Population-Level Measures to Predict Obesity Burden in Public Schools: Looking Upstream for Midstream Actions

Wasantha P. Jayawardene, MD, PhD1, David K. Lohrmann, PhD1, Stephanie Dickinson, MAS2, and Mohammad R. Torabi, PhD1

Abstract

Purpose: To estimate school-level obesity burden, as reflected in prevalence of obesity, based on the characteristics of students’ socioeconomic and geographic environments.

Design: Secondary analysis of cross-sectional data.

Setting: Public schools (N = 504) from 43 of 67 counties in Pennsylvania.

Participants: Kindergarten through grade 12 students (N = 255,949).

Measures: School-level obesity prevalence for the year 2014 was calculated from state-mandated student body mass index (BMI) measurements. Eighteen aggregate variables, characterizing schools and counties, were retrieved from federal data sources.

Analysis: Three classification variables—excess weight (BMI ≥ 85th percentile), obesity (BMI ≥ 95th percentile), and severe obesity (BMI > 35% or 120% of 95th percentile)—each with 3 groups of schools (low-, average-, and high-prevalence) were created for discriminant function analysis, based on state mean and standard deviation of school distribution. Analysis tested each classification model to reveal school- and county-level dimensions on which school groups differed from each other.

Results: Discriminant functions for obesity, which contained school enrollment, percentage of students receiving free/reduced-price lunch, percentage of black/Hispanic students, school location (suburban/other), percentage of county adults with post-secondary education, and percentage of county adults with obesity, yielded 67.86% correct classification (highest accuracy), compared to 34.23% schools classified by chance alone.

Conclusion: In the absence of mandated student BMI screenings, the model developed in this study can be used to identify schools most likely to have high obesity burden and, thereafter, determine dissemination of enhanced resources for the implementation of proven prevention policies and programs.

Keywords
demographic factors, socioeconomic factors, geographic factors, classification, schools, obesity

Introduction

In 2012, 34.2% of US elementary school-aged children and 34.5% of secondary school-aged adolescents were overweight; respectively, 17.7% and 20.5% were obese.1 Obesity burden is experienced disproportionately across demographic subgroups and differs across communities, schools, and regions.2 Since African American race, Hispanic ethnicity, low household income, and parental education plus living in inner city are associated with overweight,3 schools in neighborhoods with higher proportions of such families likely have higher rates of child overweight and obesity.2 Ethnic disparities related to childhood obesity are widening with rates increasing faster among non-Hispanic black and Mexican American boys than non-Hispanic white boys, and among non-Hispanic black girls than non-Hispanic white girls.4 These social environment characteristics (ie, upstream determinants) influence behaviors and health status of children and families.5,6

1 Applied Health Science, School of Public Health Bloomington, Indiana University, Bloomington, IN, USA
2 Epidemiology and Biostatistics, School of Public Health Bloomington, Indiana University, Bloomington, IN, USA

Corresponding Author:
Wasantha P. Jayawardene, Applied Health Science, School of Public Health Bloomington, Indiana University, PH 116, 1025 E 7th Street, Bloomington, IN 47405, USA.
Email: wajayawa@indiana.edu
Population-level obesity prevention includes multilevel school and child health policies (ie, upstream interventions), school- and community-based programs (ie, midstream interventions), and need-based individualized behavioral approaches (ie, downstream interventions). Schools are important settings for multicomponent preventive programs that improve weight status of children and adolescents by addressing attitudes and behaviors related to physical inactivity, healthy dietary habits, and mental health. Prevention policy and program implementation varies, likely contributing to overweight and obesity prevalence differences across schools. In consideration of nationwide interventions, schools involved in on-site Healthy Schools Program demonstrated a trend toward decreased overweight (−0.48%) and obesity (−0.42%) with each additional contact with the program and each additional year of exposure. From a statewide perspective, students exposed to strong, specific, competitive food laws in 2003, compared to students in states with no such laws, gained 0.25 less body mass index (BMI) units on average from fifth to eighth grade and were less likely to remain overweight or obese; also, if laws remained consistently strong over a 3-year period, students gained 0.44 fewer BMI units compared to students in states with no such laws. In consideration of school-based interventions, healthier school environments led to 15.2% relative reduction in obesity prevalence among elementary school students over 6 years, while the number of obesity prevention strategies implemented was negatively associated with overweight and obesity prevalence over time. Therefore, student overweight and obesity variability among public schools are attributed to differences in health disparities across social environments (ie, upstream determinants) or school-based obesity prevention outcomes (ie, midstream interventions) or both.

