



Published in final edited form as:

Soc Sci Med. 2007 December ; 65(12): 2458–2463. doi:10.1016/j.socscimed.2007.07.001.

Disparities in obesity rates: Analysis by ZIP code area

Adam Drewnowski,

University of Washington Seattle, WA UNITED STATES [Proxy], adamdrew@u.washington.edu

Colin D Rehm, BA, and

University of Washington, crehm@u.washington.edu

David Solet, PhD

Public Health - Seattle & King County, David.Solet@metrokc.gov

Abstract

Obesity in the USA has been linked to individual income and education. Less is known about its geographic distribution. The goal of this study was to determine whether obesity rates in King County, Seattle, Washington state, at the ZIP code scale were associated with area-based measures of socioeconomic status and wealth. Data from the Behavioral Risk Factor Surveillance System were analyzed. At the ZIP code scale, crude obesity rates varied six-fold. In a model adjusting for covariates and spatial dependence, property values were the strongest predictor of the area-based smoothed obesity prevalence. Geocoding of health data provides new insights into the nature of social determinants of health. Disparities in obesity rates by ZIP code area were greater than disparities associated with individual income or race/ethnicity.

Keywords

obesity; disparities; socioeconomic status; wealth; USA; Zip Code area

Introduction

Obesity rates in the U.S. are more prevalent among minority and other disadvantaged groups (Schoenborn, Adams & Barnes, 2002). However, the observed socioeconomic gradient has not been very steep (Truong & Sturm, 2005) and, other than for white women, not always readily apparent (Mokdad, Ford, Bowman, Dietz, Vinicor, Bales et al. 2003). The impact of socioeconomic factors on obesity rates remains a topic for debate (Flegal, Carroll, Ogden & Johnson, 2002).

Geographic mapping of obesity rates at a sufficiently fine geographic scale may offer new insights into the social determinants of health (Marmot, 2000; Reidpath, Burns, Garrard, Mahoney & Townsend, 2002). However, with some exceptions there are virtually no obesity data that are both geocoded and based on small-area studies (Mobley, Finkelstein, Khavjou & Will, 2004; Morland, Diez Roux & Wing, 2006). Further, the current measures of education and income used to measure socioeconomic status (SES) in the United States may not adequately reflect the impact of race and social class on health parameters (Braveman, Cubbin,

Correspondence to: Adam Drewnowski.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Egerter, Chideya, Marchi, Metzler et al. 2005; Marmot, 2000). Here too, area-based measures of SES may provide alternative ways of assessing economic resources of study participants and of their communities. Neighborhood residential property values may be a more accurate measure of socioeconomic position than either education or income (Braveman et al., 2005). This case study aimed to determine whether obesity rates in King County, WA, at the ZIP code scale were linked to area-based measures of race/ethnicity, income, poverty, and property values.

Methods

King County is a large metropolitan county in Washington State with a population of more than 1.7 million residents. The population and demographics of King County are described elsewhere (King County GIS Center, 2005). Aggregated multiple-year (1999–2003) data from the Washington State Behavioral Risk Factor Surveillance System (BRFSS) for King Co, with an oversample of King Co. respondents, were used for analyses. The BRFSS, conducted by the Centers for Disease Control and Prevention (CDC) and local and state health departments in all 50 states, is a random-digit-dial telephone survey of adults ≥ 18 years. Body mass index (BMI) was calculated from self-reported height and weight; respondents with a BMI of ≥ 30 were classified as obese. Aggregate data for 8,803 individuals from multiple years were used to estimate the prevalence of obesity for each ZIP code area. Given the small sample size of some ZIP code areas, an Empirical Bayes tool was used to estimate the smoothed obesity prevalence. This method allowed us to include ZIP code areas with small sample sizes that would otherwise have been excluded. This method tends to smooth prevalence estimates, especially those based on small sample size or with extreme values, in the direction of the countywide prevalence.

The final analysis was based on 74 ZIP code areas within King County. Forty-three ZIP code areas ($n=93$), were excluded from the analysis due to insufficiently small sample size within these ZIP code areas ($n<10$). A majority of the ZIP code areas excluded from the analysis had only one observation. A total of 8,803 individuals were included in the final analysis. The number of individuals surveyed in each ZIP code included in analysis ranged from 13 to 275. ZIP code areas used in the analysis had an average population size of 24,543 with a population range of 3,822–64,214. More than 95% of the King County adult population reside in the ZIP code areas included in the present study.

