Diabetes Prevalence is Associated with Different Community Factors in the Diabetes Belt versus the Rest of the United States

Candice A. Myers[1], **Tim Slack**[2], **Stephanie T. Broyles**[1], **Steven B. Heymsfield**, **Timothy S. Church**[1], and **Corby K. Martin**[1,*]

[1]Pennington Biomedical Research Center, Baton Rouge, LA, 70808

[2]Louisiana State University, Baton Rouge, LA, 70803

Abstract

Objective—This study investigated differences in community characteristics associated with diabetes prevalence between the Diabetes Belt and the rest of the contiguous U.S.

Methods—We used county-level adult diabetes prevalence estimates (i.e., percent of people [20 years] with diagnosed diabetes, 2009) from the Centers for Disease Control and Prevention, in addition to data from the U.S. Census Bureau, U.S. Department of Agriculture, and U.S. Department of Health and Human Services, to carry out a spatial regime analysis to identify county-level factors correlated with diabetes prevalence in the Diabetes Belt versus the remainder of the U.S.

Results—Counties outside of the Diabetes Belt demonstrated stronger positive associations between diabetes prevalence and persistent poverty and greater percentages of unemployed labor forces. For counties in the Diabetes Belt, diabetes prevalence showed a stronger positive association with natural amenities (e.g., temperate climate and topographic features) and a stronger negative association with fitness/recreation facility density.

Conclusions—Community-level correlates of diabetes prevalence differed between the Diabetes Belt and elsewhere in the U.S. Economic hardship was shown to be more relevant outside the Diabetes Belt, while recreational context effects were more pronounced among counties within the region. Prevention and treatment targets are geographically unique and public health efforts should acknowledge these differences in crafting policy.

Keywords

diabetes prevalence; Diabetes Belt; spatial analysis; ecological analysis; community factors

Introduction

In the United States (U.S.), upward trends in diabetes prevalence witnessed in previous decades (1990 – 2008) have more recently plateaued.(1–3) While this is hopeful news, diabetes prevalence remains an important public health issue and requires continued research to further identify factors relevant to maintaining this progress.(1) Ecological studies exploring diabetes prevalence in the U.S. have shown geographic clusters of high and low diabetes prevalence within single communities,(4) among communities within a single state, (5, 6) as well as across the entirety of the U.S.(7–10) Among these studies, racial/ethnic populations and poverty,(7) food environment measures,(6, 8) and local business or entrepreneurial culture (9, 10) have all been shown to be associated with diabetes prevalence. Most importantly, ecological diabetes research has established a geographically distinct area known as the 'Diabetes Belt,' which is a region of counties (n=644) in the Southern U.S. characterized by diabetes prevalence of 11% or greater among the adult population.(11) Such geographic disparity in diabetes prevalence suggests that underlying community factors may be uniquely relevant within certain regions of the country compared to others. However, research to date on the Diabetes Belt has focused on the characteristics of individuals, such as their demographic characteristics and personal risk factors, that differ between residents of the region and those living elsewhere.(11) While certainly important, we maintain that it is also valuable to identify community-level factors that are differentially associated with diabetes prevalence across various regions of the U.S. This information can then be used, alongside other information, to help target health policies aimed at reducing the population health burden posed by diabetes.

In the current study, we used data from the Centers for Disease Control and Prevention's (CDC) Diabetes Interactive Atlas, which provides model-based estimates of adult diagnosed diabetes prevalence for 2009 among U.S. counties.(12) These data allowed us to address the objectives of this study, which were to: 1) investigate geographic disparities in diabetes prevalence; and 2) identify associations between county-level diabetes prevalence and community characteristics that differed between the Diabetes Belt and the remainder of the U.S. The strategy used to address these objectives advances ecological diabetes research by bringing spatial effects and additional county-level covariates into consideration and empirically establishing how place -based correlates of diabetes prevalence differ between the Diabetes Belt and others areas of the country.

Methods

Dependent Variable

Drawing on model-based estimates from the CDC, our dependent variable was the age-adjusted percent of the adult population (20 years) with diagnosed diabetes within a county for 2009.(7, 11, 13–16) County-level estimates of diabetes and selected risk factors (e.g., obesity, leisure-time physical inactivity) are model-based and derived from data using the CDC's Behavioral Risk Factor Surveillance System (BRFSS)(17) and the U.S. Census Bureau's Population Estimates Program.(18) Specifically, diabetes status was established by an affirmative response to the question "Have you ever been told by a doctor that you have diabetes?" While the BRFSS currently samples from nearly every county in the nation, small

sample sizes prevent the direct calculation of reliable county-specific estimates for most U.S. counties. To overcome this limitation, the CDC has drawn on the aforementioned data to develop county-level diabetes prevalence estimates for all U.S. counties using model-based small area estimation techniques. To increase the precision of year-specific county-level estimates, 3 years of BRFSS data are pooled for a given time point. For example, the CDC estimates for 2009 were based on data from 2008, 2009, and 2010, totaling approximately 1.3 million respondents. Validation-studies have compared estimates produced by this modeling technique against direct estimates from counties with large enough sample sizes and have shown little disagreement between the direct and model-based estimates.(15) Moreover, those involved in the production of the CDC's diabetes and associated risk factors estimates have encouraged research that explicitly incorporates spatial effects to describe and account for county-level patterns in these data.(7, 11) We analyzed all counties and county equivalents (including parishes in Louisiana and independent cities in Virginia) in the contiguous U.S. (counties in Alaska and Hawaii were excluded) (n=3,109).