Contextually, BMI-based identification of children at greatest health risk due to excess body fat is essential. Prior studies found that the obesity prevalence is higher in states with school-based BMI surveillance mandates; although causality was not established, this is likely due to the circumstance that such states consider the obesity problem as more serious and, consequently, enact BMI-screening policies. Despite limitations, estimation of childhood obesity burden through school-based BMI measurement by trained personnel can identify high-risk groups, monitor progress toward achieving objectives, and foster policy and/or program changes. Such initiatives provide specific obesity burden estimates used to inform design and implementation of school-level efforts; interventions should be responsive to the existing obesity burden within schools. Twenty states require BMI surveillance and 9 recommend periodic body composition or fitness screenings. However, even within states with BMI screening mandates, all schools do not consistently collect BMI data. Thus, development of a clearly articulated model of health disparities' rational links between social environments (ie, upstream determinants) and school-level BMI outcomes is warranted to allow better approximation of obesity burden within schools when BMI data are unavailable.

Interventions that are funded and designed based on well-informed obesity burden estimates, accounting for empirically established links between health disparities and varying obesity rates across schools, can ultimately improve outcomes, cost-effectiveness, and community benefits.

**Purpose**

The objective of this study was to determine school-level obesity burden, based on the characteristics of the social, economic, and geographic environments of students. The research question was: “Can publicly available school-level and population-level variables predict school classification by student obesity prevalence?”

**Methods**

**Design**

The study involved a secondary analysis of cross-sectional data.

**Sample**

Multisource data from 43 of 67 Pennsylvania counties were accessed. Geographically, counties included 19 of the 20 most populous as well as all of Pennsylvania’s 10 largest cities. The sample consisted of 504 public elementary and secondary schools that reported measured height and weight data of 255,949 students in grades K-12 for the year 2014.

**Measures**

Eighteen potential discriminating (predictor) variables were identified—9 at school level and 9 at county level. Public school characterization variables were total students, percentage of students eligible for free lunch (≤130% of federal poverty threshold), percentage of students eligible for reduced-price lunch (≤185% of federal poverty threshold), percentage of black students, percentage of Hispanic students, percentage of male, school-level (eg, elementary, middle, high), student/teacher ratio, and school location based on urban-centric locale codes (1 = city large, 2 = city midsize, 3 = city small, 4 = suburb large, 5 = suburb midsize, 6 = suburb small, 7 = town fringe, 8 = town distant, 9 = town remote, 10 = rural fringe, 11 = rural distant, and 12 = rural remote). County-related variables were county population, percentage of African American (black), percentage of Hispanic, percentage of adults with postsecondary education, county median household income, percentage of population with limited healthy foods access, percentage of single parent households, percentage of low-birth-weight live births, and percentage of obese adults. Table 1 indicates the data source for each discriminant variable retained in the final model. Excess weight, obesity, and severe obesity were considered classification (dependent) variables. Centers for Disease Control and Prevention (CDC)-developed growth charts for children and adolescents (2000 as the growth reference year for
calculation of percentiles) were used to classify BMIs. Obesity burden variables—percentage of excess weight students (≥85th percentile), percentage of obese students (≥95th percentile), and percentage of severely obese students (≥120% of the 95th percentile or BMI > 35 regardless of age)—were derived from 2014 student height and weight measurements. State-mandated height and weight measurements were taken by qualified personnel (eg, school nurses), using established protocols, and entered into an electronic health record.

Data were compiled in 3 repositories maintained by Population Health Innovations, LLC and supported by the Highmark School’s location according to urban-centric locale codes. Nonpublicly available de-identified BMIs were accessed under a data sharing agreement. Data were compiled in 3 repositories maintained by Population Health Innovations, LLC and supported by the Highmark.

**Table 1.** Publicly Available Data Sources for the Discriminant Variables in the Final Model.

<table>
<thead>
<tr>
<th>Discriminant Variable</th>
<th>Publicly Available Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student enrollment in school</td>
<td>US Department of Education → Institute of Education Sciences → National Center for Education Statistics. (Search for Public Schools) <a href="http://nces.ed.gov/ccd/schoolsearch/">http://nces.ed.gov/ccd/schoolsearch/</a></td>
</tr>
<tr>
<td>Number of African American and Hispanic students</td>
<td>Centers for Disease Control and Prevention (CDC) → Division of Diabetes Translation → National Center for Chronic Disease Prevention and Health Promotion. (County Data) <a href="http://www.cdc.gov/diabetes/atlases/countydata/atlus.html">http://www.cdc.gov/diabetes/atlases/countydata/atlus.html</a></td>
</tr>
<tr>
<td>School’s location according to urban-centric locale codes</td>
<td></td>
</tr>
<tr>
<td>Percentage of county adults with some postsecondary education</td>
<td></td>
</tr>
<tr>
<td>Percentage of county adults who were obese</td>
<td></td>
</tr>
</tbody>
</table>

