Environmental Research 166 (2018) 100–107

Contents lists available at ScienceDirect

Environmental Research

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

Influence of school environments on childhood obesity in California

Alberto M. Ortega Hinojosa[a,b], Kara E. MacLeod[c], John Balmes[a], Michael Jerrett[a,c,*]

[a] School of Public Health, University of California, Berkeley, Berkeley, CA 94720, United States
[b] IMPAQ International, LLC, Oakland, CA 94612, United States
[c] Fielding School of Public Health, University of California, Los Angeles, Los Angeles, CA 90095, United States

ARTICLE INFO

Keywords:
Obesity
Schools
Built environment
Social environment
Surveillance

ABSTRACT

Objective: To conduct a state-wide examination of public schools and the school neighborhood as potential targets for environmental public health tracking to address childhood obesity.
Methods: We examined the relationship of social and physical environmental attributes of the school environment (within school and neighborhood) and childhood obesity in California with machine learning (Random Forest) and multilevel methods. We used data compiled from the California Department of Education, the U.S. Geological Survey, ESRI's Business Analyst, the U.S. Census, and other public sources for ecologic level variables for various years and assessed their relative importance to obesity as determined from the statewide Physical Fitness Test 2003 through 2007 for grades 5, 7, and 9 (n = 5,265,265).
Results: In addition to individual-level race and gender, the following within and school neighborhood variables ranked as the most important model contributors based on the Random Forest analysis and were included in multilevel regressions clustered on the county. Violent crime, English learners, socioeconomic disadvantage, fewer physical education (PE) and fully credentialed teachers, and diversity index were positively associated with obesity while academic performance index, PE participation, mean educational attainment and per capita income were negatively associated with obesity. The most highly ranked built or physical environment variables were distance to the nearest highway and greenness, which were 10th and 11th most important, respectively.
Conclusions: Many states in the U.S. do not have school-based surveillance programs that collect body mass index data. System-level determinants of obesity can be important for tracking and intervention. The results of these analyses suggest that the school social environment factors may be especially important. Disadvantaged and low academic performing schools have a higher risk for obesity. Supporting such schools in a targeted way may be an efficient way to intervene and could impact both health and academic outcomes. Some of the more important variables, such as having credentialed teachers and participating in PE, are modifiable risk factors.

1. Introduction

In the U.S. more than a third of children and adolescents are overweight or obese (Ogden et al., 2014). Childhood overweight and obesity are major risks for serious youth outcomes and the effects can persist beyond childhood. Health outcomes include asthma, cancer, cardiovascular disease, type II diabetes, hypertension, and depression (Dietz, 1998). It is projected that poor diet and inactivity will soon overtake tobacco as the leading risk factor for cancer and the leading cause of preventable death in the U.S. (Mokdad et al., 2004; Eheman et al., 2012; Nichols et al., 2012).

Among children and adolescents, obesity has more than doubled since the 1970s (Hedley et al., 2004; Ogden et al., 2008). The mean body mass index of U.S. children and adolescents is increasing at a rate that is much too fast to be explained by a genetic change in the population and, therefore, is likely to be related to environmental factors (Hill and Peters, 1998). There is increasing evidence that social and physical environment influence obesity (Booth et al., 2005; Davison and Lawson, 2006; Frank et al., 2007; Dunton et al., 2009; Morland and Evenson, 2009).

Two reviews indicate that natural environment (e.g. green space) may support physical activity levels among children but the evidence is sparse or conflicting (Davison and Lawson, 2006; Dunton et al., 2009). Other physical environment components that have been identified as potential modifiable determinants of physical activity or obesity for children and adolescents include traffic and air pollution (Gauderman et al., 2007; Jerrett et al., 2010, 2013), personal safety (Committee on Environmental Health), and pedestrian facilities and traffic safety

* Correspondence to: Department of Environmental Health Sciences, Center for Occupational and Environmental Health, Fielding School of Public Health, University of California, Los Angeles, 650 Charles E. Young Drive S, Rm. 56-070 CHS, Mail Code: 177220, Los Angeles, CA 90095, United States
E-mail address: mjerrett@ucla.edu (M. Jerrett).

https://doi.org/10.1016/j.envres.2018.04.022
Received 16 January 2018; Received in revised form 19 April 2018; Accepted 19 April 2018
0013-9351/

(Davison and Lawson, 2006; Timperio et al., 2005). The food environment may also affect diet. Research on the food environment around the home (Galvez et al., 2009; Powell et al., 2007) and school (Jerrett et al., 2009) supports this notion. The within school environment can also support childhood nutrition (Story et al., 2009; Wang et al., 2010).

Public schools are good places to address weight-related behaviors because it is a global way to reach the youth population and children spend a substantial proportion of their time at school (Kriemler et al., 2011). Some school interventions have demonstrated increased physical activity, improved diet, and a decline in obesity (Veugelers and Fitzgerald, 2005; Zenzen and Kridli, 2009). The length of follow-up for these studies is insufficient to determine the long-term effects of the interventions (Zenzen and Kridli, 2009) and the impacts beyond the school environment are not always assessed (Kriemler et al., 2011). Nevertheless, the extant data support the notion that school interventions can be effective (Zenzen and Kridli, 2009; Foster et al., 2008; Evenson et al., 2009; Sharma, 2007; Summerbell et al., 2005; Probart et al., 2007). These interventions are especially effective when both education and environmental changes are included (Kriemler et al., 2011).

