Effects of Built Environment on Obesity: Using a Propensity Score Approach to Assess Selection and Causal Influences

Lori Kowaleski-Jones
Barbara B. Brown
Jessie X. Fan
Ken R. Smith
Ikuho Yamada
Cathleen D. Zick

University of Utah

Introduction

Obesity and overweight are major public health problems. An estimated 65% of US adults are overweight or obese (Hedley et al., 2004) with up to 280,000 annual deaths attributable to obesity (Allison et al., 1999; Flegal et al., 2005). In an effort to understand factors associated with adult obesity/overweight, attention has recently focused on the potential effects of environmental influences. Yet, studies linking the physical environment to the risk of being overweight or obese are limited by the fact that residents are not randomly distributed by neighborhood. If significant associations are found between neighborhood characteristics and individuals' body mass indices (BMI) in observational studies, one cannot confidently draw conclusions about causality. Neighborhood features may cause people to be more physically active, or physically active individuals with low BMI's may be more likely than overweight, sedentary individuals to choose neighborhoods that support their pre-existing healthy lifestyle.

Our prior research (Smith et al, 2008) has identified several features of the built environment to be important in influencing overweight status among adults. Specifically, we conceptualized the walkable-environment measures to include two established predictors—higher density and pedestrian-friendly design (intersections within 0.25 mile of each address)—

and two new census-based, land-use diversity measures: the proportion of residents walking to work and the median age of housing. Linear regressions of BMI and logistic regressions of overweight and obesity include controls for individual-level age and neighborhood-level racial/ethnic composition, median age of residents, and median family income. We found that adding a decade to the average age of neighborhood housing decreases women's risk of obesity by about 8% and men's by 13%. However, this paper did not address the possibility of non random selection into neighborhoods and how that might influence the results. As a result, we adopt a propensity score approach to test for the robustness of the previously reported associations of neighborhood housing with risk of overweight status.

Propensity scores can be used to match participants statistically to equate the effect of covariates that reflect the selection effect. The approach identifies subgroups of individuals who are similar on covariates that may represent the selection effect, but different on a binary outcome of interest, such as obesity. Use of this technique in examining neighborhood effects on pregnancy and dropping out of school has shown robust neighborhood effects. This approach can use a variety of matching procedures to estimate the size of selection effects. The use of propensity scores does not overcome the problem of omitted variables and it depends on having many overlapping cases in the "treatment' and control group.

Data

To assess the potential for non-random selection into neighborhoods, we use data from the Utah Population Database (UPDB) that has been linked to current driver licenses to undertake a propensity score approach to modeling the neighborhood determinants of youth and young adult BMI. From the UPDB, we have 520,973 respondents (age range 16-64), all of whom lived in Salt Lake County in 2007.

The driver license data from the Utah Department of Public Safety were obtained from the Utah Population Database (UPDB), a health-related research database. To protect confidentiality of driver license holders, all personal information from the Driver License Division was removed before the data were provided to the investigators on this research project. This project was approved by the University of Utah IRB and the Utah Resource for Genetic and Epidemiologic Research. As part of this process, the UPDB staff retained identifying address information, linked driver license data (height, weight, gender, and age) to census-block groups via Universal Transverse Mercator (UTM) coordinates, and then provided the researchers with a data set without individual addresses.

Method

Through the use of propensity score matching methods, we will ask the counterfactual question: What would the proportion of high BMI persons be among individuals if they did not receive the "treatment", which in this example would be residing in neighborhoods that did not have older housing stock. The propensity score method will involve several basic steps. First, a logistic regression will be estimated with a set of covariates predicting whether an individual resides in a neighborhood with older homes. Table 1 presents the BMI of men and women in the sample by housing age. From this table, it appears that a logical inflection point to choose for constructing the dichotomous treatment variable will be homes that are over 40 years old. Covariates include individual-level characteristics such as age and gender as well as a set of covariates, measured prior to the outcome, of neighborhood diversity, density, and design. We will also include measures that measures proximity to kin given that potential kin support may be an important predictor of residential location. Second, the predicted probabilities of living in an older neighborhood, as defined by homes over 40 years old, will be calculated. These predicted probabilities are the propensity scores. Third, respondents in older housing stock neighborhoods will be matched to controls based on these propensity scores. Various algorithms for matching exist and have been employed in various investigations. These matching strategies generally produce similar estimates. Given this, we will follow the model of Gibson-Davis (2006) and Harding (2003) of considering the nearest (statistical) neighbor matching method. Fourth, once the data are matched, we will estimate conditional logistic regressions where individual-level overweight/obesity is the dependent variable and the only covariate is a dummy measuring indicating residence in a older neighborhood or not. Propensity scores are used to create the matched strata on which the conditional logistic regressions are based. Finally, we will compare propensity-score-based estimates of the effects of neighborhood housing age on individual overweight/obesity to results from traditional regression models that include all covariates (neighborhood BMI, 3Ds, demographic and familial measures) and that exclude any type of matching.