**Results**

For 43 counties, average prevalence of excess weight, obesity, and severe obesity was 36.00%, 18.85%, and 6.42%, respectively. For each retained school, percentage of excess weight, obese, and severely obese students were calculated; SDs of those distributions were 4.77%, 3.83%, and 1.74%, respectively. Ranges were 18.42% to 52.04%, 6.25% to 32.43%, and 0.00% to 19.46%. Obesity-based classification, compared to excess-weight- and severe-obesity-based classifications, yielded the highest rate of correctly classified schools (67.86% correct classification vs 34.23% classified by chance alone; 33.63% improvement; Table 2). Cross validation was done by the leave-one-out method, where each school was classified by the functions derived from all schools other than that school. Cross validation of obesity-based classification revealed a 66.18% correct classification (31.95%
Table 2. Classification of Schools Based on Excess Weight, Obesity, and Severe-Obesity Burdens.\(^{a,b}\)

<table>
<thead>
<tr>
<th>Actual Group Membership</th>
<th>School Group</th>
<th>Predicted Group Membership</th>
<th></th>
<th></th>
<th></th>
<th>Total Count (%)(^{c})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low-Prevalence Count (Row %)</td>
<td>Average-Prevalence Count (Row %)</td>
<td>High-Prevalence Count (Row %)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excess weight classification</td>
<td>Original Low excess weight</td>
<td>99 (74.4)</td>
<td>22 (16.5)</td>
<td>12 (9.0)</td>
<td>133 (27.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average excess weight</td>
<td>31 (17.6)</td>
<td>100 (56.8)</td>
<td>45 (25.6)</td>
<td>176 (37.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High excess weight</td>
<td>6 (3.6)</td>
<td>59 (35.3)</td>
<td>102 (61.1)</td>
<td>167 (35.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correct classification</td>
<td>99 (32.9)</td>
<td>100 (33.2)</td>
<td>102 (33.9)</td>
<td>301 (63.2)</td>
<td></td>
</tr>
<tr>
<td>Cross validated Low excess weight</td>
<td>99 (74.4)</td>
<td>21 (15.8)</td>
<td>13 (9.8)</td>
<td>133 (27.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average excess weight</td>
<td>32 (18.2)</td>
<td>97 (55.1)</td>
<td>47 (26.7)</td>
<td>176 (37.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High excess weight</td>
<td>6 (3.6)</td>
<td>66 (39.5)</td>
<td>95 (56.9)</td>
<td>167 (35.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correct classification</td>
<td>99 (34.0)</td>
<td>97 (33.3)</td>
<td>95 (32.6)</td>
<td>291 (61.1)</td>
<td></td>
</tr>
<tr>
<td>Obesity classification</td>
<td>Original Low obesity</td>
<td>98 (75.4)</td>
<td>27 (20.8)</td>
<td>5 (3.8)</td>
<td>130 (27.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average obesity</td>
<td>22 (11.4)</td>
<td>139 (72.0)</td>
<td>32 (16.6)</td>
<td>193 (40.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High obesity</td>
<td>2 (1.3)</td>
<td>65 (42.5)</td>
<td>86 (50.6)</td>
<td>152 (32.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correct classification</td>
<td>98 (30.3)</td>
<td>139 (43.0)</td>
<td>86 (26.6)</td>
<td>322 (67.9)</td>
<td></td>
</tr>
<tr>
<td>Cross validated Low obesity</td>
<td>98 (75.4)</td>
<td>27 (20.8)</td>
<td>5 (3.8)</td>
<td>130 (27.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average obesity</td>
<td>25 (13.0)</td>
<td>134 (69.4)</td>
<td>34 (17.6)</td>
<td>193 (40.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High obesity</td>
<td>4 (2.6)</td>
<td>66 (43.1)</td>
<td>83 (54.2)</td>
<td>153 (31.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correct classification</td>
<td>98 (31.1)</td>
<td>134 (42.5)</td>
<td>83 (26.3)</td>
<td>315 (66.2)</td>
<td></td>
</tr>
<tr>
<td>Severe-obesity classification</td>
<td>Original Low severe obesity</td>
<td>116 (71.2)</td>
<td>38 (23.3)</td>
<td>9 (5.5)</td>
<td>163 (34.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average severe obesity</td>
<td>38 (21.8)</td>
<td>95 (54.6)</td>
<td>41 (23.6)</td>
<td>174 (36.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High severe obesity</td>
<td>12 (8.6)</td>
<td>74 (53.2)</td>
<td>53 (38.1)</td>
<td>139 (29.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correct classification</td>
<td>116 (43.9)</td>
<td>95 (36.0)</td>
<td>53 (20.1)</td>
<td>264 (55.5)</td>
<td></td>
</tr>
<tr>
<td>Cross validated Low-severe obesity</td>
<td>115 (70.6)</td>
<td>38 (23.3)</td>
<td>10 (6.1)</td>
<td>163 (34.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average-severe obesity</td>
<td>39 (22.4)</td>
<td>91 (52.3)</td>
<td>44 (25.3)</td>
<td>174 (36.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High-severe obesity</td>
<td>12 (8.6)</td>
<td>76 (54.7)</td>
<td>51 (36.7)</td>
<td>139 (29.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correct classification</td>
<td>115 (44.7)</td>
<td>91 (35.4)</td>
<td>51 (19.8)</td>
<td>257 (54.0)</td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\)N = 476.