Independent Variables

Drawing from the literature examining community correlates of population health, we focused on six community-level contexts in relation to diabetes prevalence: economic context, healthcare context, recreational context, food environment, population structure, and educational levels.(7, 8, 13, 14, 19, 20) Data for these measures were obtained from multiple national sources, including the CDC, U.S. Census Bureau, U.S. Department of Agriculture (USDA), and U.S. Department of Health and Human Services (HSS).

Variables that reflect the *economic context* of counties included: 1) the percentage of the population living at or below the federal poverty thresholds, 2) the percentage of the labor force that were unemployed, 3) residential segregation of the poor from the non-poor, and 4) persistent poverty. Data for these measures were obtained from the U.S. Census Bureau's 2005–2009 American Community Survey (ACS) 5-year Estimates. Measures of the *health and healthcare context* of counties included: 5) age-adjusted percentage of adults (20 years) with obesity, 6) the age-adjusted percentage of adults (20 years) who were physically inactive, 7) the number of outpatient visits per 1,000 people, 8), the percentage of the population without health insurance and 9) the number of physicians per 1,000 people. Obesity and physical inactivity prevalence data for adults were obtained from the CDC's Diabetes Interactive Atlas. Health insurance data were obtained from the U.S. Census Bureau's Small Area Health Insurance Estimates. Both physician and outpatient visits data were from the HHS Area Resource File. The *recreational context* of counties was captured by: 10) the number of fitness and recreation centers per 1,000 people, and 11) an index of natural amenities. Data for fitness and recreation centers was drawn from the USDA's Economic Research Service (ERS) Food Environment Atlas. Created by the USDA's ERS, values for the natural amenities index range from 1 (lowest) to 7 (highest) and capture the relative presence of factors such as a varied topography; access to water; warm, sunny winters; and temperate, low-humidity summers based on data from 1999.(21) Measures of the *food environment* included: 12) the number of grocery stores and supercenters per 1,000 people and 13) the number of fast food restaurants per 1,000 people. Both variables were drawn from the USDA's ERS Food Environment Atlas. The *population structure* of counties

was captured by: 14) the percentage of families headed by single mothers, 15) the percentage of the population   65 years, 16) the percentage of the population African American, 17) the percentage of the population Hispanic, and 18) urban influence using three dummy variables: metropolitan (reference), micropolitan, or non-core areas. Sociodemographic variables were obtained from the 2005–2009 ACS 5-year Estimates. Defined by the Office of Management and Budget (OMB), metropolitan (metro) areas are central counties with one or more urbanized areas (urbanized areas are densely-settled urban entities with 50,000 or more people), plus outlying counties that are economically tied to core counties as measured by labor-force commuting patterns; micropolitan (micro) areas are counties outside of metro areas (nonmetropolitan) that are centered on smaller urban clusters (10,000–49,999 people) and defined with the same criteria used to define metro areas; and noncore areas are all remaining counties not fitting the metro or micro criteria. (22) Last, *educational levels* were measured by: 19) the percentage of the population   25 years without a high school diploma or equivalent. This variable was from the 2005–2009 ACS 5-year Estimates.

Statistical Analysis

Previous research aimed at geographically defining the Diabetes Belt used individual-level data to examine the characteristics of people living in counties with 11% or higher diabetes prevalence compared to individuals living elsewhere. (11) Here we take an explicitly ecological approach by examining county-level adult diabetes prevalence across the contiguous U.S. using Local Indicators of Spatial Association (LISA) analysis. Rather than relying upon a simple descriptive cutoff, LISA results reveal distinct spatial patterns in county-level adult diabetes prevalence by identifying statistically significant geographic clusters of counties with similar levels of diabetes prevalence. Such geographic patterns are suggestive of different spatial regimes, a form of spatial heterogeneity where a specified parameter's values differ by location (e.g., ecological covariates differ by region).(23–25) Moreover, we carry out a spatial Chow test, which produces a statistic similar to the F-statistic, to detect if there are differences in selected covariates between counties inside and outside the Diabetes Belt and, in turn, the need for further analysis to explicate how these differences operate across regimes. (26–28)

Spatial regime analysis was next used to identify significant factors that were differentially associated with diabetes prevalence between counties in the Diabetes Belt (n=644) and the rest of the contiguous U.S. (n=2,465).(27, 29) The spatial regime analysis is essentially a fully interacted regression model with each independent variable interacted with a dummy indicator for Diabetes Belt (e.g., percent pop. poor * Diabetes Belt (1=yes)), along with main effects for each independent variable and the Diabetes Belt indicator.(30) The spatial regime model included a spatial lag term to address diagnosed issues of spatial autocorrelation (i.e., the significant clustering of counties with similar diabetes levels) present in the dependent variable (see Results). Adjusting for the spatially lagged measure of adult diabetes prevalence ensures that results are not biased by shared similarities in diabetes levels among neighboring counties.(31, 32) We used a 'rook' weights matrix, indicated by best model fit, to define each county's neighboring counties, which is a contiguity-based

measure that identifies a county's neighbors as those with shared borders.(33, 34) We utilized GeoDa 1.6.7 for the spatial diagnostics.(34)

A number of steps were taken to correctly specify the spatial regime model. Regression diagnostics did not indicate any issues of collinearity between the independent variables. Because counties are situated in states and states contain varying numbers of counties, we also included state fixed-effects to control for county-invariant variables within each state (e.g., state-specific health policies). We also grand mean centered (nation) each independent variable entered into the spatial regime model.(13) To account for multiple comparisons we reported significant associations using an alpha level of $p<0.01$. All regression analyses were carried out using IBM$^{©}$ SPSS$^{©}$ Statistics Version 20.

