Spatial and Spatio-temporal Epidemiology 21 (2017) 13–23

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

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

Original Research

Exploring racial disparity in obesity: A mediation analysis considering geo-coded environmental factors

Qingzhao Yu[a,*], Richard A. Scribner[b], Claudia Leonardi[b], Lu Zhang[c], Chi Park[b], Liwei Chen[c], Neal R. Simonsen[d]

[a] Biostatistics Program, School of Public Health, Louisiana State University Health Sciences Center, 3rd floor, 2020 Gravier Street, New Orleans, LA 70112, United States
[b] Epidemiology Program, School of Public Health, Louisiana State University Health Sciences Center, New Orleans, LA, United States
[c] Department of Public Health Sciences, Clemson University, Clemson, SC, United States
[d] New Orleans, LA, United States

ARTICLE INFO

Article history:
Received 5 December 2016
Revised 31 January 2017
Accepted 10 February 2017
Available online 17 February 2017

Keywords:
Mediation analysis
Obesity
Physical activity
Racial disparity

ABSTRACT

Research shows a consistent racial disparity in obesity between white and black adults in the United States. Accounting for the disparity is a challenge given the variety of the contributing factors, the nature of the association, and the multilevel relationships among the factors. We used the multivariable mediation analysis (MMA) method to explore the racial disparity in obesity considering not only the individual behavior but also geospatially derived environmental risk factors. Results from generalized linear models (GLM) were compared with those from multiple additive regression trees (MART) which allow for hierarchical data structure, and fitting of nonlinear and complex interactive relationships. As results, both individual and geographically defined factors contributed to the racial disparity in obesity. MART performed better than GLM models in that MART explained a larger proportion of the racial disparity in obesity. However, there remained disparities that cannot be explained by factors collected in this study.

1. Introduction

Obesity is a serious public health concern in the United States (US). More than one third of US adults have obesity. It is well documented that black people have a higher rate of obesity than white people, which results in the racial disparity in obesity and the related diseases. Explanations for the obesity epidemic and the respective racial disparity are multifactorial. In addition to the traditional individual behavior risk factors, geo-coded risk factors at the neighborhood level have also been proposed. Aspects of the built environment have been considered as modifiable risk factors that influence both energy expenditure and energy consumption and therefore can be modified to address the obesity epidemic as well as the disparity in obesity between blacks and whites (Hill et al., 2003; Khan et al., 2009; Papas et al., 2007; Roux, 2003). Efforts to

Abbreviations: US, United States; NHANES, National Health and Nutrition Examination Survey; GIS, Geographic Information System; MART, multivariate additive regression trees; NCHS, National Center for Health Statistics; CDC, the Centers for Disease Control and Prevention; MEC, Medical Examination Component; NCAIS, North American Industry Classification System; SIC, Standard Industrial Classification; ESRI, Environmental Systems Research Institute; GLM, Generalized Linear Model; BMI, body mass index; kg, kilograms; m^2, squared meters; ANOVA, analysis of variance; GLM, generalized linear models; CDI, Concentrated Disadvantage Index.

* Corresponding author.
E-mail addresses: qyu@lsuhsc.edu (Q. Yu), rscrib@lsuhsc.edu (R.A. Scribner), cleon1@lsuhsc.edu (C. Leonardi), lz3@clemson.edu (L. Zhang), liz0817@gmail.com (C. Park), liweic@clemson.edu (L. Chen), epiman@yahoo.com (N.R. Simonsen).

http://dx.doi.org/10.1016/j.sste.2017.02.001
1877-5845/

Table 1
Measures of street connectivity used in the analyses.

Measure (notation)	Definition	Calculation*
Intersection density	Number of intersections per square mile of area	# Real nodes/area
Street density	Number of linear miles of street per square mile of area	Total # roadway miles/area
Connected node ratio	Ratio of number of intersections with four or more connections over the total number intersections	# of intersections associated with four or more links/total # of intersections

* Calculations utilize geographic information system-derived data.

understand the relative contribution of any single factor often result in equivocal findings. As a result, more sophisticated methods capable of addressing the variety and hierarchical structures are needed. Specifically, methods accounting for the contributions of both individual and environmental factors as well as the complex relationships (e.g., nonlinear) potentially involved. There has been considerable interest in creating spatially defined risk factors, and quantifying these risk factors at the neighborhood level in order to identify modifiable factors in the environment that account for obesity in general and the disparity in particular (Ding and Gebel, 2012; Gebel et al., 2007; Grasser et al., 2013).

