benefit from mile run and females from curl-ups with regard to decreasing BMI. Grades 7 and 8 are both higher than Grade 6 for BMI which is not sex and age adjusted but not on the sex-and-age-adjusted BMI percentile metric (which therefore controls for variations in stages of adolescence). The +0.5 year Age effect suggests that older students within a grade have lower BMI percentile but higher BMI than their younger classmates.

3.3.2 Fitness Differences

Table 2 and Figure 3 suggest that superior performance on each individual activity is associated with lower BMI and BMI percentile but superior performance on multiple dimensions, as modeled by Fit 4th, is associated with higher BMI and BMI percentile. Fit 4th coefficients are significant in all four models (P = .006 for the female L model and P < .001 for the other three models). A fitness bias has been established.

The estimated magnitude of the fitness bias using the Fit 4th criterion is substantial. The Fit 4th ΔL magnitudes in the upper panel and the ΔBMI magnitudes in the lower panel of Figure 3 are roughly the mean size of the 1 SD physical activity impacts for males and sixty percent of the mean size of these estimates for females. Fit 4th magnitudes are larger than the grade to grade comparisons in the BMI percentile models but are smaller than those comparisons in the BMI models. The BMI result is not surprising given that adolescence is a time when BMI naturally trends upward (Centers for Disease Control and Prevention, 2010).

As discussed in sections 2.4.2 and 2.4.3, ΔL can be transformed into ΔBMI percentile and ΔBMI can be transformed into percentage change in weight by multiplying by the appropriate scaling factors. Consider two Fit 4th students with a BMI percentile of 85 (who are therefore considered borderline overweight) (Barlow, 2007). The female's Fit 4th coefficient of 0.24 in Table 2 means that a best guess is that 3.1, 95% CI [0.9, 5.3], of her BMI percentile of 85 is

due to being in the Fit 4th because $S(85) = 85 \cdot 15/100 = 12.75$ and $3.1 = 12.75 \cdot 0.24$. The same calculation implies the male's coefficient of 0.44 means that a best guess is that 5.6, CI [3.3, 7.9], of his BMI percentile of 85 is due to being in the Fit 4th. Put another way, a more accurate description of the female's BMI percentile would be 81.9 and a more accurate description of the male's BMI percentile would be 79.4, placing both in the normal range based on their Fit 4th-adjusted BMI percentile.

If these 85 BMI percentile students are of average age for this sample (12.8 years old), then the female's BMI is 22.4 and the male's is 21.6 (Centers for Disease Control and Prevention, 2010). Dividing their respective Fit 4th BMI coefficients in Table 2 by these values yields a best guess of 3.6%, 95% CI [1.6%, 5.6%], of the female's weight is due to being in the Fit 4th and a best guess of 5.9% of the male's weight is due to being in the Fit 4th, CI [3.9%, 7.9%].

4. Discussion

A common critique of BMI is that it is it does not accurately reflect adiposity, most notably for athletes because athletes have a greater amount of muscle mass and muscle mass is more dense than fat mass (Ernsberger, 2012). This has led coaches and trainers to question the validity of BMI as a measure of health risk among athletes (Riewald, 2008; Wein & Palmer, 2008). Substantial athletic bias in elite athletes has been documented by comparing direct adiposity measures and BMI(Nevill et al., 2010); however, this is the first investigation to examine whether a similar bias occurs in adolescents.

This paper examines whether a fitness bias exists using physical activity performances among a sample of more than 9 000 students from 30 Pennsylvania middle schools. It analyzes the association between three physical activities and two measures used to describe overweight and obesity, BMI percentile and BMI, and provides indirect evidence that a fitness bias in BMI percentile and BMI exists in children.

The same hierarchical patterns continues to emerge regarding the relative importance of the three physical activities analyzed in explaining both body mass indices, even when viewed from a variety of perspectives. The most important is the mile run, followed by push-ups and curl-ups. Increasing performance on push-ups and curl-ups decreases BMI percentile and BMI at a decreasing rate but increasing performance on the mile does so at an increasing rate. The mile tends to have a greater relative effect on males while curl-ups tend to have a greater relative effect on females.

Students are placed in a partition based on their sex- and grade-adjusted performance on three activities. Two definitions of fit and unfit performance are examined, one uses all three activities and the other uses the mile run and push-ups. Histograms of BMI percentile frequency for both versions of fit suggest that little is gained from defining fitness using all three activities – it is sufficient to define fitness on the basis of the mile run and push-ups. Regression analysis using this definition of fitness confirms that a fitness bias exists for both sexes.