\(^{b}\)For excess weight, obesity, and severe obesity, percentage of correctly classified schools by chance alone was 33.8%, 34.2%, and 35.8%, respectively; discriminant functions improved these random classifications by 29.4%, 33.6%, and 19.7%, respectively. Cross validation was done by the leave-one-out method, where each school was classified by the functions derived from all schools other than that school. Each school was removed from the analysis one at a time, where the discriminant function analysis was refit to the remaining school data. The school which was removed from the analysis was then reclassified into 1 of 3 groups (ie, low, average, and high prevalence). Using the new discriminant function equations. This procedure was then repeated for each school to obtain the cross-validation frequencies and percentages. Refer to Table 4 for prediction equations.

\(^{c}\)Total number of schools (N = 476) was used as the denominator.

improvement). Since classification accuracy was established as the criterion for selecting the classification variable, only results from the obesity-based classification are presented and discussed below. School- and county-level variable ranges proved wide (Table 3).

The final model for obesity-based classification, providing the highest discriminatory power with the least number of variables, contained 6 independent variables (5 continuous and 1 binary; Table 1): square root of total students (sq-#students), square root of percentage of free and reduced-price lunch eligible students (sq-poverty), logarithm of percentage of black and Hispanic students (log-minority), school’s location as suburban according to urban-centric locale codes (locale-suburb; 1 = suburban or 0 = urban or rural), percentage of county adults with some postsecondary education (adult-educ), and percentage of county adults who were obese (adult-obese). Sq-#students, sq-poverty, log-minority, locale-suburb, adult-educ, and adult-obese separated low-, average-, and high-obesity schools (P = .001 for sq-#students and P < .001 for other variables). At school level, log-minority positively correlated with sq-poverty (r = 0.598); other correlations were very weak (r < 0.25). At the county level, adult-educ negatively correlated with adult-obese (r = −0.742).

Two discriminant functions were identified: sq-poverty (r = 0.922) alone was loaded on the first, whereas all other variables, that is log-minority (r = −0.690), adult-obese (r = 0.623), adult-educ (r = −0.536), locale-suburb (r = −0.474), and sq-#students (r = 0.332), were loaded on the second. The first function (r3 = 0.709), based on the percentage of free and reduced-price lunch eligible students, maximally separated low-obesity schools from the rest, while also separating high-obesity schools from average-obesity schools to some extent (Figure 1). The second function (r3 = 0.207), based on black/Hispanic percentage, adult obesity, adult education level, school location (ie, suburban or not), and school size, maximally separated average-obesity schools from the rest.

Standardized canonical discriminant function coefficients were estimated to standardize the distribution of scores from each function, utilizing a mean of 0 and SD of 1. The absolute value of these coefficients indicated the relative contribution of discriminating variables to function 1 and function 2. In function 1, greatest contribution was from sq-poverty (1.066),
followed by adult-educ ($-0.269$), log-minority ($-0.259$),
adult-obese ($-0.156$), and locale-suburb ($-0.116$), while the
least important predictor was sq-#students ($-0.004$). In func-
tion 2, greatest contribution was from log-minority ($0.684$),
followed by sq-#students ($0.549$), adult-obese ($0.499$), locale-
suburb ($0.347$), and adult-educ ($0.182$), while the least impor-
tant predictor was sq-poverty ($0.122$).