Binomial confidence intervals calculated by the score method were used to evaluate the statistical stability of the estimated crude obesity rates. Smoothed obesity rates were mapped using boundary files provided by the U.S. Census. ZIP code tabulation areas (ZCTA), a Census feature that approximates the geographic boundaries for ZIP code areas, was used for mapping and collecting contextual measures. The ZCTA boundary file was current for 2000.

Data from the 2000 U.S. Census included percent of adults living below the federal poverty line, median household income, median value of all owner-occupied housing units, percent African-American, and percent of Hispanic origin. The federal poverty line in the United States is based on household income and size. In 2000, the single-family household poverty threshold was \$8,959. Contextual data were used to evaluate the bivariate association between area-based obesity prevalence and ZIP code area contextual characteristics. The smoothed obesity prevalence followed a normal distribution, justifying the use of linear regression models as opposed to Poisson models. The bivariate relationship between each predictor and smoothed obesity prevalence was assessed by estimating linear regression models (using a two-tailed test, $\alpha = 0.05$). Multi-collinearity between predictors of interest was a concern. The tolerance for each bivariate covariate combination was assessed in linear regression models ($1-r^2$). In no instance did the tolerance indicate serious collinearity problems. The smallest tolerance

between predictors was between percent of adults below poverty and median household income ($r^2=0.70$; tolerance=0.30), however percent of adults living below the federal poverty line was not a significant independent predictor of obesity prevalence, and was not included in the final models.

Significant parameters from bivariate analysis were included in an exploratory ordinary least squares multivariate linear regression model. Any analysis of spatial data is complicated by potential spatial autocorrelation or dependence, where contiguous or nearby areas have correlated response or predictor values. The Moran's I statistic, a measure of spatial autocorrelation, was calculated to evaluate spatial autocorrelation for the smoothed obesity prevalence. A queen weights contiguity matrix was used for all spatial analyses. The queen weights define neighbors as ZIP code areas that share a boundary or a corner. One assumption of ordinary linear regression is independence of predictors, so the identification of significant spatial autocorrelation in our data necessitated the use of a spatial linear regression model, using a spatial error term, which corrects for spatial dependence (Anselin & Bera, 1998).

Statistical and spatial analyses were performed in Intercooled Stata 9.2 for Windows (StataCorp College Station, TX) and GeoDa (Luc Anselin and The Regents of the University of Illinois). Mapping was done in ArcGIS 9.1 (ESRI Redlands, CA).

Results

The Moran's statistic indicated strong spatial autocorrelation for the smoothed obesity prevalence by ZIP code area (Moran's I = 0.449; $p < 0.001$). Figure 1 shows the crude obesity prevalence and 95% confidence intervals for the 74 ZIP code areas included in the study. Figure 2 shows substantial differences between the lowest (10.1%) and the highest (25.2%) smoothed obesity rates by ZIP code area. Significant bivariate predictors of the smoothed obesity prevalence were: median value for owner-occupied housing units ($r^2 = 0.46$), and percent of population that was Hispanic ($r^2 = 0.19$). Median household income ($r^2 = 0.05$; $p = 0.058$) was not significant at the 0.05 level, but was included in subsequent models. Results of bivariate correlation analyses are shown in Table 1.

In both crude and spatial error linear regression models, median house value was significantly associated with the smoothed obesity prevalence. In a spatial error model adjusting for spatial dependence, median household income and percent Hispanic, median house value was a significant predictor of obesity prevalence at the ZIP code level ($p < 0.001$; model $r^2 = 0.52$). The inverse relationship between smoothed obesity rates by ZIP code and median house values is shown in Figure 2. For each additional \$100,000 in median house values, ZIP code area obesity prevalence decreased by 2.0% (95% CI 0.9% – 3.1%; $p < 0.001$) when adjusting for covariates and spatial dependence.

Discussion

This is one of the first studies to map the spatial distribution of obesity rates at a fine geographic scale that is more detailed than the state (Holtgrave & Crosby, 2006; Mokdad, Bowman, Ford, Vinicor, Marks & Koplan, 2001; Mokdad et al., 2003), Metropolitan Statistical Area (Ford, Mokdad, Giles, Galuska & Serdula, 2005) or county level (Ewing, Schmid, Killingsworth, Zlot & Raudenbush, 2003). The New York City Community Health Survey, mapped the prevalence of obesity at the United Hospital Fund neighborhood level, an aggregate of respondents' ZIP codes (New York City Department of Health and Mental Hygiene, 2006).