Results

Table 1 shows that county-level adult diagnosed diabetes prevalence in the Diabetes Belt (Mean±SD, 11.8±1.4%; n=644) was significantly higher ($p<0.001$) than the rest of the U.S. (8.7±1.6%; n=2,465). Figure 1 presents a LISA map of significant (pseudo $p<0.05$ based on a random permutation procedure (33)) contiguous clusters of counties across the U.S. with significantly higher or lower diabetes prevalence than would be expected if its distribution was spatially random. Notably, the LISA map shows the stark contrast in diabetes prevalence between the Diabetes Belt, which is clearly illustrated by significant clusters of high diabetes prevalence (shown in dark gray) counties in the South, and counties in the remainder of the U.S., particularly certain portions of the Northeast, Midwest, and West, that were disproportionately home to significant county-level clusters of low diabetes prevalence (shown in light gray). In fact, 82% of the counties in the Diabetes Belt were members of high diabetes clusters, while this was true for only 5% of counties outside the region. Conversely, 31% of counties outside the Diabetes Belt were members of low diabetes clusters, while this was true of no counties within the region.

Motivated by results from the LISA analysis and spatial Chow test (shown in Table 2), both of which showed the unequal impact of explanatory variables between the Diabetes Belt and the remainder of contiguous U.S. counties, we next conducted a spatial regime analysis to elucidate the specific determinants of diabetes prevalence that were unique to each region and significantly different between the regions. Results are shown in Table 3.

Diabetes Belt

Adult obesity prevalence, adult physical inactivity, African American populations, and natural amenities were associated with significantly higher diabetes prevalence among counties in the Diabetes Belt. In contrast, fitness/recreation facility density was significantly linked to lower diabetes levels in the region. Moreover, as suggested by the LISA map, diabetes prevalence among counties in the Diabetes Belt was significantly linked to membership in multi-county clusters with similarly high diabetes levels.

Non-Diabetes Belt

Poor populations, unemployment, persistent poverty, outpatient visits, adult obesity prevalence, physically inactive adults, families headed by single mothers, African American

populations, and less educated populations were each significantly related to higher diabetes prevalence among counties outside the Diabetes Belt. Conversely, physician density, populations age 65 and older, and Hispanic populations were each significantly associated with lower diabetes prevalence among such counties.

Significant Regional Differences

As indicated by significant interaction effects, there were a number of community factors associated with diabetes prevalence that differed between counties in the Diabetes Belt and elsewhere around the country. The impact of unemployment and persistent poverty on higher diabetes prevalence was significantly greater outside the Diabetes Belt (vs. in the Diabetes Belt). In contrast, the downward pressure on diabetes prevalence associated with recreation and fitness facilities was significantly more pronounced in the Diabetes Belt compared to counties outside of the Diabetes Belt. Moreover, natural amenities were associated with significantly higher diabetes in the Diabetes Belt than was true in other parts of the nation. Finally, the influence of regional clustering of like (high) levels of diabetes among counties was significantly more pronounced in the Diabetes Belt.

Discussion

The study undertaken here demonstrated that the Diabetes Belt is not only a distinctive contiguous region characterized by high diabetes prevalence, but also that the county-level characteristics associated with diabetes levels in this region differed in significant ways from other areas of the country. Importantly, these features go beyond predictable diabetes risk factors, such as obesity and physical inactivity levels, to highlight community features that are uniquely relevant within the region.

The results showed both similarities and differences in the ecological relationships that were significant among counties within and outside the Diabetes Belt. In both contexts, greater population shares of adults with obesity, physically inactive adults, and African Americans were associated with higher local diabetes prevalence. In the Diabetes Belt, greater natural amenities and spatial effects indicating significant geographic clustering of counties with high diabetes levels were linked to higher diabetes prevalence, while the density of recreation facilities was linked to lower local diabetes levels. Outside the Diabetes Belt, contemporary and persistent poverty, unemployment, outpatient doctor visits, single female headed families, and less educated populations were all linked to higher county-level diabetes. In contrast, greater physician density as well as older (age 65+) and Hispanic populations were negatively associated with local diabetes prevalence.

In terms of ecological relationships that were shown to be unique to counties in the Diabetes Belt compared to the rest of the U.S., greater fitness/recreation facility density was associated with significantly lower diabetes prevalence in this region, a finding that echoes the results of ecological studies on obesity prevalence.(13, 14, 20, 35) Conversely, a greater presence of natural amenities was related to significantly higher diabetes prevalence in the Diabetes Belt, a relationship that did not reach statistical significance among counties in the remainder of the country (p=0.051). This is an interesting finding given that ecological obesity studies that have shown greater natural amenities are linked to less obesity

prevalence at the county level.(14, 35, 36) The natural amenities scale captures desirable environmental qualities (e.g., temperate climate, mild humidity, varied topography, and water access) within counties. We surmise that this somewhat unexpected relationship could be due to the reduced range of natural amenity scores (2–5) in the Diabetes Belt compared to the remainder of the country (1–7). Last, there was a significant spatial effect in the Diabetes Belt showing that counties in the region were much more likely to be surrounded by neighboring counties with similar (high) levels of diabetes prevalence, even in the presence of the full range of other covariates, while no such effect was shown among counties outside the region. This finding provides further evidence that the Diabetes Belt is ecologically unique when it comes to the population burden of this disease.

Ecological relationships that were shown to be especially relevant outside the Diabetes Belt included county-level unemployment and persistent poverty. Both of these measures tap the influence of local economic hardship. Here it is also important to note that there is more variance in economic affluence and disadvantage outside the Diabetes Belt. Indeed, the Diabetes Belt is a part of the country characterized by longstanding economic challenges. For example, as shown in Table 1, 40% of counties in the Diabetes Belt were characterized by persistent poverty, while this was true of only 10% of counties outside the region.