Based on unadjusted group means (Table 3), low-obesity
schools had highest student enrollment, least poverty, least
percentage of black/Hispanic students, and were more likely
to be in suburban counties with high rates of postsecondary
education and lower rates of adult obesity. The antithetical
situation was observed for high-obesity schools; that is, the
lowest student enrollment, highest poverty, highest percentage
of black/Hispanic students, lowest percentage of suburban
schools, lowest adult postsecondary education rate, and highest
adult obesity rate.

Using classification function coefficients, an equation was
developed for each of the low-, average-, and high-obesity
groups to determine school assignment to group, based on the
scores for 6 variables (Table 4). To improve correct classifi-
cation, utilizing school- and county-level discriminant vari-
ables, a school should be assigned to the group for which it
obtained the highest classification score; for example, if a
rural school (locale-suburb $= 0$) has 225 students
(sq-students $= 15$), with 49% free and reduced-price lunch
eligibility (sq-poverty $= 0.7$), and 11% black and Hispanic
students (log-minority $= 0.9586$), and is located in a county
that has 23% of adults with postsecondary education (adult-
educ $= 0.23$) and 38% of adults who are obese (adult-obese $= 0.38$), the school will score 298 for low obesity, 309 for
average obesity, and 317 for high obesity; hence, the school
will be assigned to the high-obesity group.

Table 3. Population Characteristics (A) and Unadjusted Group Means for Discriminant Variables in the Obesity-Based Classification (B).

<table>
<thead>
<tr>
<th>A. School and County Characteristics</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students in school</td>
<td>497</td>
<td>92</td>
<td>2307</td>
</tr>
<tr>
<td>Proportion of free/reduced price lunch</td>
<td>34.51%</td>
<td>2.19%</td>
<td>99.72%</td>
</tr>
<tr>
<td>Proportion of black and Hispanic students</td>
<td>15.95%</td>
<td>0.00%</td>
<td>99.69%</td>
</tr>
<tr>
<td>County population</td>
<td>408</td>
<td>1229</td>
<td>14772</td>
</tr>
<tr>
<td>Median household income</td>
<td>$US$55 840</td>
<td>$US$38 351</td>
<td>$US$82 456</td>
</tr>
<tr>
<td>Percentage of county adults with postsecondary education</td>
<td>62.14</td>
<td>39.80</td>
<td>76.85</td>
</tr>
<tr>
<td>Percentage of county adults who are obese</td>
<td>28.56</td>
<td>22.20</td>
<td>36.20</td>
</tr>
<tr>
<td>Percentage of county adults with limited access to healthy foods</td>
<td>5.20</td>
<td>0.88</td>
<td>12.39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Discriminant Variable Retained in the Final Model for Obesity-Based Classification</th>
<th>Low-Obesity Schools</th>
<th>Average-Obesity Schools</th>
<th>High-Obesity Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sq-root (number of students in school)</td>
<td>23.51 (6.89)</td>
<td>22.59 (5.83)</td>
<td>20.85 (5.52)</td>
</tr>
<tr>
<td>Sq-root (proportion of free/reduced price lunch)</td>
<td>0.41 (0.13)</td>
<td>0.60 (0.13)</td>
<td>0.72 (0.14)</td>
</tr>
<tr>
<td>Log (proportion of black and Hispanic students)</td>
<td>0.03 (0.03)</td>
<td>0.05 (0.06)</td>
<td>0.09 (0.09)</td>
</tr>
<tr>
<td>Percentage of schools located in suburban areas</td>
<td>83.08 (37.64)</td>
<td>45.60 (49.94)</td>
<td>41.83 (49.49)</td>
</tr>
<tr>
<td>Percentage of county adults with postsecondary education</td>
<td>68.13 (9.19)</td>
<td>60.07 (9.88)</td>
<td>59.66 (10.93)</td>
</tr>
<tr>
<td>Percentage of county adults who are obese</td>
<td>27.10 (2.67)</td>
<td>28.81 (2.85)</td>
<td>29.06 (3.24)</td>
</tr>
</tbody>
</table>

Abbreviation: SD, standard deviation.

Figure 1. Group centroids for unstandardized canonical discriminant functions: Function 1 = student poverty. Function 2 = adult education level, and school location (suburban or not), percentage of black/Hispanic, school size, and adult obesity. The canonical score plot demonstrates how the function 1 (x axis) and function 2 (y axis) classify schools between low-, average-, and high-obesity groups by plotting the observation score, computed via unstandardized canonical discriminant functions. Theoretically, each school is represented by a dot on the plot (476 dots in total); included in the graph are only the group centroid for low-, average-, and high-obesity groups. The first function ($r_c = 0.709$), based on student poverty, maximally separated low-obesity schools from the rest, while it also separated high-obesity schools from average-obesity schools to some extent. The second function ($r_c = 0.207$), based on black/Hispanic percentage, adult obesity, adult education level, school location (ie, suburban or not), and school size, maximally separated average-obesity schools from the rest.
Table 4. School Assignment to Obesity Burden Group, Based on Classification Function Coefficients.