Preliminary Data and Discussion

Figures 1 and 2 present the distribution of resident body mass index and housing age by quartiles in Salt Lake County. A comparison of the two maps suggests that there is a fair amount of overlap between older neighborhoods and lower BMI, suggesting a relationship. This is also shown in Table 1 which presents the mean BMI of residents by housing age. The lower BMI for both men and women is in the category of older houses, over 40 years old. In our prior work, we suggested that older housing is associated with neighborhoods that mix business with residential land uses and that have more sidewalks as well as more interconnected streets, and thus might promote more physical activity in the community. By using propensity score techniques, we will explore the determinants of living in a neighborhood with older housing and then in turn evaluate the robustness of this linkage with adult overweight. In the discussion of the potential community determinants of adult overweight, information about the potential for non random selection to influence associations will shed light on an important health issue.

References

Allison, D.B., et al., *Annual deaths attributable to obesity in the United States.* Journal of the American Medical Association, 1999. **282**: p. 1530-1538.

Ewing, R., R.C. Brownson, and D. Berrigan, *Relationship between urban sprawl and weight of United States youth.* American Journal of Preventive Medicine, 2006. **31**(6): p. 464-474.

Flegal, K.M., et al., *Excess deaths assocatied with underweight, overweight, and obesity.* Journal of the American Medical Association, 2005. **293**: p. 1861-1867.

Gibson-Davis, Christina and E. Michael Foster. 2006. A Cautionary Tale: Using Propensity Scores to Estimate the Effect of Food Stamps on Food Insecurity. Social Service Review

Harding, D. J. (2003). Counterfactual models of neighborhood effects: The effect of neighborhood poverty on dropping out and teenage pregnancy. *American Journal of Sociology, 109*(3), 676-719.

Hedley, A.A., et al., *Prevalence of overweight and obesity among US children, adolescents, and adults, 1999-2002.* Jama, 2004. **291**(23): p. 2847-50.

Smith K. R., Brown B. B., Yamada I., Kowaleski-Jones L., Zick C. D., Fan J. X., 2008. Walkability and Body Mass Index: Density, design, and new diversity measures. American Journal of Preventive Medicine 35(3):237-244.

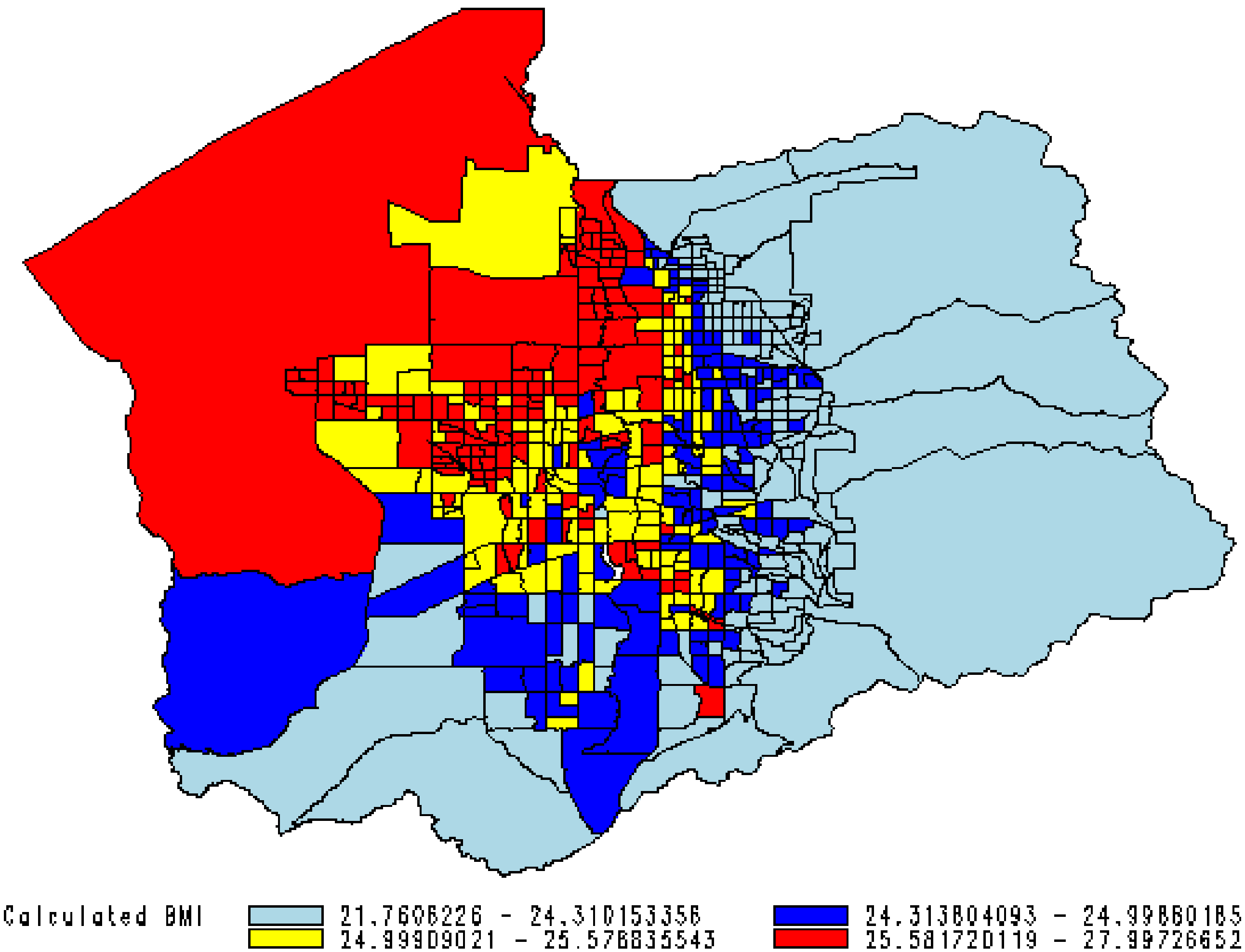
Table 1: Mean BMI by Gender and Housing Age Quartile