Existing efforts to address obesity have included the food and physical activity environment and school-based interventions for youth. Surveillance systems can be used to identify at-risk populations and to evaluate these obesity prevention efforts, including social factors; however, most systems in the U.S. collect a limited number of measures and most do not include environmental factors (Hoelscher et al., 2017). While behavioral surveys are also important for surveillance, sometimes self-reported weight and height can result in an underestimation of obesity prevalence (Hoelscher et al., 2017). School surveillance systems present an opportunity to monitor obesity as weight and height measurements can be incorporated into other routine screenings and because it is a global way to reach the youth population (Hoelscher et al., 2017). It can also be important to track characteristics at the system level that are known to determine obesity and that can be modified to reduce the risk of obesity. The objective of this paper is to examine the relationship between childhood obesity using 2003 through 2007 data from the California Physical Fitness Test (administered yearly to 5th, 7th and 9th graders) and numerous social and environmental characteristics of the within school space, the school neighborhood, and school county. This is the first state-wide study to look at public schools as potential targets for environmental public health tracking to identify at-risk populations to address childhood obesity.

2. Material and methods

2.1. Study population

By law, the State of California requires the yearly physical fitness testing of all 5th, 7th, and 9th graders in the state's public schools and maintains a database of the results through the California Department of Education. This yearly test is called the Physical Fitness Test (PFT) and is standardized through an official battery of tests called the FITNESSGRAM®. The data are considered repeated cross-sections as data from the same student are not identifiable over time due to student ID suppression in the data. Due to population size, two smaller counties did not provide data (Alpine and Trinity). Overall a majority of eligible students participate (e.g. 75% and 79% in 2003 and 2007 respectively). Participants of the PFT 2003 through 2007 constitute the study population for this paper. Data were excluded if relevant information were missing (n = 653,500); schools had fewer than 10 students (n = 2584); and weight values outside the biologically plausible range or if body mass index (BMI) was an outlier as determined from z-scores (n = 114,302). The final analytical dataset included schools with no missing data; thus, the analytic data set includes 5,265,265 student-level observations in more than 6,000 schools. All analyses of students were conducted on de-identified publicly available data. The study procedures were carried out in accordance with the Declaration of Helsinki.

2.2. Measures

Height and weight measurements from the FITNESSGRAM® were used to compute obesity rates for California public schools. Schools were allowed to choose between three methods: caliper test, bioelectric impedance, and BMI (kg/m2). BMI was the most common method (85%). BMI was recalculated from the height and weight data. Data were cleaned to exclude missing and outlier data. Childhood obesity was defined for children and adolescents based on the 2000 U.S. Centers for Disease Control and Prevention (CDC) Growth Charts comparing the BMI-for-age percentile ranking. The ranking compares the child's BMI to the distribution of BMI scores of other children of the same age and gender. Obesity is defined as being at or above the 95th percentile (CDC, 2017).

The independent variables assumed to influence these caloric consumption or energy expenditure can be divided into four hierarchies: Levels 1 (individual), 2 (school), 3 (school neighborhood or census tract), and 4 (school county) variables. The school and neighborhood change may be analyzed longitudinally using the PFT measures aggregated yearly at these ecologic scales. Additionally, other school-level attributes that are measured yearly and where available, year specific data were used. ZIP code level variables were considered but they all dropped out as part of the hierarchical screening process described below.

The only time-dependent variables available for analyses exist at the school (level 2) and neighborhood level (level 3; e.g. SES and other characteristics that can be determined from the Census), as it was not possible to obtain individual longitudinal data from the California Department of Education. The school and neighborhood change may be analyzed longitudinally using the PFT measures aggregated yearly at these ecologic scales. These are available since 2003. Additionally, there are other school-level attributes that are measured yearly and are applicable to the study such as the Academic Performance Index (API) score. Where available, year specific data were used.

2.3. Statistical analysis

Approximately 300 variables were considered. Variables were excluded in a hierarchical process. Step 1 involved assessing multicollinearity by identifying all covariates with absolute value of correlation greater than or equal to 0.8. We then compared standardized coefficients representing the association with the outcome (obesity) using univariate regression and removed the highly correlated variables that had the smaller association with the outcome. A handful of variables of potential policy interest were kept for further analysis even if they were correlated (e.g. the Academic Performance Index base score and percent of students in the Free or Reduced Meals Program), yielding a dataset with 124 variables. Step 2 involved making a decision about the remaining correlated variable pairs ($|\text{correlation}| > 0.7$-$< 0.8$). In this step, the relative importance was determined by ensemble machine learning Random Forest regression, yielding new datasets of 64, 67 and 69 variables for 5th, 7th and 9th grades, respectively. Step 3 involved a final screening of variables by eliminating variables that were captured as components of general variables (for example, murder, rape, and robbery were types of violent crime) and were not eliminated during the correlated variable screening process. Additionally, the county and ZIP code level variables were deemed relatively unimportant based on their relative low Random Forest ranking, which was supported by the evaluation of the variance explained by county-level variables. The final data sets consisted of 48 variables for 5th and 7th grade and 49 variables for 9th grade.