We note that this study does have a number of potential shortcomings. First, the analysis herein is cross-sectional, which limits our ability to assert causality in the relationships between examined community factors and diagnosed diabetes prevalence. Rather, we can only highlight correlations that reach statistical significance during a single snapshot in time. Also, because our analysis is ecological we understand that our findings may be vulnerable to both the modifiable areal unit problem (MAUP) and the uncertain geographic context problem (UGCoP), which are threats inherent in most spatial analyses, including that the spatial scale chosen can change over time (e.g., county boundaries), the units of analysis were created for purposes other than that under study (e.g., counties are government administrative units not health districts), and ultimately the proper scale is not known (e.g., diabetes prevalence could be different at smaller – e.g., neighborhoods – or larger – e.g., states – geographic scales).(37, 38) This study is also limited by the fact that is must rely on model-based estimates produced by the CDC rather than direct estimates of diabetes prevalence. Population health census data that covers the entirety of the U.S. would be invaluable for chronic disease research, policy, and intervention, but it currently does not exist. However, we did carry out sensitivity analyses using the upper and lower confidence limits of the estimates to understand the potential impact on our findings. Results in separate models using the upper and lower diabetes prevalence estimates as dependent variables largely maintained in comparison to our original model. In the Diabetes Belt, adult obesity prevalence, adult physical inactivity, natural amenities, African American populations, and diabetes prevalence among neighboring counties maintained significant associations with higher diabetes prevalence. In the non-Diabetes Belt, unemployment, persistent poverty, outpatient visits, adult obesity prevalence, physically inactive adults, and African American populations maintained significant relationships with higher diabetes prevalence, while physician density, populations age 65 and older, and Hispanic populations maintained significant associations with lower diabetes prevalence. Regarding regional differences, the impact of unemployment and persistent poverty on higher diabetes prevalence maintained

significantly greater influence in the non-Diabetes Belt; natural amenities maintained a significantly stronger association with higher diabetes in the Diabetes Belt; and the influence of regional clustering of similar (high) levels of diabetes among counties maintained stronger significance in the Diabetes Belt. While the original model and the model using upper diabetes prevalence estimates found that density of fitness/recreation facilities held a stronger negative association with diabetes prevalence, the model using lower diabetes prevalence estimates did not demonstrate a similar association. Finally, and very importantly, we would encourage readers to be mindful not to commit the ecological fallacy in interpreting this research (the error of attributing an association between variables to a level of analysis that wasn't studied). This analysis examined county-level characteristics, the attributes of places not individual persons. That is, to find that counties that are home to a higher Hispanic population share tend to be characterized by lower diabetes prevalence does not contradict studies that have shown Hispanic individuals to have higher levels of diagnosed diabetes (2, 3), the units of analysis are different.

To our knowledge this is the first ecological analysis of community covariates related to county-level adult diagnosed diabetes prevalence that focused on differential associations between the Diabetes Belt and the rest of the contiguous U.S. Moreover, rather than examining a single community context, our study measured a range of local factors that spanned multiple community dimensions to provide a detailed analysis of the aggregate correlates of diabetes prevalence. Importantly, our findings further highlighted the disproportionate burden of diabetes in the Diabetes Belt, a characteristic that sets this region apart from the remainder of the U.S. Thus, the Diabetes Belt should remain a focal point for research and health policy initiatives. Prevention and treatment efforts that apply a universal or "one-size-fits-all" approach to addressing diabetes prevalence may well miss the mark and fail to reduce regional disparities in the burden of the disease.

Acknowledgments

REFERENCES

1. Herman WH, Rothberg AE. Prevalence of diabetes in the United States: A glimmer of hope? J Am Med Assoc. 2015; 314(10):1005–1007. PubMed PMID: 26348749.

2. Geiss LS, Wang J, Cheng YJ, Thompson TJ, Barker L, Li Y, et al. Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980–2012. J Am Med Assoc. 2014; 312(12):1218–1226. PubMed PMID: 25247518.

3. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and trends in diabetes among adults in the United States, 1988–2012. J Am Med Assoc. 2015; 314(10):1021–1029. Epub 2015/09/09. PubMed PMID: 26348752.

4. Drewnowski A, Rehm CD, Moudon AV, Arterburn D. The geography of diabetes by census tract in a large sample of insured adults in King County, Washington, 2005–2006. Prev Chronic Dis. 2014; 11:E125. PubMed PMID: 25058671; PubMed Central PMCID: PMC4112927. [PubMed: 25058671]

5. Stewart JE, Battersby SE, Fede AL-D, Remington KC, Hardin JW, Mayfield-Smith K. Diabetes and the socioeconomic and built environment: geovisualization of disease prevalence and potential contextual associations using ring maps. International Journal of Health Geographics. 2011; 10(1): 18–27. PubMed PMID: 59761737. [PubMed: 21362176]

6. AlHasan DM, Eberth JM. An ecological analysis of food outlet density and prevalence of type II diabetes in South Carolina counties. BMC Public Health. 2016; 16(1):10. PubMed PMID: 26728042; PubMed Central PMCID: PMC4700568. [PubMed: 26728042]

7. Shrestha SS, Kirtland KA, Thompson TJ, Barker LE, Gregg EW, Geiss LS. Spatial clusters of county-level diagnosed diabetes and associated risk factors in the United States. The Open Diabetes Journal. 2012; 5:29–37.