The observed disparities in obesity rates across King County ZIP code areas suggest a strong relationship between obesity rates and some area-based indices of SES. Those geographic disparities were much higher than those traditionally ascribed to race or ethnicity. Analyses of

local 2000–2005 BRFSS data indicate that obesity rates among African-Americans in King County (26.5%) were 63% higher than among whites (16.3%) (Public Health - Seattle & King County, 2005). Obesity rates among persons with incomes <\$15,000 were 20.0% as compared to 15.1% among persons with incomes >\$50,000, a 32% difference (Public Health - Seattle & King County, 2005). At the national level, Hispanics are 1.2 times more likely to be obese than are whites (23.7% compared to 19.6%) (Mokdad et al., 2001). Nearly three-fold differences in obesity rates, based on geographic location, have not been reported in the literature.

Traditionally, race/ethnicity as well as education and incomes have been the focus of health disparities research (Isaacs & Schroeder, 2004). Given recent concerns that individual education and incomes may not adequately reflect social class (Marmot, 2000), there is a growing emphasis on the roles of occupation, social capital, and social context, all difficult parameters to capture in epidemiologic studies. Area-based measures of SES can provide additional information on poverty and wealth that is only rarely collected in United States health surveys (Krieger, Chen, Waterman, Rehkopf & Subramanian, 2003). In past studies, percent of residents living below the federal poverty level, based on census data, was identified as the best predictor of health outcomes (Krieger et al., 2003); whereas the present analyses point to median house values, a shorthand measure of wealth, as a strong predictor of obesity rates. The present findings are thus wholly consistent with past research on social determinants of health (Marmot, 2000).

The present study is subject to some serious limitations. First, heights and weights, used to calculate BMI values, were based on telephone self-report. Both men and women under-report weight, and men may over-report height in telephone surveys (Nawaz, Chan, Abdulrahman, Larson & Katz, 2001). However, the same BRFSS data are said to provide the best picture of the obesity epidemic in the U.S. and are the basis for national policy decisions (Mokdad et al., 2001; Mokdad et al., 2003). Second, our exclusive use of aggregate measures of SES and health outcomes does not allow us to generalize the effect of SES on obesity risk among individuals.

In addition, the standard CDC cut-points for obesity may not be appropriate for 18–19 year olds. We did not adjust BMIs for 18–19 year olds because approximately 1.5% of the sample was 18–19 year olds (n=140), and failure to adjust BMIs in the present study would not appreciably bias the results. While the current study evaluated the prevalence of obesity by area, an alternative approach would map mean BMI per area. Such an analysis would reduce potential bias due to misclassification of individuals as obese or not obese; however, data on mean BMI per ZIP code area was not available.

Perhaps more important is the issue that the ZIP code area is a problematic scale for spatial analysis. Because population counts per ZIP code area can vary widely, many ZIP code areas were too small to provide area-based prevalence estimates despite the use of an Empirical Bayes tool. Even for ZIP code areas with a large population, confidence intervals can be quite large as shown in Figure 1. We were unable to analyze the prevalence of obesity at the census tract level because respondents were only asked to report their ZIP code. Another challenge associated with the use of ZIP code areas is that they are designed to efficiently deliver mail, and the boundaries change subtly on a regular basis (Krieger, Waterman, Chen, Soobader, Subramanian & Carson, 2002).

Finally, obesity rates per ZIP code area were calculated with unweighted responses. Although the CDC provides weights for use in county- and state-level calculations, weights created for larger geographic areas are unlikely to adjust for non-response and differential probability of selection at the ZIP code level. Using unweighted data greatly simplified our analyses, since survey design effect for the BRFSS prevalence rates did not have to be taken into account when calculating the significance level of correlations and regression models. Potential sources of

bias in state-specific BRFSS data include low response rates and non-response biases within certain demographic groups. Such biases may also occur at the neighborhood level, which might make the present population samples non representative. At state level, BRFSS demographics have been compared to other data sources to determine potential sources of response bias. Fewer external data sources are available in small area studies. This is an important caution, especially since the BRFSS design weights were not devised with small area studies in mind or used in the present analysis.