<table>
<thead>
<tr>
<th>Group</th>
<th>Coefficient 1</th>
<th>Coefficient 2</th>
<th>Coefficient 3</th>
<th>Coefficient 4</th>
<th>Coefficient 5</th>
<th>Coefficient 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low obesity</td>
<td>-270.703</td>
<td>0.875</td>
<td>29.891</td>
<td>-54.249</td>
<td>log-minority</td>
<td>10.434</td>
</tr>
<tr>
<td></td>
<td>(sq-#students)</td>
<td></td>
<td></td>
<td></td>
<td>+ log-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>minority</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1068.896</td>
</tr>
<tr>
<td>Average obesity</td>
<td>-274.053</td>
<td>0.908</td>
<td>44.245</td>
<td>-64.827</td>
<td>log-minority</td>
<td>11.141</td>
</tr>
<tr>
<td></td>
<td>(sq-#students)</td>
<td></td>
<td></td>
<td></td>
<td>+ log-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>minority</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1053.822</td>
</tr>
<tr>
<td>High obesity</td>
<td>-272.331</td>
<td>0.865</td>
<td>50.553</td>
<td>-63.296</td>
<td>log-minority</td>
<td>11.002</td>
</tr>
<tr>
<td></td>
<td>(sq-#students)</td>
<td></td>
<td></td>
<td></td>
<td>+ log-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>minority</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>326.048</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1066.032</td>
</tr>
</tbody>
</table>

*For a given school, scores for each of the 6 variables can be obtained from public-use data. In the equations, sq-#students denotes square root of total students; sq-poverty denotes square root of percentage of free and reduced-price lunch eligible students; log-minority denotes logarithm of percentage of black and Hispanic students; locale-suburb denotes school’s location as suburban according to urban-centric locale codes (1 = suburban or 0 = urban or rural); adult-educ denotes percentage of county adults with some postsecondary education; and adult-obese denotes percentage of county adults who were obese. In each of the 3 equations above, a school’s score for a given variable should be multiplied by the classification function coefficient (+ or –) for that variable, and so on, for 6 variables; intercept should also be added. The school should be assigned to the group (low, average, or high obesity) for which it obtained the highest classification score.

Discussion

Study results clearly demonstrate that publicly available school-level and county-level variables can be used to predict school classification by student obesity prevalence; resultant models included a rational combination of publicly available, population-level variables for classification of schools based on obesity burden, which, in comparison with classification by chance alone, almost doubled the percentage of correctly classified schools. Critical variables for determining high-obesity schools were common: lower student numbers, higher student poverty, higher minority student enrollment, nonsuburban location plus location in communities with lower adult education, and higher adult obesity levels. Results demonstrate that reliance on a single variable (eg, percentage of free and reduced-price lunch) is unlikely to provide the best approach to identifying schools in need of more robust interventions, because obesity burden clearly depends on multiple upstream health determinants (eg, race/ethnicity, poverty, and urban/rural divide), along with interactions that may also contribute to this variation (eg, rural poverty).

Nevertheless, free or reduced-price lunch eligibility played a substantial role in classifying schools. Differentiation of low-, average-, and high-obesity schools involved a relatively stronger effect of poverty, compared to the effect of all other variables combined. This finding was unsurprising since previous studies confirmed that poverty leads to inequities in multiple factors (eg, quality housing, access to healthy food, access to quality education, and safe places to be physically active) that contribute to community obesity rates. According to 1999 to 2004 National Health and Nutrition Examination Survey (NHANES), 23% of 15- to 17-year-olds living in poverty were obese, compared to 14% obesity rate among adolescents not living in poverty. However, national data from 1971 to 2002 revealed a weakening association between poverty and childhood obesity over decades, especially among adolescents.

This study’s finding that the percentage of black and Hispanic students significantly predicted school obesity rates is supported by research which found that overweight and obesity rates are higher among black than white children. Furthermore, minority children’s obesity rates are increasing faster at earlier ages; by age 6 to 11, 26.1% of Hispanic children and 23.8% of African American children were obese compared with 13.1% of white children. Almost three-quarters of the difference in rates between Hispanic and white children occurs by third grade; three-quarters of the difference between African American and white children occurs between grades 3 and 8.

