Published in final edited form as:
Health Place. 2010 May ; 16(3): 547–556. doi:10.1016/j.healthplace.2009.12.015.

Evaluating geographic variation in Type 1 and Type 2 diabetes mellitus incidence in youth in four U.S. regions

Angela D. Liese[a], **Andrew Lawson**[b], **Hae-Ryoung Song**[a], **James D. Hibbert**[a], **Dwayne E. Porter**[c], **Michele Nichols**[a], **Archana P. Lamichhane**[a], **Dana Dabelea**[d], **Elizabeth J. Mayer-Davis**[e], **Debra Standiford**[f], **Lenna Liu**[g], **Richard F. Hamman**[d], and **Ralph B. D'Agostino Jr**[h]

Angela D. Liese: liese@sc.edu; Andrew Lawson: lawsonab@musc.edu; Hae-Ryoung Song: Hrsong14@hotmail.com; James D. Hibbert: hibbert@sc.edu; Dwayne E. Porter: porter@sc.edu; Michele Nichols: nichols@sc.edu; Archana P. Lamichhane: longroof@gmail.com; Dana Dabelea: Dana.Dabelea@ucdenver.edu; Elizabeth J. Mayer-Davis: ejmayer_davis@unc.edu; Debra Standiford: Debbie.Standiford@cchmc.org; Lenna Liu: lennall@u.washington.edu; Richard F. Hamman: richard.hamman@ucdenver.edu; Ralph B. D'Agostino: rdagosti@wfubmc.edu

[a] Department of Epidemiology and Biostatistics and Center for Research in Nutrition and Health Disparities, Arnold School of Public Health, University of South Carolina, 921 Assembly Street, Columbia, SC 29208, USA

[b] Medical University of South Carolina College of Medicine, Charleston, SC, USA

[c] Department of Environmental Health Sciences, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA

[d] Department of Epidemiology, Colorado School of Public Health, University of Colorado Denver; Aurora, CO, USA

[e] University of North Carolina at Chapel Hill, Gillings School of Global Public Health and School of Medicine, Chapel Hill, NC, USA

[f] Children's Hospital Medical Center, Cincinnati, OH, USA

[g] Seattle Children's Hospital Research Institute, Seattle, WA, USA

[h] Wake Forest University School of Medicine, Winston-Salem, NC, USA

Abstract

We evaluated geographic variation of Type 1 and Type 2 diabetes mellitus (T1DM, T2DM) in four regions of the United States.

Data on 807 incident T1DM cases diabetes and 313 T2DM cases occurring in 2002-03 in South Carolina (SC) and Colorado (CO), 5 counties in Washington (WA), and an 8 county region around Cincinnati, Ohio (OH) among youth aged 10 through 19 years were obtained from the SEARCH for Diabetes in Youth Study. Geographic patterns were evaluated in a Bayesian framework.

Incidence rates differed between the study regions, even within race/ethnic groups. Significant small area variation within study region was observed for T1DM and for T2DM. Evidence for joint spatial correlation between T1DM and T2DM was present at the county level for SC ($r_{SC=}$ 0.31) and CO

Address for corresponding author and requests for reprints: Angela D. Liese, PhD, MPH, Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, 800 Sumter Street, Columbia, SC 29208, Phone: (803) 777-9414; Fax: (803) 777-2524; liese@sc.edu.

non-Hispanic whites ($r_{CO=}$ 0.40) and CO Hispanics ($r_{CO=}$ 0.72). At the tract level no evidence for meaningful joint spatial correlation was observed ($r_{SC=}$ -0.02; $r_{CO=}$ -0.02; $r_{OH=}$ 0.03; $r_{WA=}$ 0.09).

Our study provides evidence for the presence of both regional and small-area, localized variation in type 1 and type 2 incidence among youth aged 10-19 years in the United States.

Keywords

Diabetes mellitus; youth; spatial epidemiology; Bayesian methods

Introduction

Diabetes mellitus is one of the leading chronic diseases of childhood and youth and numerous studies have documented an increase of diabetes worldwide (Onkamo et al., 1999). While in the past pediatric diabetes has been thought of largely as auto-immune, insulin-dependent type 1, the recent emergence of pediatric type 2 diabetes (non-autoimmune and non-insulin-dependent) (Pihoker et al., 1998; Pinhas-Hamiel et al., 1996; Scott et al., 1997) has raised the question whether type 1 and type 2 diabetes are truly distinct entities and whether there may be an overlay of etiologic risk factors (Wilkin, 2001). One such common factor may be obesity which has been shown to be associated with younger age at onset of type 1 and is associated with type 2 diabetes in youth. It is clear that both type 1 and type 2 diabetes are complex diseases caused by gene-environment interactions. Hence evaluation of geographic patterning of disease may reveal important environmental etiologic clues.

Based on data from the World Health Organization Multinational Project for Childhood Diabetes (DiaMond) (Karvonen et al., 1993), we know that the geographic variation in type 1 diabetes incidence is one of the largest observed for any non-communicable disease (LaPorte et al., 1995). Incidence rates follow a North-South gradient, being much higher in Scandinavian countries. In addition, marked within-country variation in incidence has also been documented (Dorman et al., 1995; Waldhor et al., 2000), increasing with northern latitude in China (Yang et al., 1998), Sardinia (Casu et al., 2004b), and Germany (Rosenbauer et al., 1999). Additionally, migration seem to reduce the age of onset (Cadario et al., 2004) and areas characterized by low population mixing seem to have the highest rates of type 1 diabetes (Parslow et al., 2001), suggesting the role of non-genetic, potentially environmental factors. Geographic variation of type 1 diabetes has not been explored systematically in the United States. It has been suggested that regional variation in type 1 diabetes may be explained to a large extent by variation in the racial/ethnic composition of the populations (LaPorte et al., 1995). Type 2 diabetes has not been explored in a geographic framework in the United States or elsewhere.

The SEARCH for Diabetes in Youth Study was initiated in 2000 to estimate the population prevalence and incidence of all types of diabetes in youth by age, sex, and race/ethnicity (SEARCH Study Group, 2004) in four geographically defined populations and two health-plan based populations using consistent methodology for case ascertainment and diabetes classification. The SEARCH study presented the unique opportunity to study regional and small-area geographic variation of pediatric diabetes risk based on data collected under a standardized and comprehensive surveillance system that ascertained both type 1 and type 2 diabetes in youth, explicitly taking into account race/ethnic differences in populations. We hypothesized that both the incidence of pediatric type 1 and type 2 diabetes would exhibit spatial correlation. We also considered that there might be joint spatial correlation (cross-correlation) between the two diseases in that areas with a high incidence of type 1 might also exhibit a high incidence of type 2 diabetes.

Research Design and Methods

Study design

Details of the SEARCH study design have been published (SEARCH Study Group, 2004). In brief, SEARCH is a six-center observational study (four geographically-based sites and two health plan membership-based sites) that began conducting population-based ascertainment of non-gestational cases of diagnosed diabetes in youth less than 20 years of age in 2001 for prevalent cases and 2002 for incident cases. SEARCH continues to ascertain incident cases through the present. The SEARCH surveillance component aims to enumerate and identify all eligible cases of diabetes based on networks of pediatric and adult endocrinologists, existing pediatric diabetes databases, hospitals, the databases of health plans, and other health care providers. Case reports are validated through physician reports, medical record reviews, or in a few instances, self-report of a physician's diagnosis of diabetes (SEARCH Study Group, 2004). Case reports are registered anonymously with the Coordinating Center at Wake Forest University in North Carolina using Health Insurance Portability and Accountability Act (HIPAA) compliant procedures. Identifying information is retained at each field center. Each center's institutional review board approved the study protocol which complies with the privacy rules of the HIPAA.

We included data on all incident cases occurring among youth aged 10-19 years in 2002 and 2003 in the four geographically defined regions comprising 1) the state of South Carolina, 2) the state of Colorado, 3) five counties around Seattle, Washington, including King, Kitsap, Pierce, Snohomish, and Thurston counties, and 4) eight counties around Cincinnati, Ohio including Butler, Clermont, Hamilton, Warren counties in Ohio, Boone, Campbell, and Kenton counties in Kentucky, and Dearborn, Indiana, which we will refer to as the Ohio region. The 10-19 year age group was chosen because type 2 diabetes is extremely rare in children under the age of 10 (Dabelea et al., 2007). Throughout this paper we refer to the four geographic areas as regions.