The present disparities at ZIP code area level stand in contrast to the well-known CDC maps, where the differences in obesity rates between the richer and the poorer states are only weakly apparent (Mokdad et al., 2001). Whereas the CDC maps have been used to support the argument that obesity rates in the U.S. are unrelated to social class, the present data show – to the contrary – that the obesity problem is concentrated in the most disadvantaged areas.

Studying the geography of obesity will require new maps of the finest spatial and statistical precision. Maps of obesity at finer spatial scales, such as the census tract or ZIP code area scale are preferable to maps at the county or state level. Mapping disease rates by community and neighborhood may very well be the future of public health assessment and surveillance.

References

- Anselin, L.; Bera, AK. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, A.; Giles, DE., editors. Handbook of Applied Economic Statistics. New York: CRC Press; 1998. p. 237-290.
- Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, Posner S. Socioeconomic status in health research: one size does not fit all. *JAMA* 2005;294(22):2879–2888. [PubMed: 16352796]
- Ewing R, Schmid T, Killingsworth R, Zlot A, Raudenbush S. Relationship between urban sprawl and physical activity, obesity, and morbidity. *Am J Health Promot* 2003;18(1):47–57. [PubMed: 13677962]
- Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999–2000. *JAMA* 2002;288(14):1723–1727. [PubMed: 12365955]
- Ford ES, Mokdad AH, Giles WH, Galuska DA, Serdula MK. Geographic variation in the prevalence of obesity, diabetes, and obesity-related behaviors. *Obes Res* 2005;13(1):118–122. [PubMed: 15761170]
- Holtgrave DR, Crosby R. Is social capital a protective factor against obesity and diabetes? Findings from an exploratory study. *Ann Epidemiol* 2006;16(5):406–408. [PubMed: 16246583]
- Isaacs SL, Schroeder SA. Class - the ignored determinant of the nation's health. *N Engl J Med* 2004;351(11):1137–1142. [PubMed: 15356313]
- King County GIS Center. King County Census Viewer. 2000 [5-4-2005]. US Census <http://www5.metrokc.gov/KCCensus/>
- Krieger N, Chen JT, Waterman PD, Rehkopf DH, Subramanian SV. Race/ethnicity, gender, and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures--the public health disparities geocoding project. *Am J Public Health* 2003;93(10):1655–1671. [PubMed: 14534218]
- Krieger N, Waterman P, Chen JT, Soobader MJ, Subramanian SV, Carson R. Zip code caveat: bias due to spatiotemporal mismatches between zip codes and US census-defined geographic areas--the Public Health Disparities Geocoding Project. *Am J Public Health* 2002;92(7):1100–1102. [PubMed: 12084688]
- Marmot M. Social determinants of health: from observation to policy. *Med J Aust* 2000;172(8):379–382. [PubMed: 10840490]
- Mobley LR, Finkelstein EA, Khavjou OA, Will JC. Spatial analysis of body mass index and smoking behavior among WISEWOMAN participants. *J Womens Health (Larchmt)* 2004;13(5):519–528. [PubMed: 15266669]
- Mokdad AH, Bowman BA, Ford ES, Vinicor F, Marks JS, Koplan JP. The continuing epidemics of obesity and diabetes in the United States. *JAMA* 2001;286(10):1195–1200. [PubMed: 11559264]

- Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, Marks JS. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA* 2003;289(1):76–79. [PubMed: 12503980]
- Morland K, Diez Roux AV, Wing S. Supermarkets, other food stores, and obesity the atherosclerosis risk in communities study. *Am J Prev Med* 2006;30(4):333–339. [PubMed: 16530621]
- Nawaz H, Chan W, Abdulrahman M, Larson D, Katz DL. Self-reported weight and height: implications for obesity research. *Am J Prev Med* 2001;20(4):294–298. [PubMed: 11331120]
- New York City Department of Health and Mental Hygiene. New York City Community Health Survey. 2006 [9-22-2006]. <http://www.nyc.gov/html/doh/html/survey/survey.shtml>
- Public Health - Seattle & King County. King County Core Indicators. Public Health - Seattle & King County. 2005. <http://www.metrokc.gov/health/reports/coreindicators/>
- Reidpath DD, Burns C, Garrard J, Mahoney M, Townsend M. An ecological study of the relationship between social and environmental determinants of obesity. *Health Place* 2002;8(2):141–145. [PubMed: 11943585]
- Schoenborn CA, Adams PF, Barnes PM. Body weight status of adults: United States, 1997–98. *Adv Data* 2002;330:1–15. [PubMed: 12661975]
- Truong KD, Sturm R. Weight gain trends across sociodemographic groups in the United States. *Am J Public Health* 2005;95(9):1602–1606. [PubMed: 16051939]