Regarding upstream race/ethnicity-related determinants, almost one-fourth of Hispanic and black families had limited healthy food access due to lack of financial or other resources, versus 11% of white families. Fewer black (11.3%) and Hispanic (9.3%) than white adolescents (4.5%) eat vegetables during the prior week. Compared to predominantly white neighborhoods, outdoor advertising for unhealthy foods was 13 times greater in predominantly black neighborhoods and 9 times greater in predominantly Hispanic neighborhoods.

Access to safe and quality public parks, green space, and sidewalks for physical activities was much lower in predominantly black and Hispanic neighborhoods. Substandard neighborhood safety had a strong negative impact on the amount of outdoor play by black girls; Hispanic children engaged in less after-school physical activity due to cost and language barriers.

Although poverty and obesity rates were higher among black and Hispanic families than white families, black race and Hispanic ethnicity are unlikely to be the sole reasons that track poverty; for example, 2005 to 2008 NHANES data revealed a significant inverse association of family income with obesity for white children but an inconsistent relationship for minority children. Contrary to expectations, a 1999 to 2004 NHANES study found higher obesity rates with higher family income among school-aged black children, especially girls. In 2005 to 2008, a similar but nonsignificant association was observed for Mexican American girls. Put differently, most obese children do not live in poverty; 62% of 12 million US obese children are not impoverished. Just 27% of 6 million obese white children are impoverished.

Although the urban-to-rural gradient and urban–rural binary variable did not improve classification of schools in the current study, use of a binary variable that categorized suburban schools versus all other schools significantly improved classification. Previous studies representing 8 states revealed that childhood obesity rates were highest in rural areas. Rural areas have the lowest food location availability, least nutrition education resources, and worst exercise facilities, while rural residents, on average, have lower incomes, lower education...
levels, and limited prevention and treatment options, all contributing to obesity.\textsuperscript{41} Meanwhile, many urban children also lack access to safe parks, playgrounds, and healthy foods and are more exposed to unhealthy foods; income modifies the relationship between food environment and BMI.\textsuperscript{52}

The current study also demonstrated that the percentage of adults with some postsecondary education, even if measured at county level, can contribute to obesity-based school classifications. Empirical evidence suggested that the household head’s education level had a significant negative relationship to obesity prevalence, although not consistent across genders and race/ethnicity groups.\textsuperscript{38} Compared to 21.1\% and 20.4\% obesity prevalence among boys and girls, respectively, living in households where the head had no high school diploma, 11.8\% and 8.3\% obesity rates were found among boys and girls, respectively, who lived in households where the head had a college degree. The relationship between household head’s education and children’s obesity was statistically significant for both white and black girls. From 1988 to 2008, with the exception of girls living in homes where the head had a college degree, obesity prevalence of girls increased significantly at all levels of household head’s education.\textsuperscript{38} Finally, several family- and community-characteristics associated with poverty and adult education potentially contribute to childhood obesity, for example dietary practices, screen time, parental behaviors and attitudes, home environment, and neighborhood physical activity opportunities.\textsuperscript{43} These may be addressed through specific obesity prevention interventions within a multilayered ecological context,\textsuperscript{43-45} although poverty reduction and adult education require broader and more upstream interventions.

In addition to upstream determinants, school characteristics may contribute to obesity prevalence. The current study revealed that, although school type (ie, elementary, middle, or high) did not improve classification of schools for obesity, school enrollment size did. Number of students in school constitutes a proxy measure of both school type and urban–rural divide combined. Elementary schools usually have the least number of students, while high schools have the greatest numbers\textsuperscript{46}; accordingly, elementary schools are more likely to be classified as high obesity; this supports the need for interventions during the earliest formative years. Also, rural schools are more likely to be classified as high obesity because they generally have fewer students than urban and suburban schools. Compared to all US schools in 2010, average student enrollment in Pennsylvania elementary schools was lower, while the averages for middle, high, and other (eg, junior high) schools were much higher, creating a wide range of within-state school sizes.\textsuperscript{46} Small schools in rural areas face unique challenges, including smaller food service programs, teacher shortages, and limited financial resources,\textsuperscript{47} although, compared to large schools, availability and purchase of competitive food and beverages by students in small schools may be favorable.\textsuperscript{48} While school program variables were not considered in this study, previous studies suggest that such midstream actions somewhat correlate with upstream determinants included in this study.\textsuperscript{49} For example, schools with higher poverty rates, compared to others, allowed students to purchase unhealthful competitive foods significantly more often.\textsuperscript{50}

Finally, the current study demonstrated that, in the absence of up-to-date county-level childhood obesity prevalence data, adult obesity prevalence data can be used to help classify schools. Obesity, once established in childhood, tends to persist through adult life\textsuperscript{51} and prevailing environmental conditions and behaviors conducive to weight gain in childhood usually operate in adult life. Thus, states and counties with relatively higher rates of adult obesity also tend to have higher rates of childhood obesity.