Demographic and clinical characteristics of cases

Basic demographic and clinical information was available for virtually all cases and generally available from a variety of data sources such as administrative data, medical record or self-report. A hierarchical approach was used to classify case characteristics (Liese et al., 2008).

Demographic information was utilized as follows. Age was categorized into two groups (10-14 years and 15-19 years) to align with published incidence rates (Dabelea et al., 2007). Race/ethnicity was classified into six groups (Hispanic, Non-Hispanic white, African American, Asian/Pacific Islander, and American Indian/Native American, Multiple and other) following the Census 2000 approach (Census Bureau (US), 2000). About 2% of our cases were non-Hispanic and multi-racial and needed to be removed after initial descriptive analyses. Diabetes type, as assigned by the health care provider, was categorized as type 1 (combining 1, 1a, and 1b), type 2, and other type (including hybrid type, maturity onset of diabetes in youth, type designated as “other”, type unknown by the reporting source, and missing).

Geocoding and geo-spatial allocation

Details of the geocoding process and success rates have been described (Hibbert et al., 2009). Geocoding was conducted in a standardized manner by a single staff person (J.H.) traveling to each field center and using ArcGIS 9.3 software (ESRI, 2008), the TIGER 2000 Road Network File complemented with Zip Code Tabulation Areas (ZCTA) data (U.S.Census Bureau, 2000). In South Carolina, this was supplemented with TIGER 2006 vintage Road Network Files, because recent land development has led to realigned street features that were not captured by the TIGER 2000 files. In brief, the vast majorities of cases were geocoded to the

street address level (overall 71%, SC 64%, CO 83%, WA 52%, OH 86%) and could thereby be allocated to a census tract. There remained 253 cases (23%) that were geocodable to a zip code level only, which were allocated by a random assignment based imputation method (Henry and Boscoe, 2008; Hibbert et al., 2009) rather than the traditional zip code centroid method (Cayo and Talbot, 2003) because the latter would create spurious clusters of cases in those tracts containing the zip code centroid. The remaining cases (6%) with information on county only could not be either geocoded or allocated to the census tract. Each case was assigned to a census tract within the boundaries of the known zip code based on a random assignment distribution process that was weighted by the proportions of the population age 0-19 residing in each of the tracts (or tract segments) that fell within the zip code boundaries. Details of this method have been described in (Henry and Boscoe, 2008) and evaluated specifically in our study areas in (Hibbert et al., 2009). In each census tract the total number of cases was determined by summing the cases of each diabetes type within that tract for both years, which constitutes the tract-specific observed cases for analyses outlined below.

Analysis sample

A total of 1,197 cases of diabetes developed in the two-year period among youth aged 10-19 years across the four regions, of whom 77 cases with diabetes other than type 1 or 2 (1 hybrid, 1 MODY, 2 other type, 22 secondary type, 37 unknown, 14 missing) were ineligible for the purposes of our study. Of the remaining 1,120, we excluded 23 cases who were multi-racial or of specified other race for the purpose of estimating race-specific incidence rates by region (n=1,097). Finally, 6 cases on whom only the state of residence was known were removed for county level analyses and modeling (n=1,091) and 65 cases with county information only were excluded for tract level analyses, leaving n=1,026, i.e. 92% of the eligible incident 2002 and 2003 cases.

Statistical analyses

Population estimates for each of the 3,138 Census tracts in our study region were obtained from Census 2000 Summary File1 (Census Bureau (US), 2000) which were age, sex, and race-group-specific. Regional annual incidence rates were computed by summing the number of incident cases of diabetes occurring in 2002 and in 2003 in a given region, dividing this numerator by two, and dividing the numerator by the 2000 Census-based denominator estimates. We present crude incidence rates by five-year age groups and by age and race. Ninety-five percent confidence intervals (CIs) were calculated on the basis of inverting the score test for a binomial proportion (Agresti A. and Coull, 1998).

In order to facilitate comparison of diabetes risks across space independent of age, gender and race-differences between populations, the following analyses adjusted for these demographic factors by including an expected number of cases into the analyses. The annual expected number of type 1 and type 2 diabetes cases was calculated by multiplying the population estimates with the corresponding published pooled (age group,- sex,- and race-specific) annual incidence rates for type 1 and type 2 diabetes from the SEARCH for Diabetes in Youth Study (Dabelea et al., 2007) and then doubled to reflect to the two year period used in this study. The total expected count for each geographic unit (i.e. county or census tract) was then calculated by summing the stratum specific expected counts within each unit. The standardized incidence ratio (SIR) was calculated by dividing the total observed cases by the total expected cases for each geographic unit. In addition, race group specific SIR's were calculated. The county or tract-specific SIRs were empirically correlated using Pearson's correlation coefficients to address the question of joint spatial correlation of type 1 and type 2. SIRs are known to be very crude estimators of risk and are highly variable over space for rare diseases such as diabetes. Therefore, our statistical analyses were largely carried out in a Bayesian framework to better address the inherent sparseness of our data.

For both type 1 and type 2 diabetes, Poisson regression models were used to a) generate spatially smoothed risk surfaces that were subsequently mapped by region, b) to evaluate spatial autocorrelation (disease clustering) within each region and c) to evaluate disease-specific cross-correlation between both diseases within regions. Each analytical step is described in more detail below. Based on extensive evaluations of a total of six Poisson models described in (Song 2008, in review), we selected the Sparse Poisson Convolution (SPC) model (which is an adaptation of the classical conditional autoregressive (CAR) model (Besag and Kooperberg, 1995; Besag et al., 1991) for sparse data) for aims a) - c) and the Sparse Poisson MCAR (SPMCAR) model (Gelfand and Vounatsou, 2003) for aim c). A detailed description of the SPC and SPMCAR models is given in Appendix 1. In brief, the SPC model is a CAR model augmented by an added term which is a function of an indicator variable denoting zero count or non-zero count in any Census tract. Hence a factored intercept is included in the SPC model. The SPMCAR is the extension of the SPC model to multivariate outcomes, i.e. a factored intercept is added to each disease model.

The model performance was assessed by the Deviance Information Criterion (DIC), which measures overall goodness of fit, and Mean Squared Prediction Error (MSPE), which evaluates predictive capability. Each model was fit to the observed number of cases in each geographic unit (county, tract), with the log of the expected number of cases in the respective unit (derived as outlined above) included as an offset. The offset in effect adjusts for demographic differences between geographic units. The risk estimates generated from these models are based on the comparison of observed to expected cases in a given geographic unit, and are conventionally referred to as relative risks. If the relative risk is greater than 1, the disease is more likely to occur than expected, and if the relative risk is less than 1, the disease is less likely to occur than expected.

We used the SPC model to evaluate the presence of spatial autocorrelation, i.e. the presence of a significant spatial patterning of disease risk over space. The SPC model includes both spatially correlated and uncorrelated random effects to explain the possible spatially correlated variations and heterogeneous patterns in the residuals. By comparing the model fit (based on DIC) between the general models (including spatial autocorrelation) with one containing only the uncorrelated random effects we tested for the presence of spatial autocorrelation. A difference of DIC greater or equal to 2 is commonly considered evidence of a significant improvement in fit, hence would be considered evidence for spatial auto-correlation if that model decreased the DIC by such an amount (Lawson, 2008, ch 5; Spiegelhalter et al., 2002). Models were fit by region and for each disease separately, first for the total population and then stratified by race.

To evaluate the joint spatial correlation (i.e. cross-correlation) between type 1 and type 2 diabetes, we calculated an empirical correlation between the RR estimates obtained for SPC models for type 1 and type 2 using the Pearson correlation coefficient. In addition, the SPMCAR model was used to provide a modeled estimate of the joint spatial correlation of type 1 and type 2 diabetes. (Gelfand and Vounatsou, 2003) The main advantage of this model is that it estimates the correlations between spatially correlated random effects of multivariate outcomes. All models were fit using WinBUGS (version 1.4.3). A detailed description of the statistical properties of the SPC and SPMCAR models is presented in Appendix 1 for the more technically inclined reader.

Results

Table 1 includes geographic and demographic characteristics of the four study regions. Colorado had the largest geographic expanse and the largest number of counties and census tracts. Colorado and South Carolina were similar in terms of total population. The Ohio and

Washington region covered a markedly smaller land area and populations. There was substantial variation in terms of race/ethnic composition of the population under surveillance and in terms of the number of incident cases between the regions. Given the very small number of cases in some of the minority population groups, all subsequent tables were limited to non-Hispanic whites in all regions, African Americans in South Carolina, and Hispanics in Colorado.