No clear rationale requires using median performance to create the partition that forms the basis for this analysis. Partitioning students using median performance on push-ups and the mile run produces a situation in which more than 30% of students are defined as fit. Erfle and Gelbaugh (2013) used this partition to examine physical activity performance differences by focal middle school students. A more restrictive definition might produce evidence of a more substantial bias, but the fact remains that statistically significant results were obtained with

this weak, inclusive, definition. Another cut-off may produce superior models to those presented here. For example, an analysis of Taiwanese students suggested using the 43% of students who scored higher than the lowest quartile on 4 physical activities to define the fitter subgroup on which BMI norms could be based (Chen et al., 2002). A similar analysis could easily be performed using performance cut-offs provided by external sources such as the *President's Challenge* which provides percentile boundaries for various activities by age.

A fitness bias may well exist for younger and older students and for adults of non-elite athletic stature. The estimated magnitude of this bias in middle-school students is consistent with, but smaller than, the elite athletic bias of more than 20% documented by direct measurement (Nevill et al., 2010). This indirect evidence should be supplemented by studies in which direct measurement of adiposity are compared with BMI and BMI percentile for individuals at various levels of physical activity performance. Such direct adiposity measurement would provide firm evidence of the existence of a fitness bias in children that has been inferred from this analysis of BMI and physical activity data.

It would be instructive to examine the components of BMI as fat mass and fat-free mass for fit and unfit performers. One would expect a substantially higher fat-free component for fit performers (Freedman et al., 2005). Of interest is whether that component remains relatively constant across BMI categories for different classes of physical activity performers. If such an association were found, it could be used to modify BMI interpretation protocols employed by health professionals and physical educators.

Some states have implemented mandatory BMI report cards for children in their state (Evans & Sonneville, 2009; Thompson & Card-Higginson, 2009). If parents are provided with BMI report cards for their children, then those report cards should be provided with the caveat that if their child is a strong physical activity performer, then the BMI on their report card may overstate their adiposity status. The present study provides some preliminary guidance regarding the size of that overstatement.

5. Conclusion

This paper provides evidence that BMI is a biased predictor of adiposity for young fit adolescents, just as coaches have long suspected (Jonnalagadda, Skinner, & Moore, 2004). This paper provides a methodology for indirectly testing whether a bias in BMI exists and provides estimates of the size of that bias. Using a definition of fitness in which more than 30% of students are defined as fit, this paper provides statistically significant support for the presence of a fitness bias of approximately 3.5% for females and 6% for males.

5.1 Practical Applications

- BMI is a population measure used to define obesity in adults and youth. In children, BMI percentile provides a sex- and age-adjusted measure of obesity. Both are indirect measures that correlate to direct adiposity measures. This paper provides empirical evidence that these measures are biased predictors of adiposity for physically fit adolescents.

- School health professionals and physical educators can use this information to target lifestyle physical activity behaviors.

- This also provides the opportunity for coaches and trainers to educate athletes about healthy body weight, what it means, and why they are considered overweight based on BMI.

- Coaches, trainers and school health professionals may be less skeptical of BMI if they are armed with empirical estimates of how large the fitness bias in BMI is likely to be. Coaches would no longer have to complain that they simply do not trust BMI because it says that they, as well as the students they are coaching, are overweight or obese despite evidence to the contrary.

Acknowledgements

References

Barlow, S. E. (2007). Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: Summary report. *Pediatrics, 120*, S164-S192. http://dx.doi.org/10.15421peds.2007-2329C

Centers for Disease Control and Prevention. (2009). Children's BMI tool for schools. Retrieved from http://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/tool_for_schools.html

Centers for Disease Control and Prevention. (2010). National center for health statistics. CDC growth charts: United States. Retrieved from http://www.cdc.gov/growthcharts/

Chen, W., Lin, C. C., Peng, C. T., Li, C. I., Wu, H. C., Chiang, J., . . . Huang, P. C. (2002). Approaching healthy body mass index norms for children and adolescents from health-related physical fitness. *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity, 3*(3), 225-232.

