Applied Geography 55 (2014) 117–126

Contents lists available at ScienceDirect

Applied Geography

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

Geospatial and geodemographic insights for diabetes in the United States

Tony H. Grubesic [a,*], Jennifer A. Miller [b], Alan T. Murray [a]

[a] *Center for Spatial Analytics and Geocomputation, College of Computing & Informatics, Drexel University, Philadelphia, PA 19104, USA*
[b] *Department of Geography and the Environment, University of Texas, Austin, TX 78712, USA*

ARTICLE INFO

Article history:
Available online

Keywords:
Diabetes
Geography
Spatial analysis
Geodemographics
Lifestyle
Clusters

ABSTRACT

Diabetes is a major public health problem in the United States. Over 29 million residents are currently diagnosed with the disease. Further, with such a large percentage of the U.S. population being diabetic, an interesting spatial pattern for the disease has emerged. The purpose of this paper is to evaluate the geodemographic correlates of Type 2 diabetes in the United States. Specifically, using a nationwide database of age-adjusted, county level estimates for diabetes prevalence, we provide an exploratory spatial analysis of lifestyle groups and their connection with diabetes. Results suggest that geodemographic information can be effective in pinpointing risky lifestyle environments and may provide basic guidance for identifying at-risk populations in order to target intervention efforts more effectively.

Introduction

The use of geodemographic data and associated lifestyle segmentation systems in public health research is relatively sparse. Broadly defined, geodemographics is the art and science of analyzing socio-economic and behavioral data about people within the context of place (i.e. geography) and local settlements (Harris, 2003). In part, the limited use of geodemographics in public health research is fueled by a longstanding, albeit waning bias toward population-based studies in epidemiology (Pearce, 2000). Potential for ecological fallacy in population level studies, where inferences about individuals are deduced for a group (Robinson, 1950), is real. Further, with the rise of biostatistical methods and the ability to develop sophisticated cohort and case–control studies that more closely fit a clinical trial paradigm, ecological and population level studies *can* be viewed as unreliable and systematically biased options in comparison to the biostatistical alternatives (Pearce, 2000). However, given the reemergence and growing popularity of ecological studies and the recognition that many risk factors for disease operate at the population level (Susser, 1994a, 1994b), geodemographics and lifestyle segmentation systems have potential to play a larger role in public health research.

Consider, for example, diabetes mellitus (DM). The Type 2 variant is regionally and demographically significant. Many of the modifiable risk factors associated with the disease, such as sedentary lifestyle and diet, are partly determined by social context and the environment where risky behavior takes place. Given the wide variation in lifestyles and local dietary habits in the United States (Popkin, Duffey, & Gordon-Larsen, 2005; Shortridge & Shortridge, 1999), heterogeneities in the spatial distribution of diabetes prevalence are to be expected. Further, many of the non-modifiable risk factors such as age, race, income and education also influence the geographies of diabetes prevalence (Barquera, Tovar-Guzman, Campos-Nonato, Conzalez-Villalpando, & Rivera-Dommarco, 2003; Green, Hoppa, Young, & Blanchard, 2003). Perhaps the most provocative facet of the geographic distribution of diagnosed diabetes in the United States is the presence of a “diabetes belt” in the Southeast (Barker, Kirtland, Gregg, Geiss, and Thompson, 2011). The diabetes belt based on county level estimates of diabetes prevalence for 2007–2008, where 644 counties are identified as “high” prevalence (>=11%) and spatially proximal. This belt includes portions of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Texas, Virginia, West Virginia and all of Mississippi. Again, while many of the risk factors for DM are well known (e.g. age, obesity, sedentary lifestyle, and non-Hispanic African-American ancestry), statistical results from Barker et al. (2011) suggest that everyone in the diabetes belt, including those that had few risk factors, were at a greater risk of diabetes than people outside the belt. Although Diamond (2003) suggests that both genetics and food history may help explain differences in diabetes prevalence, ingrained social and

* Corresponding author.
E-mail address: grubesic@drexel.edu (T.H. Grubesic).

http://dx.doi.org/10.1016/j.apgeog.2014.08.017
0143-6228/

cultural characteristics associated with local lifestyle preferences may also be contributing factors to diabetes risk.

The purpose of this paper is to evaluate geodemographic correlates of Type 2 diabetes in the United States. Specifically, using a nationwide database of age-adjusted, county level estimates for diabetes, an exploratory analysis of lifestyle groups and diabetes prevalence is undertaken. Results suggest that geodemographic systems help pinpoint risky lifestyle environments and provide guidance for targeting health intervention efforts. In the next section, geodemographics and their use in public health research are reviewed, highlighting relative strengths and weaknesses. This is followed by a brief overview of Type 2 diabetes, its known risk factors and geography. The Methods Section details the data and methods used for identifying the geodemographic correlates of diabetes in the United States, results are reported, and the paper concludes with a brief discussion detailing the policy implications for targeted health interventions.