Table 1 also illustrates the distribution of cases across geographic units of analysis. More than 70% of counties in South Carolina contained one or more case of type 1 or type 2, while in Colorado the respective proportions were 49% and 21%. Because of the small number of counties covered in the Ohio and Washington study region, spatial analyses could not be conducted in a Bayesian framework in those regions and corresponding data are not shown, As expected, data were markedly sparser at smaller geographic units such as census tracts. Across regions, only 20% of tracts contained one or more cases of type 1 and 8% contained at least one case of type 2.

Both type 1 and type 2 incidence rates varied between study regions (Table 2). The highest crude incidence rate of type 1 among the 10-14 year old non-Hispanic whites was observed in Ohio and Colorado, followed by Washington and then South Carolina. Incidence rates between the latter two and former two regions differed significantly by a factor of 1.5. In the 15-19 year non-Hispanic white group, Washington youth had the highest rates, Ohio and Colorado had very similar rates. South Carolina youth had the lowest rates of Type 1, by a factor of 2 or more, differing significantly from the rates in the other three regions. With respect to incidence of type 2 diabetes non-Hispanic white youth, the Ohio region had the highest type 2 diabetes incidence rates in both age groups and Colorado had the lowest rates. In fact, Colorado type 2 rates were significantly lower than all other regions for both age groups.

Table 3 presents the standardized incidence ratios (SIR) by study region and race/ethnicity which are simply the ratio of the total number of observed cases over the expected cases. The observed and expected cases shown here form the basis of all further statistical analyses including the model based estimation. Given that the SEARCH incidence rates obtained from all six SEARCH centers were used in the calculation of the expected, the SIR shown here indicate to what extent a given region's incidence was lower than, higher than, or similar to the overall SEARCH rate. Differences between regions in terms of age, sex or race composition of the population have been taken into account via the standardization process.

Figure 1 shows the spatially smoothed relative risks of type 1 diabetes estimated from the SPC model at the tract level for each of the four study regions. In all regions some areas with type 1 relative risks of 2.0 or higher were observed. Our test for spatial autocorrelation was significant within both Colorado and South Carolina focusing on the total populations as the difference in DIC was markedly higher than the generally used criteria of ≥ 2 (DIC difference in Colorado =5.8, South Carolina =8.4). The visual inspection of the Colorado and South Carolina maps suggests that the spatial auto-correlation was due to localized clustering of high risk tracts in the North-East of Colorado and in eastern South Carolina. Our conclusions were confirmed in race-stratified analyses as significant spatial autocorrelation was observed for South Carolina type 1 in whites (DIC difference=5.3), African Americans (DIC difference=3.0), Colorado type 1 whites (DIC difference=4.4), Hispanics (DIC difference=3.1). In Ohio, no evidence for spatial autocorrelation of type 1 diabetes was found for the aggregated total population (difference in DIC=0.54) but it emerged for the White population (DIC difference=6.5) but not for African Americans (DIC difference= 1.5). In Washington, we observed no evidence for spatial autocorrelation of type 1 diabetes (DIC difference =0.3 for total population, 0.4 for whites).

Figure 2 shows the spatially smoothed relative risks of type 2 diabetes. The sparseness of type 2 data is striking in the maps, especially in the Colorado, the Washington and the Ohio study regions. Our analyses suggested presence of spatial autocorrelation within type 2 diabetes in Washington (difference in DIC in total population =3.2, in whites =4.2) and in the Colorado total population (difference in DIC=2.7) and Hispanics (DIC difference=2.5) but not in whites (DIC difference 0.3). In Ohio no spatial auto-correlation was observed for type 2 for the total population (DIC difference =1.0) or for whites (DIC difference=1.4) but was present for the African American population DIC difference=2.8). In South Carolina, spatial autocorrelation was observed for type 2 in whites (DIC difference=3.8) and African Americans (DIC difference=3.4) but not for the aggregated total population (difference in DIC=1.3) which suggests that spatial clustering effects occur in different locations, possibly due to residential segregation.

Our evaluation of cross-correlation between type 1 and type 2 diabetes is summarized in Table 4. Because of the sparseness of our data, we present results at the level of the county and subsequently at the level of the Census tract. Because of the small number of counties included in the Ohio and Washington region, analyses were limited to South Carolina and Colorado at the county level. In a first step, the county-specific SIR for type 1 and type 2 diabetes were correlated empirically. The resulting Pearson's correlation coefficients varied from very small and non-significant for the total population (r_{SC}= 0.05, p=0.76; $r_{CO=}$ 0.04, p=0.75) to moderately large and statistically significant for race-specific analyses (non-Hispanic whites r_{SC}= 0.34, p=0.02; Hispanics $r_{CO=}$ 0.33, p=0.01). In the next step, the spatially smoothed relative risk estimates from the separate SPC models for each disease were correlated empirically. Point estimates of correlations for all groups and regions were moderate to strong, with a strong cross-correlation of type 1 and type 2 diabetes observed in the Colorado Hispanic group ($r_{CO=}$ 0.76, p=<0.0001). In a final step, we estimated the cross-correlation using the MCAR model which contains an explicit cross-correlation parameter. Consistent with the correlations of the SPC models, point estimates were moderate to strong for most groups, reaching statistical significance for South Carolina non-Hispanic whites (r_{SC}= 0.31, CI 0.09, 0.50) and Colorado non-Hispanic whites ($r_{CO=}$ 0.40, CI 0.31,0.47) and Hispanics ($r_{CO=}$ 0.72, CI 0.66,0.86).

Analyses were repeated at the tract level, now including all four study regions. The resulting Pearson's correlation coefficients were very small and non-significant (r_{SC}= -0.04, p=0.30; $r_{CO=}$ -0.04, p=0.23; $r_{OH=}$ -0.01, p=0.91; $r_{WA=}$ 0.01, p=0.90) which was similarly true for the race-specific analyses. Very similar results in terms of magnitude and level of significance were obtained by correlating the smoothed relative risk estimates obtained from separate SPC models at the tract level for each disease. One exception was Washington where the correlation became significant. However, the magnitude was still very small. In a final step, we estimated spatial correlation parameters using the MCAR model and consistent with all previous tract-level findings, correlation estimates were very close to zero (r_{SC}= -0.02, 95%CI -0.03, 0.00001; $r_{CO=}$ -0.02, 95%CI -0.03,-0.005; $r_{OH=}$ 0.03, 95%CI 0.003,0.05; $r_{WA=}$ 0.09, 95%CI 0.06,0.12).

Discussion

Worldwide, numerous studies have described geographic variation in type 1 diabetes incidence (Feltbower et al., 2005; Rytkonen et al., 2001; Samuelsson et al., 2007; Waldhor et al., 2003; Cardwell et al., 2007; Karvonen et al., 1993; Karvonen et al., 2000A), however less is known about geographic variation in the United States. Just within the Western European countries, there is a more than sevenfold difference in incidence, which is all the more noteworthy as the geographic expanse of Western Europe falls easily within North America. Similar to other countries, our study shows marked variation in incidence of type 1 diabetes as the region specific incidence rates of type 1 were quite different. However, unlike some of the

international studies conducted in Austria, China, or Germany, we did not observe a clear North-South gradient, based on comparison of regional incidence rates (Waldhor et al., 2000; Yang et al., 1998; Rosenbauer et al., 1999).

With respect to geographic variation in diabetes incidence in the United States, it has been suggested - though not systematically explored - that regional variation in type 1 diabetes may be explained to a large extent by variation in the racial/ethnic composition of the populations (LaPorte et al., 1995). Comparison of incidence rates from existing U.S. diabetes registries indicated some differences in ethnic-specific rates across the differing registry locations (Libman et al., 1998; Lipman et al., 2006; Lipton et al., 2002), but interpretation was constrained by differing ascertainment systems, differing case definitions of type 1 diabetes, age groups, time periods and - with few exceptions - fairly restricted geographical regions studied (Dorman et al., 1995). We were able to overcome these methodological issues by using data from a uniform-population-based surveillance system, the SEARCH study (SEARCH Study Group, 2004).

Our race/ethnic stratified incidence estimates suggest that there are indeed marked regional differences in diabetes incidence that are unrelated to the ethnic composition of the underlying populations. For instance, while South Carolina's type 1 incidence rates were significantly lower than those of the other sites, Colorado's type 2 incidence rates were lower than all others. These findings may be of importance to public health agencies estimating the local burden of disease or conducting localized type 1 or type 2 cluster investigations.