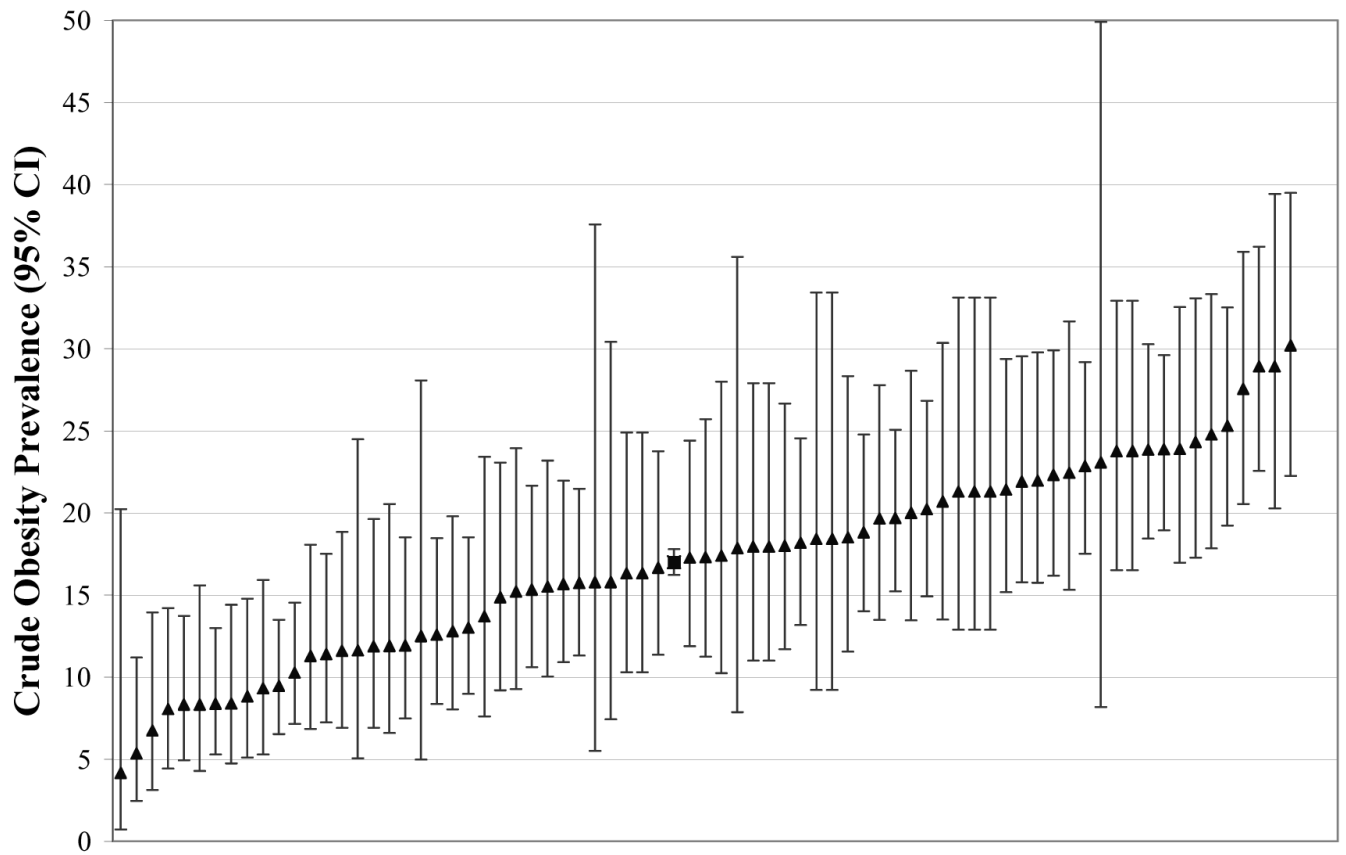


Figure 1.
Crude obesity prevalence and 95% confidence intervals for ZIP code areas in King County