Background

Geodemography is the art and science of analyzing socio-economic and behavioral data about people within the context of place (i.e. geography) and local settlements (Harris, 2003). In short, the ability to identify geographical trends or patterns within societies, such as consumer preferences or likely voting behavior, is an important step in better understanding the factors that fuel spatial outcomes. As detailed by Harris, Sleight, and Webber (2005), the strength of geodemography is generating new ideas and insights about spatial outcomes that can be investigated and debated further. This inductive approach to examining spatial outcomes is often labeled exploratory spatial data analysis, or ESDA (Bailey & Gatrell, 1995; Anselin, 1998; Rogerson & Yamada, 2009; Waller & Gotway, 2004). Thus, geodemographics is an exploratory tool, not necessarily a statistical method used to confirm or reject hypotheses (Harris et al., 2005). The core limitation of geodemographics is also well known. They cannot be used to “explain” outcomes. Disentangling complex socio-economic processes in both space and time is extremely difficult and geodemographics alone cannot account for these complexities.

The use of geodemographics in public health research is relatively sparse, but broader usage is noteworthy. Given this, epidemiological work provides a basis for perspective. First, geodemographics can suffer from a form of ecological fallacy. Although small areas are often labeled as adhering to a particular geodemographic group, behaviors of individuals within an area and the associated group vary. In this context, aggregation bias is a concern. Second, geodemography is primarily rooted in consumer and lifestyle behavior, which does not translate directly to health-related behavior. Third, geodemography emphasizes factors that make areas distinct from their peers. However, factors that make an area distinctive might be rare within an area and any effort to link a health-related behavior to these characteristics may be mismatched with the dominant qualities of individuals in a region. Finally, because geodemographics are rooted in place, they emphasize the local context in which health-related behavior occurs, potentially ignoring the multitude of individual-level factors that impact health.

Although the weaknesses of geodemographic systems are valid (Twigg, Moon, & Jones, 2000), the use of geodemography as a tool for health intelligence (Abbas, Ojo, & Orange, 2009) remains intriguing. There are several reasons for this. First, because geodemographic classifications are based on hundreds of variables, they represent a multifaceted statistical summary of lifestyle choices and preferences that provide a stronger compositional view of an area when compared to univariate snapshots such as age, gender or ethnicity. Second, it is important to remember that health care is a product. Thus, the ability to communicate, market and provide health care products plays an important role in the health care industry and ultimately to consumers. For example, geodemographics provide a means to identify variations in the spatial outcomes of health care efforts as well as to understand the satisfaction of consumer subsets with services provided. Abbas et al. (2009) also argue that geodemographics are excellent tools for measuring inequalities in health at the local level (e.g. link between poor health and socio-economic deprivation) and benchmarking the performance of health intervention efforts among geodemographic groups.

Again, while limited, geodemographics have been used with some success in public health efforts. For example, Petersen et al. (2011) explored the utility of using geodemography as a means for targeting neighborhoods in public health campaigns. Although the results indicated that geodemographic indicators were poor discriminators when compared to geographic targeting strategies based on crude disease rates in London (i.e. targeting areas was better than area types), the authors convincingly argue that the strength of geodemographics is exploration and description – not explanation of particular health outcomes. In a different effort, Farr and Evans (2005) used geodemographics to identify undiagnosed (i.e. unknown) diabetics in Great Britain. Specifically, geodemographic data were used in conjunction with hospital episode statistics for Type 2 diabetes to inform a social marketing campaign for encouraging high risk individuals to come forward for screening. In fact, several geodemographic groups were identified as representing high risk segments of the population within the Slough Primary Care Trust and the authors suggest that these groups could be targeted for intervention efforts in other communities. Sheringham, Sowden, Stafford, Simms, and Raine (2009) used a geodemographic database for monitoring inequalities in the national chlamydia screening program in England, Kimura et al. (2011) used geodemography to identify differences in neighborhood characteristics and the incidence of influenza, and Tickle, Milsom, Jenner, and Blinkhorn (2003) utilized a geodemographic system to examine variations in dental caries in Cheshire, UK.

Regardless of the substantive focus of the analysis, virtually all of the applications of geodemographics to public health questions add both depth and clarity. Again, while geodemography is not used to *explain* health outcomes, its strength in detailing socio-economic, demographic and geographic inequities between populations and disease incidence is extremely valuable in an exploratory context. In particular, the ability of geodemographic systems to deepen our understanding regional ecological structure and help identify modifiable risk factors associated with disease incidence (e.g. sedentary lifestyle and diet) make them a potentially important for epidemiological insights.

Type 2 diabetes mellitus

Type 2 diabetes is broadly defined as impaired glucose tolerance that results from a complex interaction between genetic predispositions for the disease, combined with behavioral and environmental risk factors (Neel, 1962; Tuomilehto et al., 2001; Tuomilehto & Wolf, 1987). Although the genetic factors associated with DM continue to puzzle scientists (Diamond, 2003), the modifiable risk factors associated with the disease, such as obesity, physical inactivity and dietary habits, appear to be the main non-genetic determinants (van Dam, Rimm, Willett, Stampfer, & Hu, 2002; Hu et al., 2001; Ohlson et al., 1988; Tuomilehto et al., 2001). Research suggests that the adoption of a western diet may be associated with increased incidence of Type 2 diabetes (van Dam et al., 2002; Hu et al., 2001; Paynter et al., 2006). The western diet is