Type 1 and type 2 diabetes have been viewed for a long time as having very distinct etiologies (Dorman et al., 1995). Type 1 diabetes is thought to result from beta cell loss, being characterized as an autoimmune disorder with acute onset, measurable autoantibodies, insulin dependency and affecting predominantly young people. Type 2 diabetes is thought to result from a combination of insulin resistance and insufficient insulin response, being characterized by a slow onset, not typically requiring insulin therapy, and affecting predominantly adults. The emergence of type 2 diabetes in youth has prompted the development of a new theory of diabetes development, the accelerator hypothesis, which suggests that type 1 and type 2 diabetes share a common etiology (Wilkin 2001; Wilkin 2008). This hypothesis is heavily debated, with evidence emerging on both sides (Dabelea et al. 2006; Knerr et al. 2005). A comprehensive review and criticism of the accelerator hypothesis was recently published (Fourlanos et al., 2008; Gale 2007).

By utilizing methods employed in spatial epidemiology, our study aimed to add to this discussion with data obtained at an aggregate, population level. Working from the hypothesis that areas with a high incidence of type 1 diabetes might also exhibit a high incidence of type 2 diabetes among youth, we examined the cross-correlation, or joint spatial correlation between these diseases at two geographic units, the county and the Census tract. At the level of the larger geographic unit, our study suggests a moderate cross-correlation between type 1 and type 2 diabetes in non-Hispanic whites in South Carolina and in Colorado, and a strong correlation in Hispanics in Colorado. However, at the level of the Census tract, our study findings are largely negative. This was true for all four distinct regions studied, including in subpopulations defined by race/ethnicity. Thus at the level of shared geographic patterning, our study does not provide support for a common etiology shared between type 1 and type 2 diabetes.

The discrepancy between our county and tract level analyses may be due to aggregation bias. It has been shown that the choice of the spatial unit has the potential to influence the results (Morris and Munasinghe, 1993), known as the modifiable spatial area unit problem. The spatial aggregation of data tends to increase spatial correlation between units of observation. Once

data are aggregated, they often show different characteristics from individual level data, and the lack of control for individual level confounders is a source of bias. This view would suggest that the census tract level analyses present the less biased results. On the other hand, it is conceivable that the lack of a joint spatial correlation between type 1 and type 2 diabetes at the tract level is largely due to the immense sparseness of our data. Even though the Bayesian models are inherently poised to deal with data sparseness, the extreme number of tracts without cases might have overpowered any effects. In the future, we may be able to evaluate this aspect of our study by utilizing more than two years of incidence data.

Our study furthermore explores small area geographic variation within each diabetes type. Our results suggest the presence of spatial autocorrelation of type 1 diabetes in Colorado, Ohio, and South Carolina with localized clustering of high risk tracts. It has been suggested that environmental factors such as infections or viruses may contribute to this localized clustering phenomenon of type 1, which has been shown in previous, albeit conflicting findings to be associated inversely with population density, urbanization, crowding and deprivation (Cardwell et al., 2006; Cardwell et al., 2007; Rytkonen et al., 2003; Patterson et al., 1996; Staines et al., 1997). Our study also presents evidence for the presence of spatial auto-correlation within type 2 diabetes in Colorado, Ohio, South Carolina and the Washington region. It is likely that the processes responsible for small area clustering of pediatric type 2 diabetes are very different than those for type 1 diabetes and related to low socioeconomic status, high body mass index, unhealthful dietary intake and physical inactivity. Thus, the literature on environmental correlates of diabetes seems to be consistent with the perspective that while small area variation and clustering of both types of diabetes are to expected, there would be little reason to believe that they would lead to a joint spatial pattern, i.e. cross-correlation between both diabetes types.

There are a number of limitations to our study. First, our study utilized only 2 years of incidence data, whereas other studies of geographic variation have typically utilized at least a decade's worth of incidence data if not more (Feltbower et al., 2005). The sparseness was also a function of the spatial unit chosen, particularly for the census tract. Second, we used an imputation technique to allocate persons to a tract within a known zip code. To evaluate the impact of this approach we repeated our analyses without allocation and removing these individuals with no impact on the data. A further limitation of our data is that we had to rely on contact addresses, which in some instances are not identical to the actual residence. Possibly even more important is the fact that the contact address may or may not have been the address at the relevant temporal exposure period, prior to diagnosis when exposure to potential risk factors common to type 1 and type 2 diabetes could have occurred.

A key strength of our study is that we were able to evaluate geographic variation in four distinct geographic regions including various race/ethnic groups. Our study used Bayesian methods that address several of the inherent challenges in studying geographic variation of rare diseases. Bayesian methods tap into the recognition that neighboring areas can be used in the estimation of each specific area's rates (Lawson 2008). The resulting area-specific rate estimates are thus smoothed or shrunken, with the amount of shrinkage being larger if the confidence in an area's observed rate is lower, which is generally the case for less densely populated areas. A number of studies of type 1 diabetes have utilized these techniques to address a variety of question including questions of spatial variation, space-time variation, and ecologic analyses correlating population characteristics with diabetes incidence (Cardwell et al., 2007; Casu et al. 2004a; Casu et al., 2004b; duPrel et al. 2007; Feltbower et al., 2005; Moltchanov et al., 2005; Rytkonen et al. 2001; Rytkonen et al. 2003; Thomas et al., 2008). A key strengths of a Bayesian modeling approach lies in the ability to model flexibly the disease distribution. In addition, it allows the use of special spatial random effect terms that are not available in other approaches. The main

limitation of the approach is that it assumes a parametric form for the distribution of disease. However distributional assumptions can easily be tested via sensitivity analysis.

Our study utilized Bayesian methods in evaluating the joint spatial correlation (cross-correlation) of type 1 and type 2 diabetes in youth, initially modeling each disease separately and then jointly in a multivariate SPMCAR model. They key advantage of the MCAR model is that information from multiple diseases is used to improve the estimation of incidence of each individual disease. However the amount of information that is "borrowed" across diseases depends on the amount of correlation. In addition to our own work, only very few other examples of estimated joint spatial correlation exist in the epidemiologic literature (Assuncao and Castro, 2004; Feltbower et al., 2005; Thompson et al., 2007). A study of acute lymphoblastic leukemia and type 1 diabetes found some suggestion of spatial correlation (r=0.33, 95% credible interval -0.20, 0.74) in Yorkshire (Feltbower et al., 2005). A further study evaluated a variety of childhood cancer histotypes finding moderate to high correlations (r >= 0.7) between most histotypes except when correlations with osteosarcoma were considered where correlations were markedly lower (r=0.35-0.43) (Thompson et al., 2007).

In summary, our study provides evidence for regional variation in type 1 and type 2 incidence among youth aged 10-19 years in 2002-2003 that cannot be attributed to race/ethnic differences in the underlying populations. We furthermore identified small-area, localized spatial autocorrelation within type 1 and type 2 diabetes. Both the regional data and the small area results may be important for public health agencies and future surveillance efforts for diabetes. In addition, we investigated the joint spatial patterning between type 1 and type 2 diabetes. Our results suggest that at the level of the Census tract which we consider the most unbiased level of analysis, there is currently no evidence for joint spatial patterning.

Acknowledgments

We would like to thank the SEARCH investigators, staff and participants for making this project possible.

Grant Support: The project was supported by Award Number R01DK077131 from the National Institute Of Diabetes And Digestive And Kidney Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of Diabetes And Digestive And Kidney Diseases or the National Institutes of Health

Appendix 1. Technical description of Bayesian spatial models

The Sparse Poisson Convolution (SPC) model is a mixture of two Poisson distributions of zero and non-zero observed counts and includes factored intercepts for modeling these counts.

$$\Pr(y_i)=Pois(y_i,\mu_i)$$
$$\log(\mu_i)=\log(E_i)+\alpha(j)+u_i+v_i,\ j=1,2$$

where, y_i is the observed count in the ith census tract; j is the binary factor indicating zero and non-zero observed counts (j=1 if y_i=0 and j=2 if y_i>0); $\alpha(j)$ is the factored intercept for modeling zero and non-zero counts; u_i is the correlated random effect; and v_i is the uncorrelated random effect. The expected counts were used as an offset, which is an effect observed in each area but which is regarded as a constant in the analysis. Our SPC model included spatially correlated and uncorrelated random effects to explain the possible spatially correlated variations and heterogeneous patterns in the residuals. The spatially correlated random effects, which model the extra variation that is correlated over space, are modeled by the conditional autoregressive (CAR) model (Besag and Kooperberg 1995; Besag, York, and Molliq 1991) which estimates the random effect of the i^{th} geographic unit (here: Census tract) (u_i) conditional on the sum of

the weighted adjacent geographic units (i.e. Census tracts) values. Spatial correlation is the CAR model is controlled by neighboring units with common boundaries (adjacencies). Each of the random effects is assumed to have an underlying Gaussian prior distribution controlled by a variance parameter, one for uncorrelated and one for correlated effects. These variance parameters can also have non-informative (half-Cauchy) prior distributions.