8. Ahern M, Brown C, Dukas S. A national study of the association between food environments and county-level health outcomes. The Journal of Rural Health. 2011; 27:367–3279. [PubMed: 21967380]

9. Blanchard TC, Li J, Mencken C, Tolbert CM. Entrepreneurial environment and the prevalence of diabetes in U.S. counties. ISRN Public Health. 2012; 2012:5.

10. Blanchard TC, Tolbert C, Mencken C. The health and wealth of US counties: how the small business environment impacts alternative measures of development. Cambridge Journal of Regions, Economy and Society. 2012; 5(1):149–162.

11. Barker LE, Kirtland KA, Gregg EW, Geiss LS, Thompson TJ. Geographic distribution of diagnosed diabetes in the U.S.: A diabetes belt. Am J Prev Med. 2011; 40:434–439. [PubMed: 21406277]

12. Kirtland KA, Burrows NR, Geiss LS. Diabetes Interactive Atlas. Prev Chronic Dis. 2014; 11:130300. PubMed PMID: 24503340; PubMed Central PMCID: PMC3917609. [PubMed: 24503340]

13. Myers CA, Slack T, Martin CK, Broyles ST, Heymsfield SB. Regional disparities in obesity prevalence in the United States: A spatial regime analysis. Obesity. 2015; 23:481–487. PubMed PMID: 25521074; PubMed Central PMCID: PMC4310761. [PubMed: 25521074]

14. Slack T, Myers CA, Martin CK, Heymsfield SB. The geographic concentration of U.S adult obesity prevalence and associated social, economic, and environmental factors. Obesity. 2014; 22:868–874. [PubMed: 23630100]

15. Cadwell BL, Thompson TJ, Boyle JP, Barker LE. Bayesian small area estimates of diabetes prevalence by U.S. county, 2005. Journal of Data Science. 2010; 8:173–188.

16. Centers for Disease Control and Prevention (CDC). Centers for disease control and prevention diabetes interactive atlas. 2009. Available from: http://www.cdc.gov/diabetes/atlas/

17. Centers for Disease Control and Prevention (CDC). Behavioral risk factor surveillance system (BRFSS). Available from: http://www.cdc.gov/brfss/index.htm

18. United States Census Bureau. Population estimates. Available from: http://www.census.gov/popest/index.html

19. Hillemeier MM, Lynch J, Harper S, Casper M. Measuring contextual characteristics for community health. Health Serv Res. 2003; 38:1645–1718. PubMed PMID: 12196095. [PubMed: 14727793]

20. Myers CA, Slack T, Martin CK, Broyles ST, Heymsfield SB. Change in obesity prevalence across the United States is influenced by recreational and healthcare contexts, food environments, and Hispanic populations. PLoS One. 2016; 11(2):e0148394. PubMed PMID: 26849803; PubMed Central PMCID: PMC4743954. [PubMed: 26849803]

21. McGranahan, DA. Natural Amenities Drive Rural Populaiton Change. Washington, DC: United State Department of Agriculture; 1999.

22. United States Department of Agriculture (USDA) Economic Research Service (ERS). Urban Influence Codes. Available from: http://www.ers.usda.gov/data-products/urban-influence-codes.aspx#.U2FmtoFdW0I

23. Chi G, Zhu J. Spatial regression models for demographic analysis. Population Research & Policy Review. 2008; 27:17–42. PubMed PMID: 28090041.

24. Baller RD, Anselin L, Messner SF, Deane G, Hawkins DF. Structural covariates of U.S. county homicide rates: Incorporating spatial effects. Criminology. 2001; 39:561–588.

25. Michimi A, Wimberly MC. Spatial patterns of obesity and associated risk factors in the conterminous U.S. Am J Prev Med. 2010; 39:e1–e12. [PubMed: 20621256]

26. Anselin L. Spatial dependence and spatial structural instability in applied regression analysis. Journal of Regional Science. 1990; 30:185–207.

27. Curtis KJ, Voss PR, Long DD. Spatial variation in poverty-generating processes: Child poverty in the United States. Soc Sci Res. 2012; 41:146–159. [PubMed: 23017703]

28. Chow GC. Test of equality between sets of coefficients in two linear regressions. Econometria. 1960; 28:591–605.

29. Shoff C, Yang TC. Spatially varying predictors of teenage birth rates among counties in the United States. Demographic Research. 2012; 27:377–418. PubMed PMID: 23144587; PubMed Central PMCID: PMC3493119. [PubMed: 23144587]

30. Anselin, L. Spatial Econometrics: Methods and Models. Cordrecht: Kluwer Academic Publishers; 1988.

31. Schultz J, Elliott J. Natural disasters and local demographic change in the United States. Popul Environ. 2013; 34:293–312.

32. Ward, MD., Gleditsch, KS. Spatial Regression Models. Thousand Oaks: Sage Publications; 2008.

33. Anselin L. Exploring spatial data with GeoDa. 2005 Available from: https://geodacenter.asu.edu/system/files/geodaworkbook.pdf.

34. Anselin L, Syabri I, Kho Y. GeoDa: An introduction to spatial data analysis. Geographical Analysis. 2006; 38:5–22.

35. Jilcott Pitts SB, Edwards MB, Moore JB, Shores KA, DuBose KD, McGranahan D. Obesity is inversely associated with natural amenities and recreation facilities per capita. Journal of Physical Activity & Health. 2013; 10(7):1032–1038. Epub 2012/11/09. PubMed PMID: 23136370. [PubMed: 23136370]

36. Jilcott SB, Moore JB, Shores KA, Imai S, McGranahan DA. Associations between natural amenities, physical activity, and body mass index in 100 North Carolina counties. Am J Health Promot. 2011; 26(1):52–55. PubMed PMID: 21879944. [PubMed: 21879944]

37. Saib M-S, Caudeville J, Carre F, Ganry O, Trugeon A, Cicolella A. Spatial relationship quantification between environmental, socioeconomic and health data at different geographic levels. International Journal of Environmental Research and Public Health. 2014; 11:3765–3786. PubMed PMID: doi:10.3390/ijerph110403765. [PubMed: 24705362]

38. Kwan M-P. The uncertain geographic context problem. Annals of the Association of American Geographers. 2012; 102:958–968. PubMed PMID: 78222449.