**Limitations**

This study has several limitations. Prior research revealed that children with excess body fat can be identified reasonably well using BMI criteria.\textsuperscript{52} However, interpretation of overweight (≥85th percentile but <95th percentile) in children utilizing age-sex-based BMI alone may not be fully accurate because some individuals, especially male adolescents, can be categorized as overweight due to high lean body mass rather than having excess fat.\textsuperscript{53} A similar misclassification is highly unlikely at ≥95th percentile, because assessment of body fat using standard methods (eg, dual-energy X-ray absorptiometry) revealed that almost all children identified as obese via BMI have excess body fat.\textsuperscript{54} From a study design perspective, not all public schools from each county provided student BMI, data were not available for all students in selected schools, and public schools were excluded from the analysis if demographic data were unavailable. Therefore, the possibility that systematic bias that occurred in school selection cannot be excluded. Height and weight data were collected for state-mandated health screenings, not for research purposes. Although height and weight measured by trained professionals, such as school nurses, are likely to be accurate, full-time nurses are unavailable in many schools.\textsuperscript{55} Additionally, many nurses believe that BMI surveillance is an added burden to their workload\textsuperscript{56}; personnel assigned to assist nurses with BMI surveillance need technical training for measuring height and weight accurately.\textsuperscript{57,58} Electronic and beam balance scales used to accurately measure weight (spring balance scales are not suitable) should be calibrated properly and regularly to the nearest one-fourth pound, following manufacturer’s directions, and stadiometers should measure height to the nearest one-eighth inch.\textsuperscript{58-60} Adherence to auxiliary personnel training and equipment quality-control standards for measuring student weight and height in all involved Pennsylvania schools cannot be fully guaranteed. Finally, although overall discriminant function classification accuracy was high, functions tended to overclassify low-obesity schools.

**Conclusions**

Estimation of the obesity burden in schools utilizing publicly available school-level and community-level aggregate variables identified both the magnitude of the problem in a given
school and related associations or childhood obesity risk factors. Thus, the models from this study can be used to (1) identify schools most likely to have high obesity levels, in the absence of routine student BMI screenings and (2) inform dissemination of enhanced resources for implementation of proven and robust prevention policies and programs. Public schools, clearly identified and supported by CDC as an integral part of the public health system, are expected to provide opportunities for students to adopt healthy lifestyles, regardless of socioeconomic status or ethnicity, even when communities and families are unable to do so. As mentioned in the introduction section, limited research have proven that school-based interventions (conducted at national, state, and local levels) were beneficial in reducing obesity rates. While reduction of student obesity rates is an added challenge for schools, especially high-poverty schools that struggle with meeting current academic standards, doing so is in schools’ best interest since children who were physically fit and ate healthily had lower absenteeism, concentrated better on academic tasks, and received better grades, while childhood obesity was associated with several health and social problems that can negatively affect the academic performance.

**Acknowledgments**

We thank Christina L. Wilds, DrPH, of Highmark Foundation and Robert G. Gillio, MD, of Population Health Innovations, LLC, for providing data from the Health eTools for Schools database and their support in the preparation of this article. We also thank Godfred Antwi for his contribution to data extraction from public-use databases.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**References**


---

**So What? Implications for Health Promotion Practitioners and Researchers**

**What is already known on this topic?**

Empirical studies and body mass index (BMI) surveillance programs have confirmed the magnitude and health determinants of childhood obesity plus variations in childhood obesity burden across communities.

**What does this article add?**

This is the first study to estimate school-level obesity burden based on the characteristics of the socioeconomic and geographic environments in which students live and learn.

**What are the implications for health promotion practice and research?**

Even within those 20 states having BMI screening mandates, all schools do not consistently collect BMI data. Estimation of schools’ obesity burden utilizing publicly available school- and county-level aggregate variables identified both the magnitude of the problem in a given school and related associations and/or obesity determinants. Health promotion practitioners can utilize this unique information to allocate greater physical and human resources for implementation of prevention policies and programs in schools that likely have high obesity rates.


44. Khan LK, Sobush K, Keener D, et al. Recommended Community Strategies and Measurements to Prevent Obesity in the United States. Atlanta, GA: Division of Nutrition, Physical Activity, and