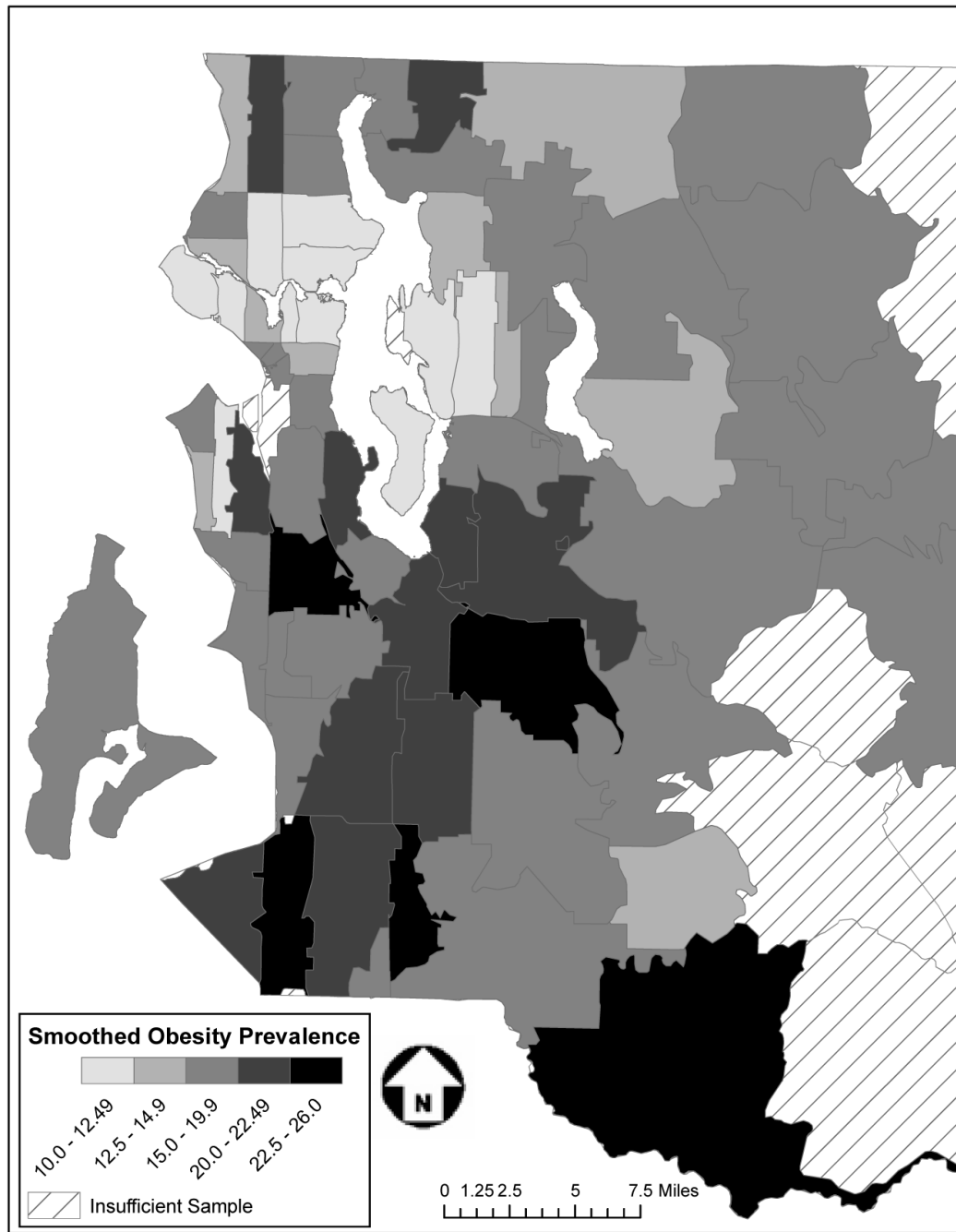


Figure 2.
Smoothed obesity prevalence by ZIP code area in King County

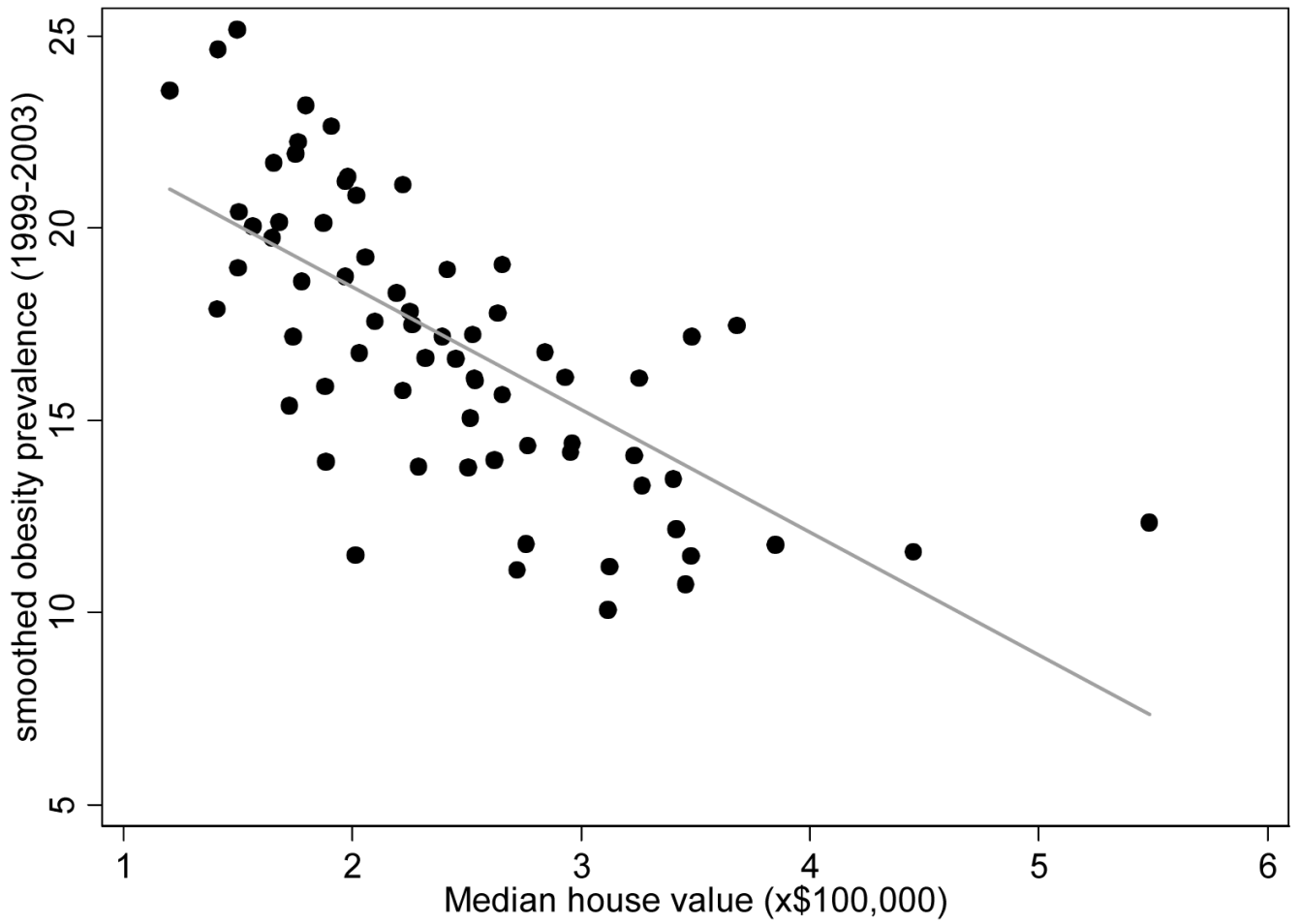


Figure 3. Smoothed obesity prevalence by King County ZIP code area plotted against median house values

Table 1

Bivariate predictors of smoothed obesity prevalence by ZIP code area

Predictor	r^2	p-value
Median House Value	0.46	<0.001
Median Household Income	0.05	0.058
% of Adults Living Below Poverty Line ¹	0.02	0.26
Percent African-American	0.02	0.18
Percent Hispanic	0.19	<0.001

¹Federal poverty line in the United States is based on household income and size. In 2000, the single-family household poverty threshold was \$8,959.