What is already known about this subject?

- County-level diabetes prevalence is not equally distributed across the United States (U.S.)

- The Diabetes Belt is a region of counties in the Southern U.S. with comparatively higher rates of diagnosed diabetes.

What does this study add?

- This study extends the literature to identify local social, economic, and environmental features that are differentially associated with county-level diabetes prevalence inside and outside the Diabetes Belt.

- The results suggest community factors that might be more relevant in the Diabetes Belt versus the remainder of the contiguous U.S. and vice versa.

- These community factors go beyond the usual risk factors (e.g., obesity and physical inactivity prevalence) and highlight geographically unique local features to be considered for targeted public health efforts.
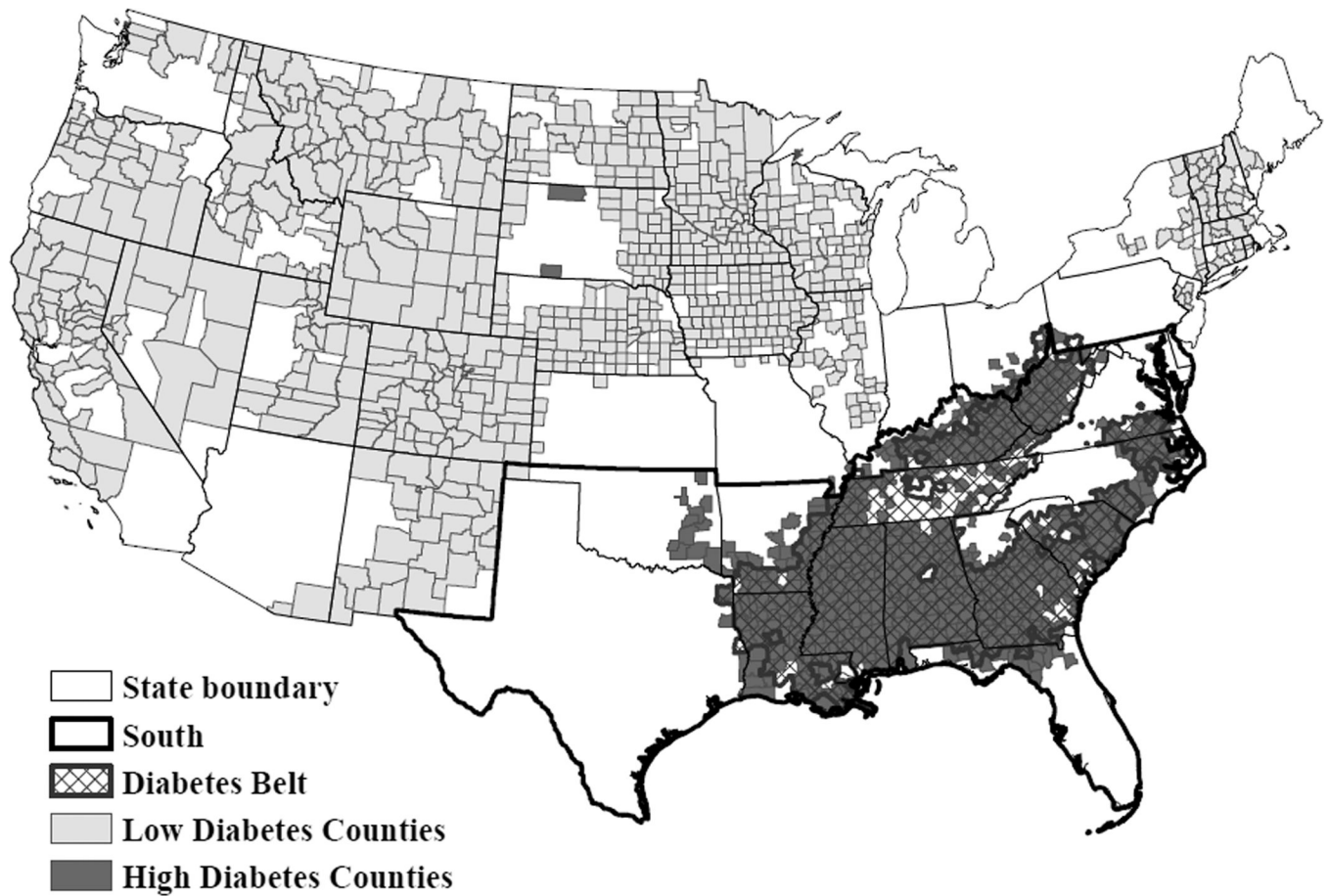
Figure 1. Local Indicators of Spatial Association (LISA) map of significant geographic concentrations of county-level diagnosed diabetes prevalence, 2009

Moran's I = 0.74; p<0.05. Light gray-shaded counties are core members of geographic clusters with significantly (p<0.05) lower adult diagnosed diabetes prevalence than would be expected at random. Dark gray-shaded counties are core members of geographic clusters with significantly (p<0.05) higher adult diagnosed diabetes prevalence.