Walkability represents one of the factors that has a complex relationship with both obesity and the racial disparity in obesity rates among Americans. We created the variables: intersection density, street density, and connection node ratio (defined in Table 1, geospatial definitions are provided in the Method section) to measure the overall construct of neighborhood walkability. While the overall neighborhood walkability construct is widely considered as a factor associated with increased physical activity and lower rates of obesity, the research linking measures of the sub-construct street connectivity to obesity in general (Ball et al., 2012; Heinrich et al., 2008; Li et al., 2008; McDonald et al., 2012; Wang et al., 2013; Wen and Kowaleski-Jones, 2012), and physical activity in particular (Berrigan and Troiano, 2002; Eriksson et al., 2012; Frank et al., 2008; Li et al., 2005; Oakes et al., 2007; Owen et al., 2007; Pearce and Maddison, 2011; Saelens et al., 2003; Witten et al., 2012) has been equivocal, evidenced in the conclusions of some reviews on the topic (Grasser et al., 2013; Saelens and Handy, 2008). Street connectivity (e.g. intersection density) appears to be associated with both obesity and physical activity. However, when the effect of other measures of walkability (e.g., street density) are controlled, the association is attenuated. It is not clear whether the effect of street connectivity is explained by these other constructs or whether other statistical issues are involved. For example, the effect of street connectivity may have a nonlinear or even non-monotone relationship with the prevalence of obesity, leading to the equivocal results. Alternatively, the effect of street connectivity may be explained in terms of its' strong association with street density, which is actually driving the effect and apparent street connectivity effects arise from this multicollinearity. In any case, the relationship is complex and hierarchical. It poses serious problems for jointly using individual and geo-spatial factors in explanation in terms of existing analytic designs.

With regard to the disparity in obesity between whites and blacks, street connectivity also has a complex relationship. While blacks in general are more obese than whites, they are more likely to reside in urban centers. However, urban centers tend to be areas with high street connectivity suggesting that blacks living in these areas should be less obese than their white counterparts in low street connectivity areas (e.g., suburbs). In addition, King (2013) recently concluded that the increased density of urban centers may overcome the benefits of walking. Once again the relationship is a complex one involving a variety of factors that are difficult to characterize with existing analytic models.

Mediation effect refers to the effect conveyed by an intervening variable to an observed relationship between an exposure and a response variable of interest. To explore the racial disparity in obesity, mediation analysis is used where race is the exposure variable and obesity or its absence is the binary response. All other risk factors that may explain the racial disparity are considered as potential mediators. There are several key challenges in the exploration of racial disparity in obesity. First, the model should be able to deal with different types of mediators, where the potential factors can be continuous, binary or categorical with or without order. Second, the indirect effect from each mediator should be differentiable so the indirect effect conveyed by different factors can be compared. Third, given multiple measures of walkability with influences that might vary by measure, their mediating effect on the racial disparity needs to be assessed jointly rather than additively, so we need a method that can measure the joint effects from combined factors. Finally, there are potential nonlinear relationships and interactions among race, the mediators, and the obesity risk, therefore the fitted model should be able to represent the most reasonable adequately complex one that represents the relationships existing among variables.

There are generally two settings for mediation analysis. One is based on linear models (Baron and Kenny, 1986; MacKinnon et al., 1995), and the other is based on the counterfactual framework (Albert, 2008; Have et al., 2007; Pearl, 2001; Robins and Greenland, 1992). In this paper, we adapted a general definition of mediation effect by Yu et al. (2014) based on the counterfactual framework. The derived mediation analysis proposed by Yu et al. is promising in that the indirect effects contributed by different mediators are separable, which enables the comparison of relative mediation effects carried by different third variables. The mediation analysis is generalized so that we can deal with binary, multi-categorical or continuous exposure, mediator and response variables. Moreover, general predictive models, as well as general linear models can be used