We furthermore extended the analysis to a model with multivariate outcomes, the Sparse Poisson MCAR model (SPMCAR) (Gelfand and Vounatsou, 2003). In this case, the intercept is applied to multiple disease models. The main advantage of this model is that it estimates the correlations between spatially correlated random effects of multivariate outcomes. Instead of considering individual diseases, here we consider the two diseases (type 1 and type 2) as components of a vector of outcomes and apply a multivariate model. The SPMCAR model is presented below:

$$\Pr(Y_i)=Pois(Y_i,\mu_i)$$
$$\log(\boldsymbol{\mu}_i)=\log(\boldsymbol{E}_i)+\alpha(j)+\underline{U}_i+V_i,\ j=1,2.$$

where, multivariate outcome Y_i follows a Poisson distribution with the mean μ_i; μ_i is a vector of the means of the Poisson distribution of the 'm' multivariate health outcomes ($\boldsymbol{\mu_i}$ =(μ_{i1},…, μ_{im})); $\boldsymbol{\alpha(j)}$ is a vector of factored intercepts (= $\alpha_1(j)$,…, $\alpha_m(j)$) where j denotes the class of the observed count (zero or positive); U_i is a vector of correlated random effects (U_i = (u_1,…, u_m)); and V_i is a vector of uncorrelated random effects (V_i = (v_1,…, v_m)). Since the RR of type 1 and type 2 are estimated within the model at the same time, we can estimate the correlation of RR of type 1 and type 2 within the model, which is the parameter of interest in this study.

All models were fit using WinBUGS (version 1.4.3) which can be used to provide posterior samples of parameter values from Bayesian models. Samples were generated based on the Gibbs sampling, and we ran 3 different chains of parameter values to check the sensitivity to the initial values. Convergence was assessed using the Brooks-Gelman-Rubin diagnostics (Brooks and Gelman, 1998) as well as visual inspection of trace plots. 10,000 samples were generated from which the first 8,000 samples were discarded, the remaining 2,000 samples were used to summarize the posterior estimates. For prior distributions, the normal prior distribution (N(0,1000)) was assigned to α, and uniform prior distributions are assigned to σ_u and σ_v (σ_u, σ_v ~*Unif* (0,10),).

References

Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat 1998;52:119–126.

Assuncao RM, Castro MS. Multiple cancer sites incidence rates estimation using a multivariate Bayesian model. Int J Epidemiol 2004;33(3):508–516. [PubMed: 15044411]

Besag J, Kooperberg C. On conditional and intrinsic autoregressions. Biometrika 1995;82(4):733–746.

Besag J, York J, Molliq A. Bayesian image restoration, with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics 1991;43(1):1–20.

Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics 1998;7:434–455.

Cadario F, Vercellotti A, Trada M, Zaffaroni M, Rapa A, Iafusco D, Salardi S, Baldelli R, Bona G. Younger age at diagnosis of type 1 diabetes mellitus in children of immigrated families born in Italy. J Endocrinol Invest 2004;27(10):913–918. [PubMed: 15762037]

Cardwell CR, Carson DJ, Patterson CC. Higher incidence of childhood-onset type 1 diabetes mellitus in remote areas: a UK regional small-area analysis. Diabetologia 2006;49:2074–2077. [PubMed: 16868747]

Cardwell CR, Carson DJ, Patterson CC. Secular trends, disease maps and ecological analyses of the incidence of childhood onset Type 1 diabetes in Northern Ireland, 1989-2003. Diabetic Medicine 2007;24(3):289–295. [PubMed: 17305789]

Casu A, Pascutto C, Bernardinelli L, Songini M. Bayesian approach to study the temporal trend and the geographical variation in the risk of type 1 diabetes. The Sardinian Conscript Type 1 Diabetes Registry. Pediatr Diabetes 2004a;5(1):32–38. [PubMed: 15043688]

Casu A, Pascutto C, Bernardinelli L, Songini M. Type 1 diabetes among Sardinian children is increasing: the Sardinian diabetes register for children aged 0-14 years (1989-1999). Diabetes Care 2004b;27 (7):1623–1629. [PubMed: 15220238]

Cayo MR, Talbot TO. Positional error in automated geocoding of residential addresses. Int J Health Geogr 2003;2(10)

Census Bureau (US). Census 2000: Summary File 1 (SF1). Washington: Census Bureau (US); 2000.

Dabelea D, Bell RA, D'Agostino RB Jr, Imperatore G, Johansen JM, Linder B, Liu LL, Loots B, Marcovina S, Mayer-Davis EJ, Pettitt DJ, Waitzfelder B. Incidence of diabetes in youth in the United States. JAMA 2007;297(24):2716–2724. [PubMed: 17595272]

Dabelea D, D'Agostino RB Jr, Mayer-Davis EJ, Pettitt DJ, Imperatore G, Dolan LM, Pihoker C, Hillier TA, Marcovina SM, Linder B, Ruggiero AM, Hamman RF. Testing the accelerator hypothesis: body size, beta-cell function, and age at onset of Type 1 (autoimmune) diabetes. Diabetes Care. 2006

Dorman, JS.; McCarthy, BJ.; O'Leary, LA.; Koehler, AN. Diabetes in America. 2. National Diabetes Data Group of the National Institute of Diabetes and Digestive and Kidney Diseases; Bethesda, MD: 1995. Risk Factors for Insulin-Dependent Diabtes; p. 165-178.

du Prel JB, Icks A, Grabert M, Holl RW, Giani G, Rosenbauer J. Socioeconomic conditions and type 1 diabetes in childhood in North Rhine-Westphalia, Germany. Diabetologia. 2007

ESRI. ArcGIS 9.3. Redlands, CA: Environmental Systems Research Institute (ESRI); 2008.

Feltbower RG, Manda SO, Gilthorpe MS, Greaves MF, Parslow RC, Kinsey SE, Bodansky HJ, McKinney PA. Detecting small-area similarities in the epidemiology of childhood acute lymphoblastic leukemia and diabetes mellitus, type 1: a Bayesian approach. Am J Epidemiol 2005;161(12):1168–1180. [PubMed: 15937026]

Fourlanos S, Harrison LC, Colman PG. The accelerator hypothesis and increasing incidence of type 1 diabetes. Curr Opin Endocrinol Diabetes Obes 2008;15(4):321–325. [PubMed: 18594271]

Gale EA. To boldly go--or to go too boldly? The accelerator hypothesis revisited. Diabetologia 2007;50 (8):1571–1575. [PubMed: 17579829]

Gelfand AE, Vounatsou P. Proper multivariate conditional autoregressive models for spatial data analysis. Biostatistics 2003;4(1):11–25. [PubMed: 12925327]

Henry KA, Boscoe FP. Estimating the accuracy of geographical imputation. International Journal of Health Geographics 2008;7(1):3. [PubMed: 18215308]

Hibbert JD, Liese AD, Lawson A, Porter DE, Puett RC, Standiford D, Liu L, Dabelea D. Evaluating geographic imputation approaches for zip code level data: an application to a study of pediatric diabetes. Int J Health Geogr 2009 Oct 8;8(54)

Karvonen M, Tuomilehto J, Libman I, LaPorte R. A review of the recent epidemiological data on the worldwide incidence of type 1 (insulin-dependent) diabetes mellitus. World Health Organization DIAMOND Project Group. Diabetologia 1993;36(no 10):883–892. [PubMed: 8243865]

Karvonen M, Viik-Kajander M, Moltchanova E, Libman I, LaPorte R, Tuomilehto J. Incidence of childhood type 1 diabetes worldwide. Diabetes Mondiale (DiaMond) Project Group. Diabetes Care 2000;23:1516–1526. [PubMed: 11023146]

Knerr I, Wolf J, Reinehr T, Stachow R, Grabert M, Schober E, Rascher W, Holl RW. The 'accelerator hypothesis': relationship between weight, height, body mass index and age at diagnosis in a large cohort of 9,248 German and Austrian children with type 1 diabetes mellitus. Diabetologia 2005;48 (12):2501–2504. [PubMed: 16283240]

LaPorte, RE.; Matsushima, M.; Chang, Y. Diabetes in America. 2. National Diabetes Data Group of the National Institute of Diabetes and Digestive and Kidney Diseases; Bethesda, MD: 1995. Prevalence and Incidence of Insulin-Dependent Diabetes; p. 37-46.

Lawson, AB. Bayesian Disease Mapping: Hierarchical modeling in Spatial Epidemiology. CRC Press; New York, NJ: 2008.

Libman IM, LaPorte RE, Becker D, Dorman JS, Drash AL, Kuller L. Was there an epidemic of diabetes in nonwhite adolescents in Allegheny County, Pennsylvania? Diabetes Care 1998;21(8):1278–1281. [PubMed: 9702433]

Liese AD, Liu L, Davis C, Standiford D, Waitzfelder B, Dabelea D, Bell R, Williams D, Imperatore G, Lawrence JM. Participation in pediatric epidemiologic research: the SEARCH for Diabetes in Youth Study experience. Contemp Clin Trials 2008;29(6):829–836. [PubMed: 18573350]

Lipman TH, Jawad AF, Murphy KM, Tuttle A, Thompson RL, Ratcliffe SJ, Levitt Katz LE. Incidence of type 1 diabetes in Philadelphia is higher in black than white children from 1995 to 1999: epidemic or misclassification? Diabetes Care 2006;29(11):2391–2395. [PubMed: 17065673]

Lipton R, Keenan H, Onyemere KU, Freels S. Incidence and onset features of diabetes in African-American and Latino children in Chicago, 1985-1994. Diabetes Metab Res Rev 2002;18(2):135–142. [PubMed: 11994905]

Moltchanova E, Penttinen A, Karvonen M. A hierarchical Bayesian birth cohort analysis from incomplete registry data: evaluating the trends in the age of onset of insulin-dependent diabetes mellitus (T1DM). Stat Med 2005;24(19):2989–3004. [PubMed: 16149124]

Morris RD, Munasinghe RL. Aggregation of Existing Geographic Regions to Diminish Spurious Variability of Disease Rates. Statistics in Medicine 1993;12:1915. [PubMed: 8272670]

Onkamo P, Vaananen S, Karvonen M, Tuomilehto J. Worldwide increase in incidence of type I diabetes-the analysis of the data on published incidence trends. Diabetologia 1999;42(12):1395–1403. [PubMed: 10651256]

Parslow RC, McKinney PA, Law GR, Bodansky HJ. Population mixing and childhood diabetes. Int J Epidemiol 2001;30(3):533–538. [PubMed: 11416078]

Patterson CC, Carson DJ, Hadden DR. Epidemiology of childhood IDDM in Northern Ireland 1989-1994: low incidence in areas with highest population density and most household crowding. Northern Ireland Diabetes Study Group. Diabetologia 1996;39:1063–1069. [PubMed: 8877290]

Pihoker C, Scott CR, Lensing SY, Cradock MM, Smith J. Non-insulin dependent diabetes mellitus in African-American youths of Arkansas. Clin Pediatr (Phila) 1998;37(2):97–102. [PubMed: 9492117]

Pinhas-Hamiel O, Dolan LM, Daniels SR, Standiford D, Khoury PR, Zeitler P. Increased incidence of non-insulin-dependent diabetes mellitus among adolescents. J Pediatr 1996;128(5):608–615. [PubMed: 8627431]

Rosenbauer J, Herzig P, von KR, Neu A, Giani G. Temporal, seasonal, and geographical incidence patterns of type I diabetes mellitus in children under 5 years of age in Germany. Diabetologia 1999;42 (9):1055–1059. [PubMed: 10447515]

Rytkonen M, Moltchanova E, Ranta J, Taskinen O, Tuomilehto J, Karvonen M. The incidence of type 1 diabetes among children in Finland - rural-urban difference. Health Place 2003;9:315–325. [PubMed: 14499216]

Rytkonen M, Ranta J, Tuomilehto J, Karvonen M. Bayesian analysis of geographical variation in the incidence of Type I diabetes in Finland. Diabetologia 2001;44(3):B37–B44. [PubMed: 11724415]

Samuelsson U, Carstensen J, Lofman O, Nordfeldt S. Seasonal variation in the diagnosis of type 1 diabetes in south-east Sweden. Diabetes Research and Clinical Practice 2007;76:75–81. [PubMed: 16963147]

Scott CR, Smith JM, Cradock MM, Pihoker C. Characteristics of youth-onset noninsulin-dependent diabetes mellitus and insulin-dependent diabetes mellitus at diagnosis. Pediatrics 1997;100(1):84–91. [PubMed: 9200365]

SEARCH Study Group. SEARCH for Diabetes in Youth: a multicenter study of the prevalence, incidence and classification of diabetes mellitus in youth. Control Clin Trials 2004;25(5):458–471. [PubMed: 15465616]

Song HR, Liese AD, Lawson AB, D'Agostino R Jr. Modeling type 1 and type 2 diabetes mellitus incidence in youth: an application of Bayesian hierarchical regression for sparse small area data. Spatial an Spatio-temporal Epidemiology. 2008 in review.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B 2002;64(4):583–639.

Staines A, Bodansky HJ, McKinney PA, Alexander FE, McNally RJ, Law GR. Small area variation in the incidence of childhood insulin-dependent diabetes mellitus in Yorkshire, UK: links with overcrowding and population density. Int J Epidemiol 1997;26:1307–1313. [PubMed: 9447411]

Thomas W, Birgit R, Edith S. Changing geographical distribution of diabetes mellitus type 1 incidence in Austrian children 1989—2005. Eur J Epidemiol 2008;23(3):213–218. [PubMed: 18210201]

Thompson JA, Carozza SE, Zhu L. An evaluation of spatial and multivariate covariance among childhood cancer histotypes in Texas (United States). Cancer Causes and Control 2007;18(1):105–113. [PubMed: 17186426]

U.S.Census Bureau. Census 2000 ZIP code tabulation areas technical documentation. Census 2000 2000:1–22.

Waldhor T, Schober E, Karimian-Teherani D, Rami B. Regional differences and temporal incidence trend of Type I diabetes mellitus in Austria from 1989 to 1999: a nationwide study. Diabetologia 2000;43 (11):1449–1450. [PubMed: 11126417]

Waldhor T, Schober E, Rami B. Regional distribution of risk for childhood diabetes in Austria and possible association with body mass index. Eur J Pediatr 2003;162:380–384. [PubMed: 12756559]

Wilkin TJ. The accelerator hypothesis: weight gain as the missing link between type I and type II diabetes. Diabetologia 2001;44(7):914–922. [PubMed: 11508279]

Wilkin TJ. Diabetes: 1 and 2, or one and the same? Progress with the accelerator hypothesis. Pediatr Diabetes 2008;9(3):23–32. [PubMed: 18540866]

Yang Z, Wang K, Li T, Sun W, Li Y, Chang YF, Dorman JS, LaPorte RE. Childhood diabetes in China. Enormous variation by place and ethnic group. Diabetes Care 1998;21(4):525–529. [PubMed: 9571336]

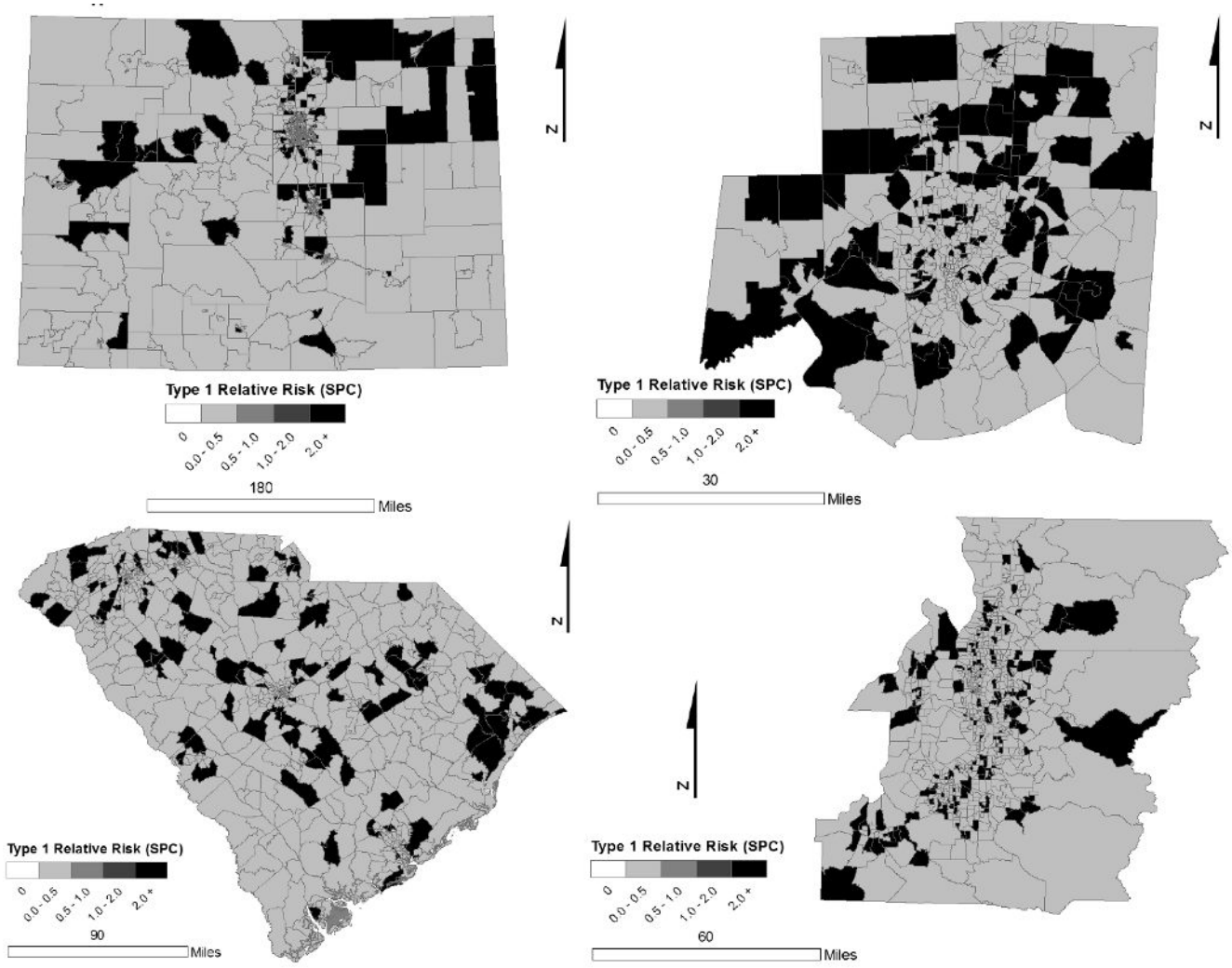

Figure 1.
Spatially smoothed relative risks of Type 1 diabetes mellitus 2002-2003 in 10-19 year olds in four geographic regions computed from Bayesian SPC model (Clockwise from upper left: Colorado, Ohio, Washington, South Carolina)

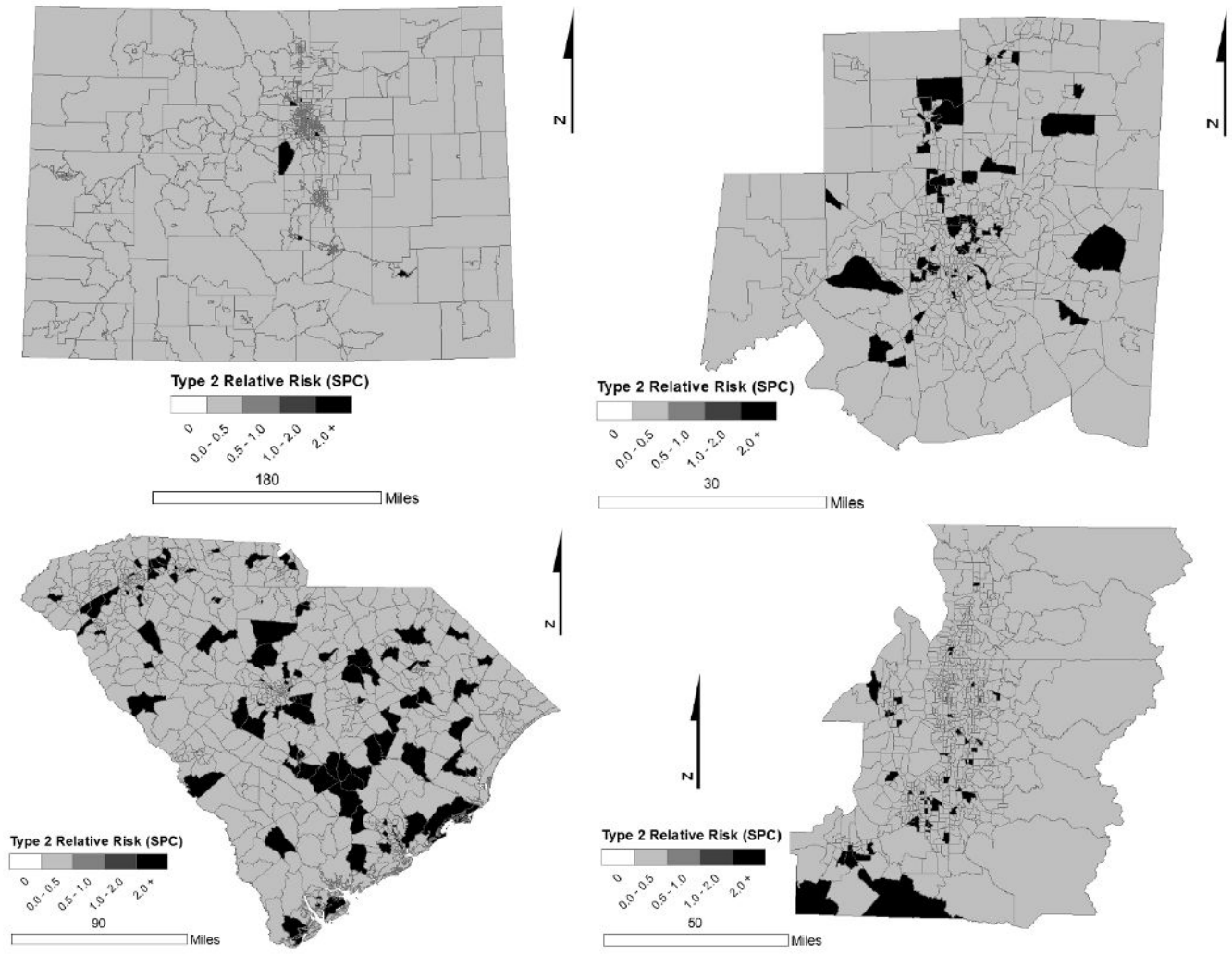

Figure 2.
Spatially smoothed relative risks of Type 2 diabetes mellitus 2002-2003 in 10-19 year olds in four geographic regions computed from Bayesian SPC model Type 2 SPC (Clockwise from upper left: Colorado, Ohio, Washington, South Carolina)

Table 1

Geographic and demographic characteristics of the study regions

	Study Region			
	South Carolina	**Colorado**	**Ohio[a]**	**Washington[b]**
Land area (square miles)	30,109	103,718	2,636	7,335
Total population in 2000 (number)	4,012,012	4,301,261	1,886,650	3,483,202
Population 10-19 years in 2000 All	585,856	618,735	280,222	483,598
Non-Hispanic White (%)	58.2	68.8	81.3	71.6
African American (%)	36.6	4.2	14.6	5.6
Hispanic (%)	2.7	21.4	1.3	6.4
Asian/Pacific Islander (%)	0.9	2.2	1.0	8.8
American Indian/Alaska Native (%)	0.4	0.8	0.2	1.4
Multiple and other (%)	1.2	2.6	1.6	6.2
Type 1 Cases in 2002-2003 (number)*				
All	170	282	145	210
Non-Hispanic White	118	232	127	175
African American	44	8	13	11
Hispanic	4	37	2	7
Asian/Pacific Islander	0	0	0	7
American Indian/Alaska Native	1	1	0	2
Multiple and other	3	4	3	8
Type 2 Cases in 2002-2003 (number)*				
All	130	51	69	63
Non-Hispanic Whlite	35	13	37	38
African American	89	5	28	8
Hispanic	3	28	1	8
Asian/Pacific Islander	1	3	0	7
American Indian/Alaska Native	1	2	0	1
Multiple and Other	1	0	3	1
Counties (number)	46	63	8	5

	Study Region			
	South Carolina	**Colorado**	**Ohio[a]**	**Washington[b]**
Counties with 1 or more type 1 cases (%)	71.7	49.2	-	-
Counties with 1 or more type 2 cases (%)	73.9	20.6	-	-
Census tracts (number)	867	1062	460	749
Census tracts with 1 or more type 1 cases (%)	16.4	20.1	24.3	20.2
Census tracts with 1 or more type 2 cases (%)	12.2	4.2	11.7	6.4

*Note: The total number of cases of type 1 and type 2 diabetes aged 10-19 years diagnosed in the years 2002-2003 was 1,120. The “All” race/ethnic group includes all 6 race/ethnic categories shown above.

[a]Ohio region includes Butler, Clermont, Hamilton, Warren counties in Ohio; Boone, Campbell, and Kenton counties in Kentucky; and Dearborn county in Indiana.

[b]Washington region includes 5 counties: King, Kitsap, Pierce, Snohomish and Thurston.

Table 2

Incidence rates of Type 1 and Type 2 diabetes mellitus per 100,000 by study region, age group, race/ethnicity

			Type 1 Diabetes Mellitus		Type 2 Diabetes Mellitus	
Site	**Age Group**	**Total No. of Youth**	**2002-03 Cases**	**Incidence Rate (95%CI)**	**2002-03 Cases**	**Incidence Rate (95%CI)**
SC	10-14 years					
	All	286,806	128	22.3 (18.8-26.5)	73	12.7 (10.1-16.0)
	Non-Hispanic White	168,389	93	27.6 (22.5-33.8)	17	5.1 (3.1-8.1)
	African American	108,756	32	14.7 (10.4-20.8)	53	24.4 (18.6-31.9)
	15-19 years					
	All	292,170	39	6.7 (4.9-9.1)	56	9.6 (7.4-12.4)
	Non-Hispanic White	172,834	25	7.2 (4.9-10.7)	18	5.2 (3.3-8.2)
	African American	105,578	12	5.7 (3.3-9.9)	36	17.1 (12.3-23.6)
CO	10-14 years					
	All	303,017	200	33.0 (28.7-37.9)	21	3.5 (2.3-5.3)
	Non-Hispanic White	214,932	169	39.3 (33.8-45.7)	7	1.6 (0.8-3.4)
	Hispanic	654,68	26	19.9 (13.6-29.1)	11	8.4 (4.7-15.0)
	15-19 years					
	All	299,626	78	13.0 (10.4-16.2)	30	5.0 (3.5-7.1)
	Non-Hispanic White	210,984	63	14.9 (11.7-19.1)	6	1.4 (0.7-3.1)
	Hispanic	66,616	11	8.3 (4.6-14.8)	17	12.8 (8.0-20.4)
OH	10-14 years					
	All	140,296	101	36.0 (29.6-43.7)	30	10.7 (7.5-15.3)
	Non-Hispanic White	115,264	93	40.3 (32.9-49.4)	15	6.5 (3.9-10.7)
	African American	21,601	7	16.2 (7.9-33.5)	14	32.4 (19.3-54.4)
	15-19 years					
	All	135,519	41	15.1 (11.2-20.5)	36	13.3 (9.6-18.4)
	Non-Hispanic White	112,705	34	15.1 (10.8-21.1)	22	9.8 (6.4-14.8)
	African American	19,203	6	15.6 (7.2-34.1)	14	36.5 (21.7-61.2)
WA	10-14 years					
	All	230,381	135	29.3 (24.8-34.7)	31	6.7 (4.7-9.5)

Site	Age Group	Total No. of Youth	Type 1 Diabetes Mellitus		Type 2 Diabetes Mellitus	
			2002-03 Cases	Incidence Rate (95%CI)	2002-03 Cases	Incidence Rate (95%CI)
	Non-Hispanic White	17,7526	115	32.4 (27.0-38.9)	15	4.2 (2.6-7.0)
	15-19 years					
	All	223,316	67	15.0 (11.8-19.1)	31	6.9 (4.9-9.9)
	Non-Hispanic White	168,626	60	17.8 (13.8-22.9)	23	6.8 (4.5-10.2)

Note: The total number of cases included is 1,026 because of exclusion of the “multiple and other” race/ethnic group. The “All” race/ethnic group includes 5 race group: Non-Hispanic White, Black, Hispanic, Asian/Pacific Islander and American Indian/Alaska Native.

Table 3

Standardized incidence ratios (SIRs) (2002-2003) for Type 1 and Type 2 diabetes mellitus among youth 10-19 years, by study region

Study Region	Type 1 Diabetes Mellitus			Type 2 Diabetes Mellitus		
	Obs. cases	Exp.cases	SIR	Obs. cases	Exp.cases	SIR
South Carolina						
All	165	233.74	0.71	124	126.71	0.98
Non-Hispanic white	118	162.77	0.72	33	29.34	1.12
African American	42	65.29	0.64	87	89.63	0.97
Colorado						
All	259	255.10	1.02	48	89.66	0.54
Non-Hispanic white	214	205.07	1.04	12	36.36	0.33
Hispanic	37	39.38	0.94	27	34.10	0.79
Ohio						
All	142	123.87	1.15	66	38.83	1.70
Non-Hispanic white	127	109.73	1.16	37	19.45	1.90
African American	13	12.55	1.03	28	17.07	1.64
Washington						
All	173	192.27	0.90	49	68.54	0.71
Non-Hispanic white	146	167.53	0.87	25	29.42	0.85

Note: Total number of cases included is1,026 due to exclusion of 71 cases that could not be assigned to census tract. The number of cases in remaining race categories were so small that analyses were not conducted in those race groups. The “All” race groups includes 5 race group: Non-Hispanic White, Black, Hispanic, Asian/Pacific Islander and American Indian/Alaska Native.

Table 4

Evaluation of cross-correlation between Type 1 and Type 2 diabetes mellitus among youth 10-19 years, 2002-2003, by study region and geographic unit

Study Region		Empirical Cross-Correlation				Modeled Cross-Correlation	
		SIR		SPC Model		SPMCAR Model	
		Correlation	P-value	Correlation	P-value	Correlation	95% CI
County-level analyses							
South Carolina	All	0.05	0.76	0.28	0.06	0.28	0.10, 0.40
Non-Hispanic white		0.34	0.02	0.28	0.06	0.31	0.09, 0.50
African American		-0.11	0.48	0.15	0.31	0.16	-0.03,0.33
Colorado	All	0.04	0.75	0.49	<0.0001	0.47	0.30, 0.60
Non-Hispanic white		0.09	0.49	0.42	<0.0001	0.40	0.31, 0.47
Hispanic		0.33	0.01	0.76	<0.0001	0.72	0.66, 0.86
Census travel level analyses							
South Carolina	All	-0.04	0.30	-0.01	0.81	-0.02	-0.03, 0.00001
Non-Hispanic white		0.06	0.06	0.01	0.72	0.01	-0.005, 0.03
African Amenican		-0.02	0.60	0.00	1.00	0.01	-0.01, 0.03
Colorado	All	-0.04	0.23	-0.02	0.45	-0.02	-0.03,-0.005
Non-Hispanic white		-0.02	0.55	0.01	0.76	0.01	-0.02,0.05
Hispanic		-0.01	0.76	0.00	0.96	0.00	-0.02,0.02
Ohio	All	-0.01	0.91	0.03	0.59	0.03	0.003,0.05
Non-Hispanic white		-0.001	0.98	-0.06	0.21	0.06	0.03,0.09
African Amenican		-0.009	0.85	0.05	0.33	0.06	0.05,0.11
Washington	All	0.01	0.90	0.10	0.01	0.09	0.06,0.12
Non-Hispanic white		-0.003	0.92	0.09	0.02	0.09	0.06,0.12

Note: Total number of cases in county level analyses is 618 for 2 sites. Total number of type 1 and type 2 cases is 1,026 for tract-level analyses (all 4 sites). No. of cases in remaining race categories were so small that analyses were not conducted in those race groups. The “All” race groups includes 5 race group: Non-Hispanic White, Black, Hispanic, Asian/Pacific Islander and American Indian/Alaska Native.