Table 1

Descriptive statistics

Variables	United States	Diabetes Belt	Non-Diabetes Belt
	Mean (S.D.)	Mean (S.D.)	Mean (S.D.)
Dependent Variable			
Percent adults with diabetes	9.1 (2.0)	11.6 (1.3)	8.5 (1.5)
Independent Variables			
Economic context			
Percent of pop. poor	15.4 (6.5)	21.2 (6.1)	13.9 (5.7)
Percent of labor force unemployed	4.1 (1.7)	5.0 (1.6)	3.9 (1.7)
Poor/non-poor segregation	18.8 (10.8)	17.5 (8.8)	19.1 (11.2)
Persistent poverty	0.1 (0.3)	0.4 (0.5)	0.1 (0.2)
Health and healthcare context			
Percent of adults with obesity	30.3 (4.2)	34.2 (3.2)	29.3 (3.9)
Percent of adults physically inactive	26.9 (4.9)	31.6 (3.6)	25.7 (4.5)
Percent of pop. uninsured	18.3 (5.6)	20.0 (3.1)	17.9 (6.2)
Number of physicians/1,000 pop.	1.5 (1.8)	1.2 (1.2)	1.6 (1.9)
Number of outpatient visits/1,000 pop.	2,431.2 (3,326.9)	1,773.7 (2,152.8)	2,602.9 (3,550.9)
Recreational context			
Number of recreation facilities per 1,000 pop.	0.1 (0.1)	0.1 (0.1)	0.1 (0.1)
Natural amenities (low of 1 to high of 7)	3.5 (1.0)	3.4 (0.6)	3.5 (1.1)
Food environment			
Number of grocery stores & supercenters/1,000 pop.	0.3 (0.2)	0.2 (0.1)	0.3 (0.2)
Number of fast food restaurants/1,000 pop.	0.6 (0.3)	0.5 (0.3)	0.6 (0.3)
Population structure			
Percent of families headed by single mothers	9.6 (3.8)	12.6 (4.6)	8.9 (3.2)
Percent of pop. aged 65 and older	15.4 (4.2)	14.4 (2.7)	15.6 (4.5)
Percent of pop. African American	8.9 (14.4)	25.2 (21.1)	4.6 (7.7)
Percent of pop. Hispanic	7.6 (12.8)	2.4 (2.6)	8.9 (14.0)
Metropolitan	0.4 (0.5)	0.3 (0.5)	0.4 (0.5)
Micropolitan	0.2 (0.4)	0.2 (0.4)	0.2 (0.4)
Noncore	0.4 (0.5)	0.5 (0.5)	0.4 (0.5)
Educational level			
Percent of adults with less than high school	17.7 (7.3)	24.2 (6.0)	15.2 (6.4)
High diabetes	21%	82%	5%
Low diabetes	24%	0%	31%
N	3,109	644	2,465

Source: Obesity and physical inactivity prevalence data are from the CDC County-level Estimates of Obesity and Leisure-Time Physical Inactivity. Uninsured data are from the U.S. Census Bureau's Small Area Health Insurance Estimates. Physician and outpatient visits data are from the Area Resource File, via the Department of Health and Human Services. Recreation facilities, grocery stores and supercenters, and fast food restaurant data are from the USDA Economic Research Service's Food Environment Atlas and the U.S. Census Bureau's County Business Patterns. Metropolitan, micropolitan, and noncore status data are from the USDA Economic Research Service. All other data were obtained from the U.S. Census Bureau's 2005–2009 American Community Survey 5-year Estimates.

Table 2

OLS regression models of county-level age-adjusted adult diagnosed diabetes prevalence, 2009

Variables	United States		Diabetes Belt		Non-Diabetes Belt	
	b	95% CI	b	95% CI	b	95% CI
Economic context						
Percent of pop. poor	0.016[a]	0.009, 0.024	0.008	−0.010, 0.027	0.012[a]	0.004, 0.020
Percent of labor force unemployed	0.086[a]	0.065, 0.107	−0.035	−0.083, 0.012	0.123[a]	0.100, 0.147
Poor/non-poor segregation	0.003	−0.001, 0.006	0.003	−0.006, 0.012	0.002	−0.001, 0.006
Persistent poverty	0.257[a]	0.154, 0.370	0.127	−0.068, 0.012	0.446[a]	0.289, 0.603
Health and healthcare context						
Percent of adults with obesity	0.129[a]	0.117, 0.141	0.130[a]	0.099, 0.160	0.122[a]	0.109, 0.135
Percent of adults physically inactive	0.052[a]	0.042, 0.127	0.053[a]	0.027, 0.079	0.047[a]	0.035, 0.059
Number of physicians/1,000 pop.	−0.041[a]	−0.062, −0.021	−0.023	−0.102, 0.056	−0.045[a]	−0.066, −0.024
Percent of pop. uninsured	0.007	−0.005, 0.020	0.003	−0.039, 0.044	0.007	−0.005, 0.020
Number of outpatient visits/1,000 pop.	0.033[a]	0.023, 0.042	0.010	−0.028, 0.048	0.033[a]	0.024, 0.042
Recreation context						
Number of recreation facilities/1,000 pop.	−0.457	−0.809, −0.104	−2.355[a]	−3.770, −0.940	−0.341	−0.689, 0.006
Natural amenities (low of 1 to high of 7)	0.084[a]	0.042, 0.127	0.221[a]	0.095, 3.48	0.048	0.004, 0.092
Food environment						
Number of grocery stores & supercenters/1,000 pop.	0.059	−0.090, 0.208	0.465	−0.123, 1.054	0.053	−0.096, 0.201
Number of fast food restaurants/1,000 pop.	−0.032	−0.125, 0.062	0.072	−0.261, 0.405	−0.039	−0.132, 0.055
Population structure						
Percent of families headed by single mothers	0.013	0.002, 0.025	−0.010	−0.036, 0.015	0.020[a]	0.007, 0.032
Percent of pop. 65 and older	−0.032[a]	−0.041, −0.024	−0.024	−0.055, 0.007	−0.030[a]	−0.038, −0.021
Percent of pop. African American	0.020[a]	0.017, 0.024	0.031[a]	0.023, 0.038	0.022[a]	0.016, 0.027
Percent of pop. Hispanic	−0.012[a]	−0.016, −0.008	0.023	−0.007, 0.053	−0.016[a]	−0.020, −0.012
Metropolitan (ref.)	----		----		----	
Micropolitan	−0.112[a]	−0.189, −0.034	−0.122	−0.318, 0.073	−0.089	−0.172, −0.006
Noncore	−0.049	−0.132, 0.035	−0.126	−0.323, 0.071	−0.025	−0.115, 0.066
Education level						