Department of Health, P. (2013). Procedures for the growth screening program for PA's school-age population manual. Retrieved from http://www.portal.state.pa.us/portal/server.pt/community/schools/14130/mandated_school_health_program_(exams_screens)/556692

Dumith, S. C., Gigante, D. P., Domingues, M. R., & Kohl, H. W.,3rd. (2011). Physical activity change during adolescence: A systematic review and a pooled analysis. *International Journal of Epidemiology,* http://dx.doi.org/10.1093/ije/dyq272

Erfle, S. E., & Gamble, A. (2015). Effects of daily physical education on physical fitness and weight status in middle school adolescents. *J Sch Health, 85*(1), 27-35. http://dx.doi.org/10.1111/josh.12217

Erfle, S. E., & Gelbaugh, C. M. (2013). Physical activity performance of focal middle school students. *Measurement in Physical Education & Exercise Science, 17*(2), 150-166. http://dx.doi.org/10.1080/1091367X.2013.761034#.UjXzxH9huSo

Ernsberger, P. (2012). BMI, body build, body fatness, and health risks. *Fat Studies, 1*(1), 6-12. http://dx.doi.org/10.1080/21604851.2012.627788

Evans, E. W., & Sonneville, K. R. (2009). BMI report cards: Will they pass or fail in the fight against pediatric obesity? *Current Opinion in Pediatrics, 21*(4), 431-436. http://dx.doi.org/10.1097/MOP.0b013e32832ce04c

Freedman, D. S., Wang, J., Maynard, L. M., Thornton, J. C., Mei, Z., Pierson, R. N., . . . Horlick, M. (2005). Relation of BMI to fat and fat-free mass among children and adolescents. *International Journal of Obesity, 29*(1), 1-8. http://dx.doi.org/10.1038/sj.ijo.0802735

Jonnalagadda, S. S., Skinner, R., & Moore, L. (2004). Overweight athlete: Fact or fiction? *Current Sports Medicine Reports, 3*(4), 198-205.

Knutson, K. L. (2005). Sex differences in the association between sleep and body mass index in adolescents. *The Journal of Pediatrics, 147*(6), 830-834. http://dx.doi.org/10.1016/j.jpeds.2005.07.019

Mei, Z., Grummer-Strawn, L. M., Pietrobelli, A., Goulding, A., Goran, M. I., & Dietz, W. H. (2002). Validity of body mass index compared with other body-composition screening indexes for the assessment of body fatness in children and adolescents. *The American Journal of Clinical Nutrition, 75*(6), 978-985.

Nevill, A. M., Winter, E. M., Ingham, S., Watts, A., Metsios, G. S., & Stewart, A. D. (2010). Adjusting athletes' body mass index to better reflect adiposity in epidemiological research. *Journal of Sports Sciences, 28*(9), 1009-1016.
http://dx.doi.org/10.1080/02640414.2010.487071

Ode, J. J., Pivarnik, J. M., Reeves, M. J., & Knous, J. L. (2007). Body mass index as a predictor of percent fat in college athletes and nonathletes. *Medicine and Science in Sports and Exercise, 39*(3), 403-409. http://dx.doi.org/10.1249/01.mss.0000247008.19127.3e

Prentice, A. M., & Jebb, S. A. (2001). Beyond body mass index. *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity, 2*(3), 141-147.

Reichert, F. F., Baptista Menezes, A. M., Wells, J. C., Carvalho Dumith, S., & Hallal, P. C. (2009). Physical activity as a predictor of adolescent body fatness: A systematic review. *Sports Medicine (Auckland, N.Z.), 39*(4), 279-294.

Riewald, S. (2008). Does the body mass index accurately reflect percent body fat in athletes? *Strength & Conditioning Journal (Lippincott Williams & Wilkins), 30*(1), 80-81.

Sweeting, H. N. (2007). Measurement and definitions of obesity in childhood and adolescence: A field guide for the uninitiated. *Nutrition Journal, 6*, 32.
http://dx.doi.org/10.1186/1475-2891-6-32

Thompson, J. W., & Card-Higginson, P. (2009). Arkansas' experience: Statewide surveillance and parental information on the child obesity epidemic. *Pediatrics, 124 Suppl 1*, S73-82.
http://dx.doi.org/10.1542/peds.2008-3586J

Wein, D., & Palmer, K. (2008). Does a high BMI mean an unhealthy athlete? *NSCA's Performance Training Journal, 7*(4), 20-21.

Copyright Disclaimer