	Mean BMI*	
Housing Age	Women	Men
0 - 17	23.91	25.71
18 - 25	24.32	25.92
26 - 40	24.16	25.77
41 - 60	24.14	25.59

*Age Adjusted

Average BMI by Block Group

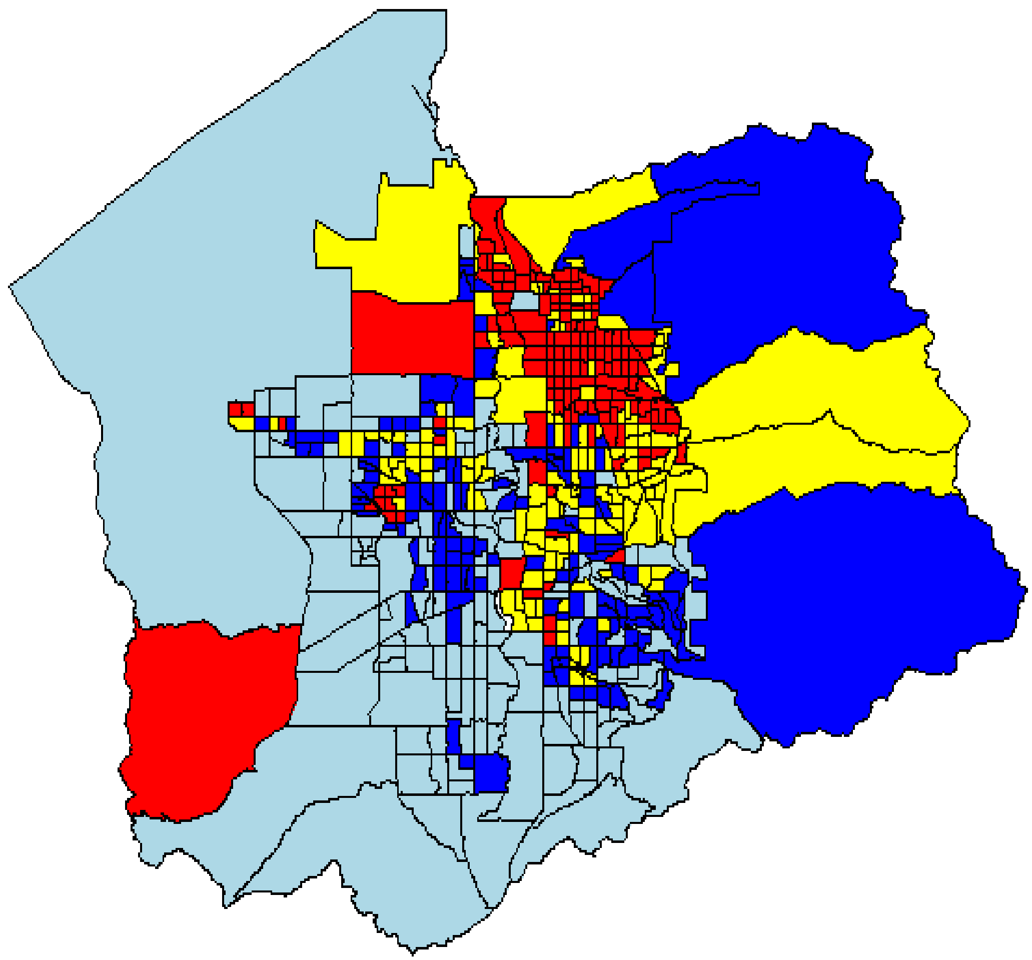
Quartiles



Source: UPDB Driver License Data

Median Housing Age by Block Group: Salt Lake County

Quartiles



Source: 2000 Census data