Variables	United States		Diabetes Belt		Non-Diabetes Belt	
Percent of adults with less than high school	0.007	−0.001, 0.015	−0.004	−0.023, 0.014	0.012[a]	0.004, 0.021
Spatial Lag	0.072[a]	0.032, 0.113	0.148[a]	0.046, 0.249	0.013	−0.032, 0.058
Intercept	8.675[a]	8.437, 8.914	10.144[a]	9.787, 10.501	8.605[a]	8.369, 8.841
Adjusted R^2	0.865		0.635		0.800	
Sum of Squares Residual	10,257.488		723.158		4,555.407	
N	3,109		644		2,465	
Spatial Chow test for differences between Diabetes Belt vs. Non-Diabetes Belt					131.41[a]	

Notes: 'Pop.' is an abbreviation for 'population. Unstandardized regression coefficients from regression models with state fixed effects. Number of outpatient visits per 1,000 pop. coefficient multiplied by 1,000. CI=confidence interval.

[a] p<.01.

Source: Obesity and physical inactivity prevalence data are from the CDC County-level Estimates of Obesity and Leisure-Time Physical Inactivity. Uninsured data are from the U.S. Census Bureau's Small Area Health Insurance Estimates. Physician and outpatient visits data are from the Area Resource File, via the Department of Health and Human Services. Recreation facilities, grocery stores and supercenters, and fast food restaurant data are from the USDA Economic Research Service Food Environment Atlas and the U.S. Census Bureau's County Business Patterns. Metropolitan, micropolitan, and noncore status data are from the USDA Economic Research Service. All other data were obtained from the U.S. Census Bureau's 2005–2009 American Community Survey 5-year Estimates.

Table 3

Spatial regime model of county-level age-adjusted adult diagnosed diabetes prevalence, 2009

Variables	Diabetes Belt		Non-Diabetes Belt	
	b	95% CI	b	95% CI
Economic context				
Percent of pop. poor	0.011	−0.004, 0.027	0.013^a	0.004, 0.021
Percent of labor force unemployed	−0.027	−0.069, 0.015	0.119^{ab}	0.095, 0.143
Poor/non-poor segregation	0.001	−0.007, 0.009	0.002	−0.002, 0.005
Persistent poverty	0.092	−0.075, 0.259	0.444^{ab}	0.280, 0.607
Health and healthcare context				
Percent of adults with obesity	0.141^a	0.115, 0.167	0.122^a	0.109, 0.135
Percent of adults physically inactive	0.062^a	0.040, 0.084	0.045^a	0.033, 0.058
Number of physicians/1,000 pop.	0.007	−0.060, 0.074	−0.049^a	−0.071, −0.028
Percent of pop. uninsured	−0.005	−0.036, 0.026	0.011	−0.002, 0.023
Number of outpatient visits/1,000 pop.	0.009	−0.025, 0.042	0.034^a	0.025, 0.044
Recreation context				
Number of recreation facilities/1,000 pop.	−2.384^{ab}	−3.632, −1.137	−0.328	−0.689, 0.034
Natural amenities (low of 1 to high of 7)	0.224^{ab}	0.114, 0.334	0.045	0.000, 0.090
Food environment				
Number of grocery stores & supercenters/1,000 pop.	0.312	−0.201, 0.824	0.055	−0.099, 0.209
Number of fast food restaurants/1,000 pop.	0.099	−0.189, 0.386	−0.036	−0.133, 0.061
Population structure				
Percent of families headed by single mothers	−0.009	−0.032, 0.013	0.020^a	0.007, 0.033
Percent of pop. 65 and older	−0.029	−0.055, −0.003	−0.030^a	−0.039, −0.021
Percent of pop. African American	0.026^a	0.020, 0.031	0.022^a	0.017, 0.028
Percent of pop. Hispanic	0.008	−0.015, 0.032	−0.016^a	−0.020, −0.012
Metropolitan (ref.)	----		----	
Micropolitan	−0.056	−0.222, 0.110	−0.112	−0.198, −0.027
Noncore	−0.052	−0.221, 0.117	−0.052	−0.146, 0.041
Education level				
Percent of adults with less than high school	−0.003	−0.018, 0.012	0.012^a	0.003, 0.021
Spatial Lag	0.223^{ab}	0.148, 0.299	0.004	−0.041, 0.050
Intercept	8.609^a[8.364, 8.853]			
Adjusted R^2	0.869			
N	3,109			

Notes: Unstandardized regression coefficients produced by a single OLS regression model which induces all main effects, interaction effects, state fixed effects and spatial lag. 'Pop.' is an abbreviation for 'population. Number of outpatient visits per 1,000 pop. coefficient multiplied by 1,000. CI=confidence interval.

^{a}p<.01.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript